AHALIA SCHOOL OF ENGINEERING & TECHNOLOGY
Approved by AICTE & Affiliated to A. P. J. Abdul Kalam Technological University
ISO 9001: 2015 Certified Institution | NBA accredited B.Tech. ECE & CSE programs (till 31-12-2028)
Ahalia Campus, Palakkad I Ph: 04923 226666 | www.ahalia.ac.in

INTERNATIONAL CONFERENCE ON

ADVANCEMENTS IN SCIENCE, ENGINEERING & TECHNOLOGY

ICASET 2026 [2nd Edition]

25-26
APRIL
2026

ISBN NO:978-81-987389-3-6
978-81-987389-3-6

PROCEEDINGS

ABOUT THE INSTITUTION

Ahalia School of Engineering & Technology (ASET), Palakkad, is a premier institution committed to excellence in technical education and research. Established in 2012, the institution is approved by the All India Council for Technical Education (AICTE) and affiliated with A. P. J. Abdul Kalam Technological University, Kerala. Located in the serene Ahalia Health, Heritage and Knowledge Village, the institution provides a conducive environment for academic learning, innovation, and holistic development. ASET is an ISO 9001:2015 certified institution, reflecting its commitment to quality education and continuous improvement. The institution has been accredited by the National Assessment and Accreditation Council (NAAC) in 2023, and the B.Tech programmes in Computer Science & Engineering and Electronics & Communication Engineering are accredited by the National Board of Accreditation (NBA) from January 2026, signifying adherence to high academic and professional standards. ASET offers a range of undergraduate programmes in engineering including Civil Engineering, Computer Science and Engineering, Electronics and Communication Engineering, Electrical and Electronics Engineering, Mechanical Engineering, and emerging areas such as Artificial Intelligence and Machine Learning. The institution also promotes research, industry interaction, and innovation-driven learning. With a strong emphasis on academic excellence, ethical values, and societal responsibility, ASET strives to develop competent professionals capable of addressing global challenges and contributing to technological advancement.

ABOUT ICASET 2026

The *International Conference on Advancements in Science, Engineering & Technology (ICASET 2026)* is a premier academic event organized by Ahalia School of Engineering & Technology, Palakkad, aimed at bringing together researchers, academicians, industry professionals, and students from across the globe to share and exchange knowledge in emerging areas of science, engineering, and technology. ICASET 2026 serves as a multidisciplinary platform for presenting innovative research findings, discussing contemporary challenges, and exploring future directions in various domains of engineering and applied sciences. The conference encompasses a wide spectrum of tracks including Civil Engineering, Computer Science & Engineering, Electronics & Communication Engineering, Electrical & Electronics Engineering, Mechanical Engineering, and other interdisciplinary fields. The conference features keynote addresses by distinguished experts, technical paper presentations, and interactive sessions that foster collaboration and intellectual exchange. All submitted papers undergo a rigorous peer-review process to ensure quality and relevance, and the accepted papers included in this proceedings volume reflect significant contributions to research and innovation. ICASET 2026 aims to promote research culture, strengthen academia-industry interaction, and provide a platform for young researchers to showcase their work. The conference is expected to inspire new ideas, encourage collaborative efforts, and contribute to the advancement of science and technology for societal benefit.

MESSAGE FROM THE PRINCIPAL

It is a matter of great pride and satisfaction to present the proceedings of the ***International Conference on Advancements in Science, Engineering & Technology (ICASET 2026)*** organized by Ahalia School of Engineering & Technology.

ICASET 2026 serves as a significant platform for researchers, academicians, industry experts, and students to disseminate their findings, exchange innovative ideas, and engage in meaningful discussions on emerging trends in science, engineering, and technology. The diverse range of papers included in this volume reflects the depth of research and the commitment of the contributors towards advancing knowledge in their respective fields.

I am confident that this compilation of scholarly works will serve as a valuable reference for researchers and practitioners, and will inspire further innovations and collaborations.

I extend my heartfelt appreciation to all authors for their valuable contributions, the reviewers for their meticulous evaluation, and the organizing committee for their dedicated efforts in bringing out this proceedings volume successfully.

I congratulate everyone involved and wish ICASET 2026 continued success in the years to come.

Dr. P R SURESH
Principal
Ahalia School of Engineering & Technology
Palakkad

AHALIA SCHOOL OF ENGINEERING & TECHNOLOGY

Approved by AICTE & Affiliated to A. P. J. Abdul Kalam Technological University
ISO 9001: 2015 Certified Institution | NBA accredited B.Tech. ECE & CSE programs (till 31-12-2028)
Ahalia Campus, Palakkad I Ph: 04923 226666 | www.ahalia.ac.in

NAAC NATIONAL ASSESSMENT AND ACCREDITATION COUNCIL

NBA NATIONAL BOARD OF ACCREDITATION

ICASET 2026

INTERNATIONAL CONFERENCE ON ADVANCEMENTS IN SCIENCE, ENGINEERING & TECHNOLOGY 2026

Dr. Arvind Ajoy
Dean ICSR, IIT Palakkad

25th April 2026

09.15 AM

VISVESVARAYA HALL

INSTITUTION'S INNOVATION COUNCIL
(Ministry of HRD Initiative)

AHALIA SCHOOL OF ENGINEERING & TECHNOLOGY
Approved by AICTE & Affiliated to A. P. J. Abdul Kalam Technological University
ISO 9001: 2015 Certified Institution | NBA accredited B.Tech. ECE & CSE programs (till 31-12-2028)
Ahalia Campus, Palakkad I Ph: 04923 226666 | www.ahalia.ac.in

ICASET 2026

INTERNATIONAL CONFERENCE ON ADVANCEMENTS IN SCIENCE, ENGINEERING & TECHNOLOGY 2026

KEY NOTE SPEAKERS

Topic: "Driving Creativity for Innovation"

Nisha Nair

Clinical Associate Professor of Business Administration
University of Pittsburgh

25-04-2026, 10.00 am

Topic: "The Rise of AI Engineers: From Coders to System Thinkers".

Roji Philip

Sr. Director, AI & Automation
Speridian Technologies, Mumbai

25-04-2026, 11.00 am

Topic: "Trends in automotive electronics and embedded system"

Shamin Dudu

Section Head and Technical director
Bosch Global Software and Technologies Pvt ltd
Bangalore

25-04-2026, 11.35 am

All keynote sessions will be held in online mode via Google Meet.

Bridging Engineering and Management – Driving Creativity for Innovation

In an era defined by rapid technological advancement and increasing organizational complexity, the ability to effectively bridge engineering and management has become a critical determinant of innovation and success. Engineers are traditionally equipped with strong analytical, logical, and problem-solving skills; however, the application of these capabilities within organizational settings demands an additional layer of managerial competence. This presentation focuses on connecting engineering expertise with management practices through the lens of creativity, highlighting its role as a foundational driver of innovation.

The talk explores the transition from purely technical problem-solving to a more holistic approach that incorporates human behavior, collaboration, and leadership. While engineering emphasizes precision and efficiency, management requires the ability to influence, coordinate, and inspire teams. This intersection creates an opportunity to harness creativity as a tool for transforming technical ideas into impactful solutions. The session examines how creative thinking acts as a precursor to innovation, enabling individuals and teams to move beyond conventional approaches and generate novel, value-driven outcomes.

A key aspect of the presentation is the exploration of the creative process from a behavioral perspective. It provides insights into how ideas are generated, developed, and implemented within teams, and how organizational environments can either foster or hinder creativity. Emphasis is placed on understanding team dynamics, motivation, and communication as critical factors that shape innovative outcomes. The talk also outlines practical strategies for managers and engineers to cultivate a culture that encourages experimentation, open dialogue, and psychological safety, thereby empowering individuals to contribute creatively without fear of failure.

Furthermore, the session highlights the evolving role of engineers as not only technical experts but also facilitators of innovation and change. By adopting managerial perspectives, engineers can better align technical solutions with organizational goals, influence team behavior, and drive collective performance. The presentation underscores the importance of balancing structured processes with flexibility, ensuring that efficiency does not come at the cost of creativity.

Ultimately, this talk provides a framework for integrating engineering rigor with managerial insight to drive sustainable innovation. It encourages professionals to embrace a dual mindset—combining analytical depth with creative and behavioral awareness—to effectively lead in dynamic and competitive environments. By bridging this gap, organizations can unlock the full potential of their talent and create a strong foundation for continuous innovation and growth.

The Rise of AI Engineers – From Coders to System Thinkers

The rapid advancement of Artificial Intelligence (AI) is fundamentally transforming the role of software engineers, shifting expectations from traditional coding skills to holistic system thinking. This talk explores how the conventional goals of education—securing a degree, job, or financial stability—are becoming insufficient in a world where the rules of the game have been completely rewritten . As AI increasingly automates routine programming tasks, engineers are now required to evolve beyond implementation roles into problem definers, system designers, and critical thinkers.

The session introduces a structured framework for developing a "1% Engineer" through five essential layers: thinking, building, communication, presence, and growth. Emphasis is placed on first-principles thinking rather than template-based learning, encouraging individuals to break down complex problems into fundamental components. Practical capability is highlighted through project-based learning, where real-world problem solving takes precedence over mere certification. Additionally, the ability to communicate complex ideas in simple terms is identified as a key differentiator in the modern engineering landscape.

The talk also underscores the importance of professional presence and continuous learning as habits that drive long-term success. With AI already embedded in everyday workflows, the role of engineers is no longer to code more, but to think more deeply and strategically . This paradigm shift marks the transition from coders who execute instructions to system thinkers who design solutions and question assumptions.

Finally, the session advocates for a "founder mindset," where engineers proactively identify problems, build solutions, and take ownership of outcomes. It concludes with a call to action for learners to focus on building meaningful systems and shaping technology, rather than merely using it. This talk provides a roadmap for aspiring engineers to remain relevant and impactful in an AI-driven future.

Trends in automotive electronics and embedded system

Electronic systems have become an increasingly large component of the cost of an automobile, from only around 1% of its value in 1950 to around 30% in 2010.Powertrain electronics, safety, chassis electronics, and infotainment systems are the four major categories of automotive electronic systems.Future autonomous cars will rely on powerful computer systems, an array of sensors, networking, and satellite navigation, all of which will require electronics.

Automotive architecture is evolving from decentralized, function-specific Electronic Control Units (ECUs) toward centralized, zonal, and software-defined platforms. This shift, driven by "CASE" trends (Connected, Autonomous, Shared, Electric), reduces wiring complexity and weight while supporting high computing power for advanced software. Two main technologies are now primarily used in ADAS system are: LiDAR (Light Detection and Ranging), and visual sensors (cameras) which capture images and video like human eyes. These are combined with systems such as GPS, neural networks, artificial intelligence, and established ADAS engineering to deliver levels of driving autonomy.

An autonomous car is a vehicle capable of sensing its environment and operating without human involvement. A human passenger is not required to take control of the vehicle at any time, nor is a human passenger required to be present in the vehicle at all. The Society of Automotive Engineers (SAE) currently defines 6 levels of driving automation ranging from Level 0 (fully manual) to Level 5 (fully autonomous). These levels have been adopted by the U.S. Department of Transportation.

A Software-Defined Vehicle (SDV) is a vehicle whose features, functionality, and user experience are primarily enabled and updated through software rather than hardware. Using centralized, powerful computers rather than hundreds of individual, fixed ECUs, SDVs allow for continuous, over-the-air (OTA) updates to enhance safety, performance, and personalization throughout the vehicle's lifespan. IBM research predicts that 90% of all vehicle-related innovations are expected to consist of software in 2030. And 75% of auto industry executives anticipate that the software-defined experience is going to be the core of brand value by 2035.

The future of automated driving presents transformative opportunities across economic, societal, and technological domains, with the potential to create $300 billion to $400 billion in revenue by 2035. Beyond the immediate consumer appeal of convenience, automated driving (AD) is expected to revolutionize logistics, enhance safety, and enable new, data-driven business models in the automotive sector.

Organising Committee – ICASET 2026

Patron

- Dr. V. S. Gopal, Chairman, Ahalia Group

Conference Chair

- Dr. P. R. Suresh, Principal, ASET

Conference Co-Chair

- Dr. Krishna Kumar Kishor, Vice Principal, ASET & Executive Director, Ahalia Group

Department Conveners

- Dr. V. Balamurugan – Professor & Head, ECE
- Dr. C. Pradip – Professor & Head, EEE
- Dr. Bhavin K Bharath – Professor & Head, ME
- Dr. S. Gunasekaran – Professor & Head, CSE
- Dr. Mathew B. V. – Professor & Head, CE
- Mr. Sundar V – Professor & Head, AIML

Executive Conveners

- Dr. Rajeev N – Dean Academics
- Dr. Binoy Balan K – Professor & Head, S&H

Department Conference Coordinators

- Dr. Reshmi B – Associate Professor, CSE
- Mr. Manickavasagam A – Assistant Professor, EEE
- Mr. Jishnoop J – Assistant Professor, ME
- Dr. Aneesh K – Associate Professor, ECE
- Mr. Jayakrishnan R – Assistant Professor, CE
- Mrs. Reshma P. K – Assistant Professor, AIML

Publication & IT Support

- Mr. Lal S – Librarian
- Ms. Amrita Devadasan – Assistant Professor, CSE

International Advisory Committee

- Dr. Zakariya Kaneesamkandi – King Saud University, KSA
- Prof. Dr. Afsin GUNGOR – Akdeniz University, Turkey
- Dr. Mathew T Mathew – University of Illinois, Chicago, USA

- Dr. Narayanan Neithalath – Arizona State University, USA

National Advisory Committee

- Dr. Arun P – NIT Calicut
- Dr. Vinay Panicker – NIT Calicut
- Dr. Shanthakumar Mohan – IIT Palakkad
- Dr. N. Sendhil Kumar – NIT Puducherry
- Dr. A. M. Pradeep – IIT Bombay
- Dr. S. Murugan – NIT Rourkela
- Dr. Kiran Vijayan – IIT Kharagpur
- Dr. A. V. Prajeesh – VIT Vellore
- Dr. Vipin B – IIT Madras
- Dr. Santosh G. Thampi – NIT Calicut
- Dr. Kumaravel S – NIT Calicut
- Dr. J. Christopher Clement – VIT Vellore

Track Chair/In-Charge

Track	Discipline	Track Chair / In-Charge	Designation & Affiliation	
Track 1	Civil Engineering	Dr. Surya S S	Asistant Professor, Department of Civil Engineering, NSS College of Engineering, Palakkad	
Track 2	Computer Science & Engineering	Dr. Maya Mohan	Associate Professor, Department of Computer Science and Engineering NSS College of Engineering, Palakkad	
		Dr. Geethu Mary George,	HoD, Department of Computer Science & Engineering (Cyber Security), Jyothi Engineering College.	
Track 3	Electronics & Communication Engineering (ECE)	Dr. Sajitha A S	Associate Professor, Department of AI & DS Jyothi Engineering College,Thrissur	
Track 4	Electrical & Electronics Engineering (EEE)	Dr. Sujith S	Professor, Department of EEE NSS College of Engineering, Palakkad	
Track 5	Mechanical Engineering (ME)	Dr. B. Santhosh	Associate Professor, Department of Mechanical Engineering Amrita School of Engineering, Coimbatore	

Track 6	Others	Dr. E K Girisan	Dean, IQAC Head – Department of Computer Technology Sri Krishna Adithya College, Coimbatore Hon. Secretary – CSI, Coimbatore Chapter	

S.No	Paper No.	Title	Authors	Page No.
CIVIL ENGINEERING				
1	CE-01	Performance Evaluation of a Multi-Media Filtration System for Household Greywater Reuse	Dr. B V Mathew, Neethu John; Greeshma K, Shahina K, Nitin Jitenderji C (BTech Students)	2
2	CE-02	Use of Near Surface Mounted and Externally Bonded Techniques for Structural Upgrading of Building	Sreya K S; Aswathy V, Akash B, Anushwal Krishna S, Anurudh R	6
3	CE-03	Traffic Congestion Mitigation and Signal Optimization in Palakkad Town	Jayakrishnan R; Aiswarya K, Ajanya R, Hrithika R, Niranjan R P	11
4	CE-04	Defluoridation of Water Using Electrocoagulation	Thejaswi S; BhavanaPrahladh, Sreeraghav	16
5	CE-05	Composite Gypsum Boards Reinforced with Agro-Industrial Wastes	Ayana S Devan; Akhil M, Akshaya Krishnan M, Rohith M, Akhil Krishnan MK	23
6	CE-06	Life Cycle Analysis of Pavement in Kanjikode Industrial Area for Next 15 Years	Vidhun N; Sulfath H, NifaThaslim M, Basudev P D	29
7	CE-07	Catalytic Pyrolysis of Polypropylene Using Eggshell-Derived CaO: Effect of Catalyst Loading on Fuel Yield and Quality	Vimal Raj J	36
8	CE-08	Performance Evaluation of an Integrated Ultrafiltration–Reverse Osmosis System for Municipal Wastewater Reuse in Industrial Applications	Vignesh S	39
9	CE-09	Flow modelling of aquifer for tracing out of contaminants in karst topography	Pradeepa F, KavisriManikannan M	45
COMPUTER SCIENCE & ENGINEERING				
10	CSE-01	PreviDict: An AI-Based Framework for Pre-Upload View Prediction and Content Optimization of YouTube Shorts	JavharaSherin V; Anupama K T, Navarang S Manoj, Yash Krishna K P, Aswani N M, Arya A Surya	52
11	CSE-02	Integrating Human Activity Recognition and Affective Computing for Elderly Well-Being: A Comprehensive Review	Dona Maria Mani	57
12	CSE-03	Sustainable Route Planning in Maritime Transport Using Machine Learning: A Comprehensive Review	Aneesha K Jose	63

13	CSE-04	LURNZO: An AI-Based Technology Learning Platform	LamiyaFathima; AyishaNajiya P, FathimaFidha C P, Sameeha T, Ruby Fathima A C	75
14	CSE-05	ASTHRA: An Adaptive AI Framework for Automated Bulk Document Synthesis	Midhun M; Dr. Reshmi B, Nithya S, Rithin B	81
15	CSE-06	Real-Time College Bus Tracking and Route Assistance System (BusMate)	Anagha S Anil; Aneeta P, Nandana R, Aiswarya K V, Dhanya P G	85
16	CSE-07	HealHer: Web Platform for PCOS/PCOD Risk Assessment & AI Health Assistance	Hareeshna T G; AmrithaDevadasan, Ann Milena J, Diya Krishna M D, Teertha K K	88
17	CSE-08	Data Analysis with Generative AI: The D-Studio	CP Vishnu Subramanian; AtulKarthick CK, Naveen S	94
18	CSE-09	AI-Based Skin Disease Detection Using CNN	Neha U; Anisree P G, Nandhagopan M S, Prathyush P Nambiar, Muhammad Sahil K	97
19	CSE-10	Web App Security Analyzer with AI Framework	Nivedh V; Kavitha S, Milan Krishna, KassimNoushin N	101
20	CSE-11	Breach Checker	Anirudh G Menon; Rishika, HidhunBabu, Sreedev K A	106
21	CSE-12	Smart Vision: AI-Based Face Recognition Attendance System	Arjun Sunil; Harikrishna S, Sharon Manoj, Sukesh R	111
22	CSE-13	Skin Conseiller: Personalized Skincare for Indian Skin	Dhanya P; Amrita CJ, Sreelaya S, Varuna Varma V	115
23	CSE-14	Hybrid Post-Quantum Key Encapsulation for Mobile Messaging	Ajay M; Dhananjay Krishna S, R Gautham Krishna, Aswathy A	118
24	CSE-15	Explainable AI-Based Real-Time Phishing Detection	Nithya R; Alzeena A, Asha K, Riya Lakshmi S, Dr. Reshmi B, Dr. Gunasekaran S	125
25	CSE-16	Inside Ahalia : Indoor smart campus navigation system	Sreehari Athiyarath; Abhiram R, Mohammed Sanjidh S, Ajuwad B, Happy A, Gunasekaran S	133
26	CSE-17	AI-Assisted UI Builder Reducing Design Ambiguity	Abhiram K; Vishnu VS, Poojitha S, SomyaRanjanMohanty	140

27	CSE-18	AI Debugging Assistant using OCR & Local LLMs	D Haritha; Dharshin R, AfreenYusaf, Keerthana Mohan	145
28	CSE-19	Rubik's Cube Solver Using Computer Vision &OpenCV	Sayuj S; Rahul R, Bholu Kumar D, Hashim E H	152
29	CSE-20	ImpactMatch: AI-Driven Volunteer Platform	Pavithra M; Sneha G, Asitha R Prajith, M Nisthula	157
30	CSE-21	Civic Fix: Digital Platform for Civic Grievance Management	Smrithi G Iyer; Muhammed Sherif I, N P Prajwalitha, Sreenish S	164
31	CSE-22	Nexora AI	Saniya Fathima A; Anjal S, AfrahFathima R, Varsha G,Dr. S Gunasekaran	169
32	CSE-23	Automatic Exam Hall Seating Arrangement Using AI	Arya P M; Lidhin M, Neethu Suresh, ThansilThanish K S	172
33	CSE-24	AI-Orchestrated Digital Fabrication Workflow	H A Rahulnath; SumeshBalan	178
34	CSE-25	AI-BASED IMAGE AUTHENTICITY DETECTION SYSTEM	Akshaya M; Deepthy PS, Aryananda P, Sriparvathy C, Sayana V	186
35	CSE-26	YouTube Sentiment Analyzer using NLP	Gopika B; Aswan K Prakash, Nivedh A Nair, Rithu P	192
36	CSE-27	Lost and Found System	Adarsh A; Athul Krishna K, Midhun T Chandran	197
37	CSE-28	AI-Based Exam Evaluation System	Adhip Suresh; Sreeyesh Sunil Kumar, Abhijith PG, MithunYuvanah G	202
38	CSE-29	Smart Learn AI: Intelligent Study Assistant	Pranav Krishnan S; Nandana TA, Nikhil C, Sreeshankar KS	209
39	CSE-30	Personalized AI Chatbot for KTU CSE Students	IffahHussainbi S; NasibaRahana S, Arjitha A, NikithaJayan Nair	213
40	CSE-31	Realtime Collaborative Code Editor	Mrinal V; Midhun V, Mohammed Shahal A, Navaneeth Krishna M M	222
ELECTRONICS & COMMUNICATION ENGINEERING				
41	ECE-01	Brain Tumor Detection Using Multimodal MRI Scans with Deep Learning and Edge AI	Pooja P Menon; Subiksha S; Nikhil V; Mohammed Haddiq S; Dr. Leesha Paul; Dr. V. Balamurugan	226

42	ECE-02	Design and FPGA Implementation of a 32-bit RISC Processor Using Verilog HDL	Arya P R; Ashrin A; Athish R; Chaithra K C; Dr. Aneesh K; Dr. V. Balamurugan	231
43	ECE-03	A Compact Clover-Shaped Slot-Loaded Microstrip Patch Antenna for Sub-6 GHz Applications	Nandana N; Dr. V. Balamurugan; Mr. Sanish VS; Rijul R Das; Shakira U; Sivani PS	236
44	ECE-04	CATRON: A Cable Tunnel Inspection Robot	Dhanush P; Arya Sasikumar; Gowthami S; Mahitha M; Dr. V. Balamurugan; Swetha C	241
45	ECE-05	VLSI Architecture for Video Edge Detection	Aditya Kiran S R; AnirudhSreejithesh; Ashik M; Jishnu R; Gayathri P S; Dr. V. Balamurugan	246
46	ECE-06	AI Powered Robot for Microplastic Detection and Extraction	Sreerag P; Sanu C S; Swetha R; Thripthi M; Ms. Vijitha Khan; Dr. V. Balamurugan	252
47	ECE-07	AI Powered Laser-Based Autonomous Rover for Weed Detection and Removal	Sooraj S; Sidharth K; Vineeth M; Muhammad Aslam KM	257
48	ECE-08	Solar Powered Oxygen Concentrator with Real-Time Monitoring and Battery Backup	Abhishek K V; Ajil J; Ajmal T A; Athul G; Vinod Kumar P; Dr. V. Balamurugan	262
49	ECE-09	Intelligent Front Desk Framework for Enhancing Patient Flow in Hospitals	Anamika S; Akshaya Sri P S; MehrinAyisha S; Manu S Vignesh; Dr. V. Balamurugan	265
50	ECE-10	Non-Contact Voltage Detector Gloves Using 555 Timer	AkhibAhamed M; Darsan V; Ann Maria Neefa; HarishmaChandran	270
51	ECE-11	Real-Time ECG Monitoring and Analysis System Using Arduino	AjilMadhav C K; Abhinandh PN; Ajay NJ; Ananthakrishnan S; M Mohammed Niyas	273
52	ECE-12	Automatic Indicator System Using ESP32 and GPS	Mohammed Irfan B; Prasin Krishna E; SainulAbidh PS; Sree Vishnu S; Dr. V. Balamurugan; Ms. Gayathri P S	275
53	ECE-13	XRAY3D: Deep Learning-Based 3D Reconstruction from 2D Radiographs	Mridhul Krishna M; Muhammed Ibrahim S; Sabarinath M; Sriram G	280

54	ECE-14	Performance Analysis of Multiple FPGA Platforms Using Up and Down Counters	Anuja S; Akshaya A; Arya B; Arya R	283
55	ECE-15	Tomato Quality Classification	Chandru; Dr. Divya Mohan; Dr. V. Balamurugan; Sreelakshmi V; Rayan Gaffor; M Padmaja	287
ELECTRICAL & ELECTRONICS ENGINEERING				
56	EEE-01	Smart Infusion Pump Using IoT	Mohammed A; Karthika G; Rahul R; Ms. RosuVarkeyachan; Akshay Kumar M	291
57	EEE-02	Fingerprint Electronic Voting Machine	Adithyan P; Midhun J	303
58	EEE-03	Design and Implementation of Smart Bag Tracking and Safety System for Kids	Babhitha Rosh B; Gowtham S; Hridya R; Rikas M; A. Manickavasagam	309
59	EEE-04	Design of Smart Oxygen Concentrator with Reduced Grid Reliance	Ajith M; AnirudhReji; Fasir Ali; Sreya P; Ms. Lakshmi Suresh	314
60	EEE-05	Smart Home Illumination Using Smart Sensors and IoT	Neha Babu; Arjun A; Thejus Das V; Sudarsh S; Subin Chacko	320
61	EEE-06	Short-Term Load Forecasting in Smart Grid Using Dragonfly Algorithm with BPN	Sivakumar R; Pradip C; Manickavasagam A; John Alexis S; Muruganath Gopal Raj	327
MECHANICAL ENGINEERING				
62	ME-01	Performance Monitoring of a Savonius Vertical Axis Windmill	Prathyush M; Sadanandan R; Dr. Vivek Gopi; Karthikeyan V; Muhammed Kareem F; Ramith R	337
63	ME-02	Studies on the Production of Compressed Air Using Wave Energy-Based Air Compressor System	Akhilesh P; Rebin Husain M; Mohammed Noufal N; Naveen K, Dr.Johnalexis	342
64	ME-03	Parametric Study on the Vibration Characteristics of Neoprene Rubber Isolators	Anna InasuChungath; Charles Michael Augustine; Dawn Cletus Francis; Srutheesh T P	346

65	ME-04	Experimental Investigation of an Innovative System for Solar Desalination	Shabith A; Dr. P R Suresh	352
66	ME-05	Solar-Based Continuous Hot Water Supply System Enhanced with Phase Change Materials	Abhijith M M; Hrithikkrishna V D; Fevin Francis; Justin Thomas; SarathBabu Ramachandran; Dr. Renjish Vijay	359
67	ME-06	Experimental Investigation of PLA-Based Natural Composites in 3D Printing	Adarsh S; Shiyas S; Vishnu M; JishnoopJayaprakash:Venkitesh K S	363
68	ME-07	Integration of Total Quality Management (TQM) in Mechanical Engineering Processes	Dr. ShindeGobaDevman; Dr. Sancheti Santosh D; SanchetiDipak D	368
69	ME-08	Performance Evaluation of Various Feed Water Approaches in Passive Solar Still Desalination	D'Cotha J S; Suresh P R; Sajeesh P	374
70	ME-09	Acoustic Streaming Effects on the Atomization of a Spray	Kasif Ehsan; AadilKureshee; Deepak Kumar Mandal; S. Narayanan	380
71	ME-10	A Critical Summary on Drop Impact Icing and the Role of Superhydrophobic Surfaces in Anti-Icing Applications	K KKrishnaram; Rahul Sharma; A R Harikrishnan	385
72	ME-11	Next-generation lightning protection using conductive CFRP laminates	Mohan Kumar Pitchan, Balachandran	390
73	ME-12	Design Development and Testing of motor test rig for an EV with continuously variable transmission (cvt)	Vinod V. Muhammed Jaseem Muhammed Shibil, Sreelakshmy PK Sradha Shibu	396
OTHERS				
74	OT-01	Data vs. Model Machine Learning Fairness Testing: An Empirical Study	Vithul M; Reshma P K; Sankar Krishna P; Sreya S; Darshan R M	402
75	OT-02	Regenerative Effects of α-Mangostin-Loaded ZnO Nanoparticles on Wound Healing: In Vitro and Zebrafish In Vivo Insights	Taniya Mary Martin; MeenakshiSundaram Kishore Kumar	404
76	OT-03	Determining the Dominant Social Media Platform using Hexa Graph Topology and Matrix Representation	Santhiya S	416
77	OT-04	Analytical Computation of Atangana–BaleanuDifferintegrals of Classical Special Functions	Prabha R; Sunitha K G; Sangeetha P S; Preeja V	420

78	OT-05	Graph Analytics for Large Scale Networks	BinoyBalan K; G. Murugananth; Rajitha M; Theertha K	424
79	OT-06	Comparative Structural Analysis of ERCO Ercoupe Wing Rib with Different Cut-Out Geometries	Dr. Ajith V S; Dr. Rajeev N; FathimaFidha K; Geethika K G; Ardra P; Haritha R S	428

TRACK-1

CIVIL ENGINEERING

PERFORMANCE EVALUATION OF A MULTI-MEDIA FILTRATION SYSTEM FOR HOUSEHOLD GREYWATER REUSE

Greeshma K
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
greeshmakumar2004@gmail.com

Shahina K
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
shahinakaja986@gmail.com

Nitin Jitendergi C
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
nitinjitendergic2019@gmail.com

Neethu John
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
neethu.john@ahalia.ac.in

Dr. B V Mathew
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
hodce.aset@ahalia.ac.in

***Abstract*—**Water scarcity is becoming a serious concern in many regions, especially in water-stressed areas where sustainable water management is essential. A significant portion, nearly 50–70%, of domestic wastewater generated in households is greywater, which originates from bathrooms, wash basins and kitchens. In most situations, this greywater is discharged without any form of treatment, resulting in environmental pollution and unnecessary loss of potentially reusable water resources. This study presents the, development and performance evaluation of three laboratory-scale multi-media filtration systems constructed using locally available and low-cost natural materials. The systems were designed to function under simple gravity flow conditions, eliminating the need for electricity and making them suitable for decentralized household applications. The treated greywater was assessed for non-potable reuse purposes such as gardening and toilet flushing. To evaluate the effectiveness of the systems, important water quality parameters including pH, turbidity, total dissolved solids (TDS), total hardness, biological oxygen demand (BOD), chemical oxygen demand (COD), oil and grease, and E.coli were tested before and after treatment. The findings revealed a considerable reduction in pollutant levels across all systems, with System 2 exhibiting the highest overall removal efficiency. The results highlight the potential of low-cost multi-media filtration as a practical and sustainable solution for decentralized greywater treatment and reuse.

Keywords*—Greywater treatment, Multi-media filtration, Water reuse, Sustainable water management, Gravity-based treatment*

INTRODUCTION

Water scarcity has become one of the most pressing environmental challenges of our time. Rapid urbanization, continuous population growth and the impacts of climate change are placing increasing pressure on available freshwater resources. In India, especially in semi-arid regions such as Palakkad district, the demand for water is steadily rising while the availability remains limited. At the same time, households generate a considerable amount of wastewater every day, and a major portion of this is greywater.

Greywater, which comes from bathrooms, wash basins and kitchens, contains suspended solids, organic matter, detergents, oil and grease, nutrients and microorganisms. Although it is comparatively less polluted than blackwater from toilets, its direct discharge into the environment can still lead to soil contamination, groundwater pollution and unpleasant surroundings. If properly treated, greywater can be safely reused for non-potable purposes, thereby reducing household freshwater demand by nearly 30–40%.

Traditional centralized wastewater treatment plants involve high construction and operational costs, require skilled manpower and depend on continuous energy supply. These limitations make them impractical for small-scale or household-level applications. As a result, decentralized and low-cost treatment options are receiving greater attention. Multi-media filtration is one such simple and effective technique that uses layers of natural materials to remove impurities through physical filtration, adsorption and limited biological processes.

In this context, the present study focuses on evaluating the performance of three different multi-media filtration combinations. The objective is to identify the most efficient and practical configuration that can be implemented at the household level for safe greywater reuse.

LITERATURE REVIEW

Several earlier studies have highlighted the effectiveness of natural filtration systems in improving wastewater quality. Techniques such as slow sand filtration and constructed wetlands have been widely adopted for decentralized wastewater treatment, particularly in rural and semi-urban areas. These systems are appreciated for their simplicity, low operational cost and minimal energy requirement. Research findings indicate that natural materials like rice husk, coconut coir, charcoal and gravel have good filtration and adsorption properties, enabling them to remove suspended solids and reduce turbidity and organic load effectively.

In the context of greywater treatment, many researchers have reported that multi-layer filtration systems can significantly reduce parameters such as Biological Oxygen Demand (BOD), Chemical Oxygen Demand (COD) and Total Dissolved Solids (TDS). The performance of such systems is influenced by several factors, including the type and combination of filter media, thickness of layers, flow rate and hydraulic retention time. Proper selection and arrangement of media play a crucial role in achieving higher treatment efficiency.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Although a number of studies have explored natural filtration techniques, limited research has been conducted on comparative performance evaluation of different media combinations using locally available materials, particularly under rural Kerala conditions. Most existing studies focus either on a single filtration setup or on large-scale treatment systems.

The present study aims to address this gap by developing and comparing three different low-cost multi-media filtration models using materials that are easily available locally. The objective is to identify the most efficient and practical system suitable for decentralized household greywater reuse.

METHODOLOGY

The methodology of the study consisted of fabricating and evaluating three gravity-based multi-media filtration systems for greywater treatment. Greywater samples were collected from domestic sources within the campus premises to ensure that the study reflected real household wastewater conditions.

A variety of locally available filter media materials were selected based on their filtration and adsorption properties. These included gravel, river sand, charcoal, coconut coir, burnt coconut shell, wood chips, crushed bricks, tamarind shell, rice husk and sugarcane bagasse. Before use, all materials were thoroughly washed and cleaned to remove dust, fine particles and other impurities that could affect filtration performance.

The filtration units were fabricated using plastic trays . The filter media were carefully arranged in layered form, with coarser materials placed at the bottom and finer materials positioned towards the top. This arrangement helped in achieving gradual filtration, where larger particles were removed first, followed by finer contaminants as the water passed through successive layers.

Greywater was poured into the top of the column and allowed to flow downward under gravity, eliminating the need for pumps or external energy sources. Samples were collected before and after filtration for analysis. Laboratory tests were conducted to determine important water quality parameters such as pH, turbidity, total dissolved solids (TDS), total hardness, biological oxygen demand (BOD), chemical oxygen demand (COD), oil and grease, and E.coli levels. The treatment performance of each system was evaluated by calculating the percentage removal efficiency using standard analytical methods.

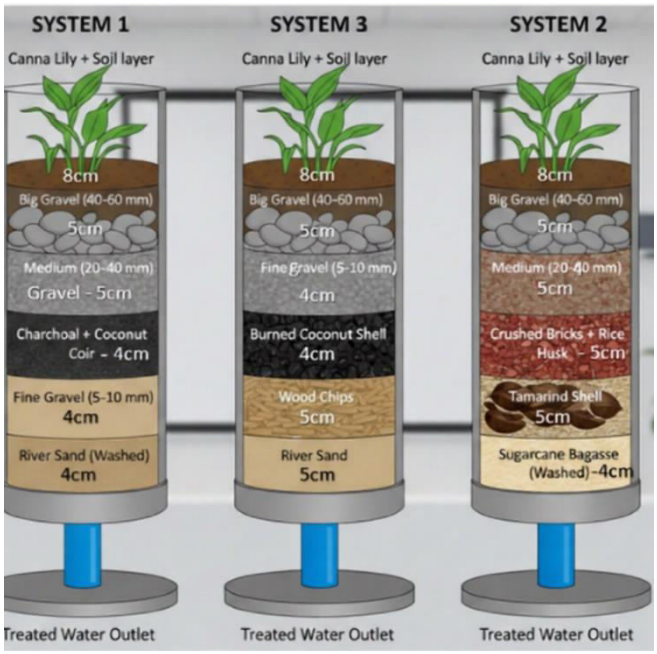
DEVELOPMENT OF FILTRATION SYSTEM

Each filtration unit was carefully designed as a vertical arrangement fitted with a bottom outlet valve to collect the treated water. The vertical arrangement allowed greywater to flow downward under gravity, ensuring simple operation without the need for pumps or electricity. The design also provided proper drainage and minimized the chances of clogging by maintaining adequate spacing between filter layers.

Three different filtration systems were developed using varying combinations of natural filter media. System 1 was constructed using canna lilies planted with soil, layers of big gravel, medium gravels, charcoal mixed with coconut coir, fine gravel, sand. System 2 consisted of canna lilies, big gravel, medium gravel, crushed brick + rice husk, sugarcane bagasse, tamarind shell. System 3 was consists of canna lilies, big gravel, fine gravel, burned coconut shell, wood chips, sand.

The selection of filter media was primarily based on local availability, porosity, adsorption capacity, durability and overall cost-effectiveness. Coarse materials were placed at the bottom to support drainage and remove larger suspended particles, while finer and more porous materials were arranged above to enhance adsorption and filtration of smaller contaminants.

The layered configuration allowed the system to remove pollutants through multiple mechanisms, including mechanical straining of suspended solids, adsorption of dissolved impurities onto media surfaces and limited biological degradation of organic matter within the filter bed. This combined action improved the overall treatment efficiency of the filtration units.



RESULT

- The below table shows the water quality parameters before treatment and after three different treatment systems, compared with the guidelines from the Central Pollution Control Board (CPCB), India.
- Before treatment, the water quality was poor and above permissible limits for most parameters. The pH was 9.2, which is higher than the recommended range of 5.5–7.0, indicating alkaline water. Total dissolved solids (TDS) were 948 mg/L, which is more than double the recommended limit of 450 mg/L, showing a high amount of dissolved minerals or salts. Turbidity was 125.5 NTU, far above the acceptable range of 10–50 NTU, indicating the presence of suspended particles. Total hardness (212 mg/L), BOD (248 mg/L), COD (228 mg/L), and oil and grease (18 mg/L) were also higher than the CPCB limits, indicating significant organic pollution and contamination. However, E. coli/coliform bacteria were not detected (BDL – Below Detection Limit).

- After applying Treatment System 1, some improvement occurred but many parameters were still above acceptable limits. The pH decreased slightly to 7.7, TDS reduced to 903 mg/L, and turbidity decreased to 117.6 NTU, but these values were still higher than recommended. BOD and COD decreased to 104 mg/L and 152 mg/L, showing partial removal of organic pollutants. Oil and grease were reduced to below detection limit, which is a positive outcome.
- Treatment System 2 showed the best performance overall. The pH improved to 7.1, which is close to the recommended range. TDS dropped significantly to 386 mg/L, meeting CPCB guidelines. Turbidity decreased to 47.2 NTU, which falls within the acceptable range. Total hardness decreased to 152.7 mg/L, also within limits. Most importantly, BOD reduced to 26 mg/L and COD to 45.8 mg/L, both meeting CPCB standards, indicating effective removal of organic pollutants. Oil and grease remained below detection limits.
- After Treatment System 3, the water quality improved compared to the untreated sample but was generally worse than Treatment System 2. The pH increased again to 8.34, exceeding the recommended range. TDS (482 mg/L) and turbidity (79.9 NTU) were higher than the limits. BOD (56 mg/L) and COD (186 mg/L) also exceeded permissible levels, although oil and grease remained below detection limits.
- Overall, the results indicate that Treatment System 2 is the most effective method, as it reduces most parameters to within CPCB standards and significantly improves the water quality compared to the untreated sample and the other treatment systems.

SL NO	PARAMETERS	UNIT	BASED ON GUIDELINES FROM CENTRAL POLLUTION CONTROL BOARD (CPCB) IN INDIA	BEFORE TREATMENT REULTS	AFTER TREATMENT SYSTEM 1 RESULTS	AFTER TREATMENT SYSTEM 2 RESULTS	AFTER TREATMENT SYSTEM 3 RESULTS
1.	pH		5.5 – 7.0	9.2	7.7	7.1	8.34
2.	Total dissolved solids	Mg/l	450	948	903	386	482
3.	Turbidity	NTU	10 - 50	125.5	117.6	47.2	79.9
4.	E.Coli / Coliforms	mL	100	BDL	BDL	BDL	BDL
5.	Total hardness	Mg/l	200	212	161.5	152.7	171.1
6.	BOD	Mg/l	<30	248	104	26	56
7.	COD	Mg/l	<50	228	152	45.8	186
8.	Oil and	Mg/l	<10	18	BDL	BDL	BDL

EQUATION

$$\textbf{Efficiency} = \frac{Before\ test\ value - After\ test\ value}{Before\ test\ value} * 100$$

1. $\textbf{System 1} = \frac{4.75+6.29+23.82+58.06+33.33}{5} = 25.25 \sim 25\%$

2. $\textbf{System 2} = \frac{59.28+62.39+27.97+89.52+79.91}{5} = 63.81 \sim 64\%$

3. $\textbf{System 3} = \frac{49.15+36.33+19.29+77.42+18.42}{5} = 40.12 \sim 40\%$

ADVANTAGES OF THE SYSTEM

The developed multi-media filtration system offers several practical advantages, making it suitable for household-level greywater treatment. One of its major strengths is its low construction cost, as it utilizes locally available and inexpensive natural materials. This makes the system affordable and accessible, particularly in rural and semi-urban areas.

The system is environmentally friendly, since it relies on biodegradable and natural filter media rather than chemical treatment processes. It operates entirely under gravity flow, eliminating the need for electricity or mechanical pumps. This not only reduces operational costs but also makes the system reliable in areas with limited power supply.

In addition, the filtration unit is simple in design and easy to construct using basic materials such as Plastic trays and natural filter layers. Installation can be carried out at the household level without requiring advanced technical skills. Maintenance requirements are minimal, mainly involving periodic cleaning or replacement of filter media to maintain efficiency.

Overall, the system provides a cost-effective, sustainable and user-friendly solution for decentralized greywater treatment and reuse.

LIMITATIONS OF THE SYSTEM

Despite its effectiveness and simplicity, the developed multi-media filtration system has certain limitations. The treated water is suitable only for non-potable reuse applications such as gardening and toilet flushing, and it is not recommended for direct human consumption without additional advanced treatment and disinfection processes.

Regular maintenance is essential to ensure consistent performance. Over time, suspended particles and organic matter may accumulate within the filter layers, leading to clogging and reduced flow rate. Therefore, periodic cleaning and proper maintenance are necessary to maintain filtration efficiency.

In addition, the performance of the system may gradually decline during long-term operation if the filter media are not replaced or regenerated at appropriate intervals. The adsorption capacity of natural materials decreases over time as active sites become saturated. Hence, scheduled replacement of media is required to sustain treatment effectiveness.

CONCLUSION

The present study clearly demonstrates that multi-media filtration using natural and locally available materials can serve as an effective decentralized solution for greywater treatment. All three filtration systems developed in this study showed noticeable improvement in key water quality parameters, confirming the practicality of layered natural media in removing physical, chemical and organic contaminants.

Among the three configurations tested, System 2 achieved the highest overall removal efficiency and exhibited better performance in reducing turbidity, TDS, BOD and COD levels. Based on the experimental findings, System 3 is recommended for household-level non-potable reuse applications such as gardening and toilet flushing.

Overall, the study emphasizes the importance of adopting simple, sustainable and low-cost treatment technologies to reduce freshwater consumption and promote water management. The developed system provides a practical approach to addressing water scarcity challenges, particularly in regions facing increasing pressure on water resources.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to our guide and and supervisor for their continuous support, valuable guidance, and encouragement throughout the completion of this project titled “Performance Evaluation of a Multi-Media Filtration System For Household Greywater Reuse.” Their insightful suggestions and constructive feedback helped us successfully complete this work.

We would also like to thank the faculty members of the department for providing the necessary facilities, resources, and academic support required for this study.

REFERENCE

[1]. Oteng-Peprah M, Acheampong M. A, and DeVries, H. (2018). Greywater characteristics, treatment systems and reuse potentials:A review. Environmental Science and Pollution Research (review), 2018, 25.

[2]. Pradhan S, and Sharma R. (2019). Greywater treatment by ornamental plants and media for reuse in buildings. Journal of Cleaner Production, 2019, 236.

[3]. Feitosa A. P, et al. (2024). Enhancing greywater treatment: high-efficiency vertical flow filters with alternative substrates. Applied Sciences, 14(19).

[4]. Guti´errez V, Monsalves N., G´ omez, G., and Vidal G. (2023). Performance of a full-scale vermifilter for sewage treatment. Sustainability, 15(8), 6842.

[5]. Bani-Melhem, K, et al. (2023). Impact of a sand filtration pretreatment step on high-strength greywater electrocoagulation. Water (MDPI), 15(5), 990.

[6]. Busgang A, et al. (2015). Epidemiological study for the assessment of health risks associated with greywater reuse for irrigation and toilet flushing. Science of the Total Environment, 2015, 533.

[7]. Miito G. J, et al. (2024). Influence of earthworm population density on vermifilter performance for domestic wastewater. Journal of Environmental Quality, 53(2).

[8]. Nerlekar T, and Patil Y. M. (2020). Design of greywater treatment system based on vermifiltration (pilot study). ISCA International Journal (conference/paper).. ResearchGate authors (2019). Greywater treatment using sand filtration — pilot study. (Conference / report), 2019, volume N/A.

[9]. Vymazal, J. (2022). The regime of constructed wetlands in greywater treatment. Water Science and Technology, 85(11), 3169–3178.

[10]. Coppens K, et al. (2025). Performance analysis and impact of operating conditions in vermifilters treating domestic greywater. Water Research (article, 2025), volume pending/online.

[11]. Hamidi M. N, et al. (2025). Greywater reuse for irrigation: a critical review of suitability and risks. Science of the Total Environment, 2025, (in press/online).

[12]. World Health Organization (2006). Guidelines for the safe use of wastewater, excreta and greywater, Volume II: Wastewater Use in Agriculture. WHO, Volume II (guideline).

[13]. United States Environmental Protection Agency (EPA). (2012). Guidelines for Water Reuse. EPA/600/R-12/618 (technical guideline/report, 2012). (no journal volume).

[14]. Comprehensive review (2023). Greywater treatment technologies: a comprehensive review (physical, chemical, biological and nature-based solutions). Journal/Report, 2023, volume N/A (review).

[15]. (2025) Greywater treatment using constructed wetlands modified with biochar/ceramics: performance study. Environmental Technology and Innovation, 2025.

USE OF NEAR SURFACE MOUNTING AND EXTERNALLY BONDING METHOD FOR STRUCTURAL UPGRADING OF BUILDING

*Note: Sub-titles are not captured in Xplore and should not be used

SREYA K S
DEPARTMENT OF CIVIL ENGINEERING
AHALIA SCHOOL OF ENGINEERING AND TECHNOLOGY
PALAKKAD,INDIA
KSSREYA04@GMAIL.COM

ASWATHY V
DEPARTMENT OF CIVIL ENGINEERING
AHALIA SCHOOL OF ENGINEERING AND TECHNOLOGY
PALAKKAD,INDIA
ASWATHY.V@AHALIA.AC.IN

AKASH B
DEPARTMENT OF CIVIL ENGINEERING
AHALIA SCHOOL OF ENGINEERING AND TECHNOLOGY
PALAKKAD,INDIA
AKKUAKASH69@GMAIL.COM

ANUSHWAL KRISHNA S
DEPARTMENT OF CIVIL ENGINEERING
AHALIA SCHOOL OF ENGINEERING AND TECHNOLOGY
PALAKKAD,INDIA
ANUSHWALKRISHNA9539@GMAIL.COM

ANURUDH R
DEPARTMENT OF CIVIL ENGINEERING
AHALIA SCHOOL OF ENGINEERING AND TECHNOLOGY
PALAKKAD,INDIA
ANURUDHR33@GMAIL.COM

***Abstract*—Deterioration of reinforced concrete (RC) structures due to aging, environmental exposure, and increased service loads necessitates effective strengthening techniques to enhance structural performance and extend service life. Fiber Reinforced Polymer (FRP) systems have emerged as a promising solution for retrofitting because of their high strength-to-weight ratio, corrosion resistance, and ease of installation. This study investigates the effectiveness of two FRP strengthening techniques, namely Near Surface Mounted (NSM) and Externally Bonded (EB) reinforcement, for upgrading RC beams. The structural condition of the selected beam was evaluated using Non-Destructive Testing (NDT) methods including Rebound Hammer, Ultrasonic Pulse Velocity, and Electrical Resistivity tests. Reinforced concrete beam specimens were prepared and strengthened using FRP materials to assess improvements in structural behavior. The study aims to compare the performance of NSM and EB techniques in terms of strength, durability, and applicability, providing insights for selecting appropriate retrofitting methods for existing RC structures.**

Keywords: - Reinforced concrete (RC), Fiber Reinforced Polymer (FRP), Near Surface Bonded (NSM), Externally Bonded (EB), Non-Destructive Test (NDT).

I. INTRODUCTION

Structures made of reinforced concrete (RC) frequently deteriorate as a result of ageing, exposure to the environment, reinforcement corrosion, and higher service loads. As a result, many existing structures need to be strengthened in order to satisfy modern performance and safety standards. Infrastructure can be strengthened and rehabilitated to prolong its useful life at a lower cost than demolition and reconstruction. Steel plate bonding, section expansion, and external post-tensioning are examples of conventional strengthening techniques that have been extensively employed. These techniques, however, have a number of drawbacks, such as higher structural weight, corrosion susceptibility, and installation and maintenance challenges. These shortcomings have prompted the creation of more effective strengthening materials and methods. Fiber Reinforced Polymer (FRP) composites, especially Carbon Fiber Reinforced Polymer (CFRP), have become popular materials for structural retrofitting because of their high strength-to-weight ratio, resilience to corrosion, and simplicity of use. RC members' flexural performance and durability can be greatly improved by CFRP systems. The two most popular FRP strengthening procedures are Near Surface Mounted (NSM) and Externally Bonded (EB). Using epoxy adhesives, CFRP sheets or laminates are bonded to the surface of concrete members in the EB process. Despite being easy to use, this approach may have problems with environmental exposure and early debonding. By inserting CFRP rods or strips into grooves made in the concrete cover, the NSM approach, on the other hand, enhances bond performance and offers superior reinforcement protection. The efficiency of NSM and EB strengthening methods for reinforced concrete beams is the main topic of this study. Investigating the structural performance and usefulness of these techniques in fortifying current RC structures is the driving force behind this study.

II. LITERATURE REVIEW

Previous research mostly concentrated on the Externally Bonded (EB) method, which uses epoxy adhesives to connect CFRP sheets or laminates to the concrete surface. EB strengthening greatly increases the flexural strength and stiffness of RC beams, according to research. However, one major issue limiting its effectiveness is the early debonding

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

of FRP from the concrete surface. The Near Surface Mounted (NSM) method, which involves embedding CFRP rods or strips into grooves cut in the concrete cover and adhering them with epoxy, was created to address this problem. In contrast to EB strengthening, studies by Waag and Belarbi and Seleem et al. (2023) found greater bond performance, increased load carrying capability, and better use of FRP materials in NSM systems.

While FRP strengthening approaches have been the subject of several study, few studies directly compare NSM and EB procedures under comparable circumstances. Thus, the efficacy of these two methods for reinforcing concrete beams is assessed and contrasted in this study.

III. METHODOLOGY

A. Site Investigation and Condition Assessment

The condition of the chosen reinforced concrete beam was assessed through a preliminary site survey. Rebound Hammer, Ultrasonic Pulse Velocity (UPV), and Electrical Resistivity tests are examples of Non-Destructive Testing (NDT) techniques used to evaluate corrosion risk, estimate compressive strength, and check concrete quality without causing structural damage.

B. Material Testing

The characteristics of the aggregates and cement used in the manufacture of concrete was determined using basic material tests. These included the fineness test of cement, specific gravity test of cement and fine aggregate, and normal consistency test of cement.

C. Sample Preparation

Specimens of reinforced concrete beams were cast for investigation in experiments. To guarantee uniformity, the identical mix proportions and reinforcing details were used to prepare both control beams and enhanced beams.

D. NSM Strengthening Technique

The tension face of the beam was carved with grooves for the Near Surface Mounted (NSM) technique. To improve flexural capacity, CFRP rods were inserted into the grooves and joined with epoxy adhesive.

E. EB Strengthening Technique

In the Externally Bonded (EB) method, the beam surface was cleaned and roughened before bonding CFRP sheets to the tension face using epoxy resin.

F. Flexural Testing

The beams were tested under single-point and two-point loading conditions to evaluate flexural behavior, load carrying capacity, and failure characteristics.

IV. ADVANTAGES AND LIMITATIONS

ADVANTAGES

- The strengthening techniques significantly increased the load carrying capacity and flexural strength of the RC beams compared to the control specimen.
- Strengthened specimens showed delayed crack initiation and reduced crack width, improving the durability of the structure.
- CFRP materials provide high strength-to-weight ratio and corrosion resistance, making them suitable for rehabilitation of existing structures.
- Minimal Increase in Structural Weight

LIMITATIONS

- CFRP materials are relatively expensive compared to conventional strengthening materials.
- The tests were performed under controlled laboratory conditions, which may differ from real field conditions.
- The study considered limited loading cases, and the behavior under different loading conditions such as cyclic or fatigue loading was not investigated.

V. NON-DESTRUCTIVE EVALUATION OF EXISTING RC BEAM

Fig: Rebound Hammer Test

Fig: Electrical Resistivity

.

Fig: Ultra Sonic Pulse Velocity

Test Method	Measured Values	Interpretation
Rebound Hammer Test	Avg. Rebound No.: 20.7 Corrected Value: 21.7	Concrete Quality: Fair Compressive Strength: 12 MPa Status: Border
Ultrasonic Pulse Velocity (UPV)	3.64 km/s	Concrete Quality: Good
Electrical Resistivity Test	15.86 kΩ·cm	Moderate corrosion risk

Table 1. Results of NDT Tests

Interpretation:

- The concrete shows moderate strength and fair surface hardness.
- UPV indicates relatively good internal integrity.
- Electrical resistivity suggests some possibility of reinforcement corrosion in the future.

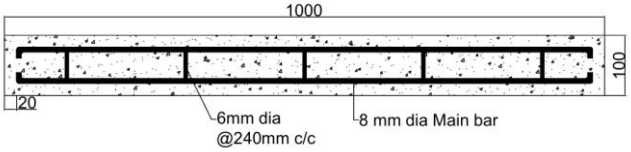
In other words, the beams were not in perfect condition but suitable for strengthening studies.

VI. MATERIAL TESTING

For determining the basic characteristics of cement and aggregates used for preparing concrete samples material quality tests were conducted. The fineness of cement is 3.5% indicated that it had been sufficiently ground and had a sufficient surface area for effective hydration. The average normal consistency of cement was found to be 31.5%, which is within the permissible limit for ordinary Portland cement and guarantees the right amount of water needed for paste preparation. The specific gravity of cement was obtained as 3.10, which is close to the standard value for ordinary Portland cement and confirms the absence of significant impurities or adulteration. The specific gravity of fine aggregate was found to be 2.62, indicating normal density and suitability for use in structural concrete.

The results indicate that the tested materials meet standard quality standards and are considered suitable for the fabrication of concrete specimens used in the experimental study. The results show that the materials have sufficient physical characteristics to provide consistent and dependable concrete for structural testing.

VII. BEAM SPECIFICATIONS



The beam specimen was designed using a concrete mix ratio of 1:2:4 (cement: sand: coarse aggregate). Fine aggregate used was M-sand, and the coarse aggregate size was 20 mm. The concrete cover provided was 20 mm.

The reinforcement details include a total beam length of 1000 mm, and the reinforcement used was Fe550 grade steel. The diameter of the main reinforcement bar was 8 mm, and the stirrup diameter was also 6 mm. The stirrup length was 560 mm, and the spacing between stirrups was 240 mm.

VIII. SELECTION OF FRP MATERIAL AND ITS IMPORTANCE IN STRENGTHENING TECHNIQUES

Fiber Reinforced Polymer (FRP) materials are widely used for strengthening and rehabilitation of reinforced concrete structures due to their high strength-to-weight ratio, corrosion resistance, and ease of installation. Among the different types of FRP materials, Carbon Fiber Reinforced Polymer (CFRP) is commonly preferred because of its superior mechanical properties, including high tensile strength, high modulus of elasticity, and excellent durability under environmental exposure. The selection of CFRP in this study is mainly based on the fact that this type of composite material is capable of significantly improving the flexural strength of reinforced concrete beams without adding significant weight to the structure. In addition, this type of composite material is resistant to corrosion and chemical effects.

In using the NSM technique, CFRP rods are embedded in grooves that are manufactured on the tension zone of the reinforced concrete beam. This type of arrangement is able to offer sufficient bond interaction between the CFRP composite material and the reinforced concrete beam, resulting in an efficient stress transfer mechanism.

In using the EB technique, CFRP sheets are externally bonded to the reinforced concrete beam. This technique is easy to use and is able to significantly enhance the load-bearing capacity and stiffness of the reinforced concrete beam.

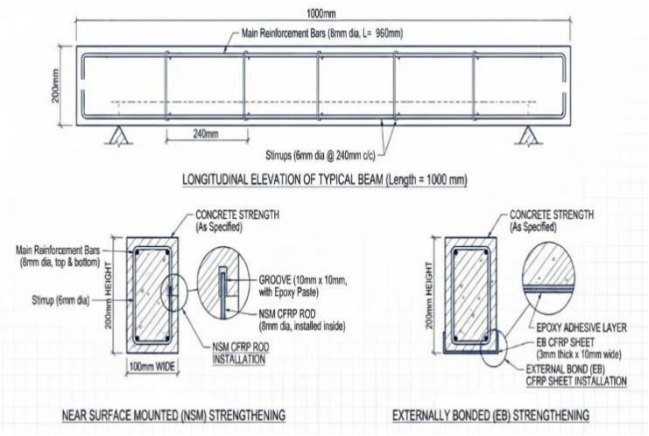
Therefore, in using CFRP composite materials in reinforced concrete beams through NSM and EB strengthening techniques, this type of composite material plays a significant role in enhancing the strength and stiffness of reinforced concrete beams.

IX. INSTALLATION OF NSM AND EB CFRP

- In the near surface mounted method of strengthening, grooves were first cut along the tension surface of the reinforced concrete beam using the required cutting tools. The width and depth of the cut grooves were about 10 mm, as required for the CFRP rods. After preparing the cut grooves, the surface was cleaned properly to remove dust and debris for proper bonding.

- Subsequently, an adhesive material made of epoxy was applied within the cut grooves, and CFRP rods of diameter 8 mm were inserted into the cut grooves. Another coating of epoxy was applied to fill the groove completely for proper encapsulation of the inserted CFRP rod. The surface was leveled and left for the required time for proper bonding between the inserted CFRP rod and the concrete surface.

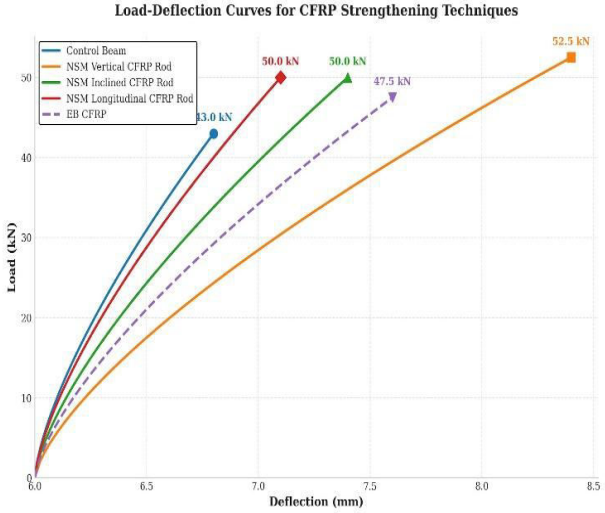
- For the Externally Bonded strengthening method, the concrete surface on the tension side of the beam was first prepared by grinding and cleaning to remove laitance, dust, and loose material. Proper surface preparation was done to enhance the bond between the CFRP sheet and the concrete surface.
- An epoxy resin layer was applied to the prepared concrete surface. CFRP sheets measuring 3 mm × 10 mm were placed on the epoxy-coated surface and pressed firmly to ensure even bonding. It was ensured that there were no air bubbles and that the sheet was in close contact with the concrete surface. The bonded system was left to cure before testing.



X. EXPERIMENTAL RESULTS AND DISCUSSION

Single-point loading was applied at the mid-point of the reinforced concrete beam, with the aim of evaluating its maximum flexural behavior under concentrated load. Single-point loading is normally used to determine the bending behavior of the beam, as well as its maximum load carrying capacity.

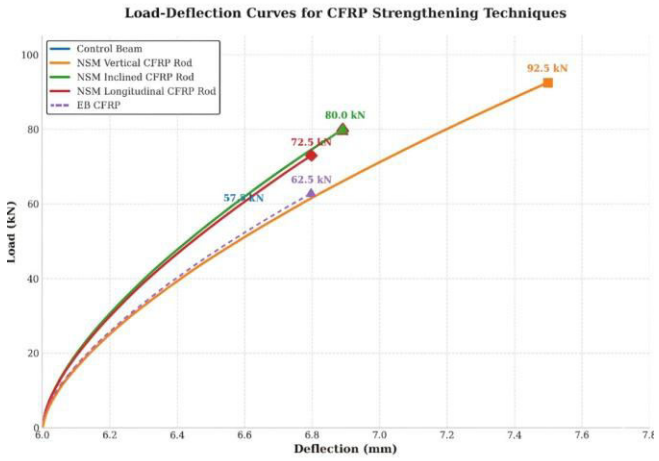
SINGLE POINT LOADING		
SPECIMEN	**LOAD (KN)**	**FLEXURAL STRENGTH (N/MM^2)**
CONTROL BEAM	43	16.13
NSM: SINGLE CFRP ROD	50	18.75
NSM: LONGITUDINAL CFRP RODS	52.5	19.68
NSM: DIAGONAL CFRP RODS	50	18.75
EB: CFRP SHEET	47.5	17.81



TWO POINT LOADING

Two-point loading was employed to ensure that a region of constant bending moment existed between the two points of loading, which would facilitate a better assessment of the beam's flexural rigidity without the effects of shear forces at the midspan region.

TWO POINT LOADING		
SPECIMEN	**LOAD (KN)**	**FLEXURAL STRENGTH (N/MM^2)**
CONTROL BEAM	57.5	16.13
NSM: SINGLE CFRP ROD	75	18.75
NSM: LONGITUDINAL CFRP RODS	92.5	23.13
NSM: DIAGONAL CFRP RODS	80	20
EB: CFRP SHEET	62.5	15.63



XI. FUTURE SCOPE OF THE PROJECT

Investigation under Cyclic and Fatigue Loading: Future studies can examine the behavior of strengthened beams under repeated or dynamic loading conditions.

Use of Different CFRP Configurations: Various orientations, spacing, and number of CFRP rods can be studied to determine the most efficient strengthening configuration.

Digital Image Correlation for Crack Analysis: Advanced techniques such as digital image correlation can be used for accurate crack propagation and strain analysis.

Long-Term Durability Studies: The long-term performance of CFRP strengthening under environmental effects such as moisture, temperature variation, and corrosion can be investigated.

Application to Full-Scale Structural Elements: Future research can extend the study to full-scale beams, slabs, and bridge elements to evaluate real-life structural behavior.

XII. CONCLUSION

In this study, the effectiveness of the various Carbon Fiber Reinforced Polymer (CFRP) strengthening techniques for reinforced concrete beams was investigated, focusing particularly on the Near Surface Mounted (NSM) and Externally Bonded (EB) techniques. The condition of the selected structural member was first evaluated using non-destructive tests such as the rebound hammer test, ultrasonic pulse velocity test, and the electrical resistivity test, which provided useful information regarding the in-situ quality and durability of the concrete.

As revealed through the experimental investigation of the control and strengthened beam specimens, the use of CFRP has been observed to have a positive impact on the performance of reinforced concrete beams. The two techniques employed in this study, the NSM and EB techniques, were observed to have enhanced the flexural strength and stiffness of the beam compared to the control specimen.

The results revealed that the NSM technique offers better bond performance and increased resistance to loads due to the embedding of CFRP rods into the concrete cover. This technique ensured better anchorage and minimized the risk of debonding. Even though the EB technique is easier to apply, it is more prone to debonding on the surface and environmental conditions.

Overall, it is clear that CFRP-based techniques are efficient solutions to repair deteriorated reinforced concrete structures. Among all techniques, the NSM technique is more promising to strengthen deteriorated RC beams.

ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

[1] Abdullahi G. Saad; Mohammed A. Sakr; Tarek M. Khalifa; Esam A. Darwish; et al(2025). ”CFRP STRENGTHENING OF RC BEAMS WITH OPENINGS UNDER IMPACT LOADS”, International Journal of Concrete Structures and Materials,vol. 19, article 95

[2] B. Wu, Q. Xia, Y. Gong, S. Fu, H. Wang, and Z. Guo (2024),”EVALUATION OF THE STRENGTHENING EFFECTS ON PRESTRESSED CFRP-STRENGTHENED STEEL BEAM BRIDGES USING MACRO-STRAIN INFLUENCE LINES”, Buildings, vol. 14, p. 2535

[3] Jawed Qureshi (2023), ”A REVIEW OF FIBER REINFORCED POLYMER BRIDGES”, Fibers, vol. 11, no. 5

[4] Hosseini, A., Motavalli, M., and Ghafoori .E (2023), ”SMART PRESTRESSED CFRP SYSTEMS FOR SUSTAINABLE METALLIC BRIDGE REHABILITATION : LONG- TERM FIELD PERFORMANCE AND MONITORING ”, Construction and Building Materials, 402, 133083

[5] J. W. Schmidt, J. D. Sørensen, and C. O. Christensen (2022), ”IN SITU CONCRETE BRIDGE STRENGTHENING USING DUCTILE ACTIVATED NSMR CFRP SYSTEMS”,Buildings, vol. 12, no. 12, p. 2244

[6] R. Thamrin, Z. Zaidir, and S. Desharma (2021), ”DEBONDING FAILURE ANALYSIS OF REINFORCED CONCRETE BEAMS STRENGTHENING WITH CFRP PLATES”, Polymers, vol. 13, no. 16, p. 2738

[7] Saracino R, Nanni A, De Lorenzi L. (2007). ”BOND STRENGTH OF FRP STRIP TO CONCRETE JOINTS INSTALLED CLOSE TO THE SURFACE”. Journal of Composites for Construction, 11(4), 401.

[8] Subramanian K, Kalpana V. G, (2024). ”FLEXURAL BEHAVIOUR OF CONCRETE BEAMS REINFORCED WITH EXTERNALLY BONDED FABRIC AND AFRP RODS NEAR THE SURFACE”. Journal of Structural Engineering, 40(2), 163168.

[9] T. Siwowski, B. Piatek, P. Siwowska, and A. Wiater (2020), ”DEVELOPMENT AND IMPLEMENTATION OF CFRP POST-TENSIONING SYSTEM FOR BRIDGE STRENGTHENING”,Engineering Structures, vol.207,p.110266

[10] Zhang S.S, Yu T (2017). ”BOND BEHAVIOUR IN NSM-FRP SYSTEMS”. Construction and Building Materials, School of Civil, Mining and Environmental Engineering, University of Wollongong, Australia

your paper may result in your paper not being published.

TRAFFIC CONGESTION MITIGATION AND SIGNAL OPTIMISATION IN PALAKKAD TOWN

*

Aiswarya K[1], Ajanya R[2], Hrithika R[3], Niranjan RP[4]
Department of Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
aiswaryak395@gmail.com,ajanya003@gmail.com,
rhrithika.r@gmail.com,nzniranjan31@gmail.com

Jayakrishnan R[5]
Assistant Professor, Department of Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
jayakrishnan.r@ahalia.ac.in

***Abstract*—With the focus being on improving the flow of traffic and reducing the number of delays that occur on the road, the study aims to evaluate the traffic congestion and the performance of the traffic signals in Palakkad. With the use of maps and newspaper articles, the study area was identified as the Kalmandapam Junction, which is known for its traffic congestion. To evaluate the congestion that occurs during the peak hour and the flow of traffic at the junction, various surveys were conducted. To suggest ways of reducing the traffic and improving the flow of traffic in the urban area, the data collected is used for simulation using PTV Vissim.**

***Index Terms*—Traffic Congestion, Signal Optimization, PTV VISSIM, Intersection Analysis, Delay Analysis.**

I. INTRODUCTION

Rising traffic congestion has become an unavoidable issue in rapidly urbanizing cities across the world. In India, cities and towns like Palakkad are also experiencing increasing traffic related challenges due to growing vehicle populations and limited road infrastructure. One of the most severely affected areas in Palakkad is the Kalmandapam Junction, which is a key intersection witnessing heavy traffic congestion, especially during peak hours. This congestion causes significant delays, energy consumption, environmental pollution, and commuter stress. Traffic delay refers to the additional travel time expe rienced by road users due to interruptions in smooth traffic f low. These delays contribute to economic losses, particularly when assessed through the lens of opportunity cost—the value of productive time lost in traffic. Passenger occupancy (the average number of people per vehicle), value of travel time (monetary worth of time spent traveling), and traffic volume (the number of vehicles passing a section of road per unit time) are all critical parameters used to analyze congestion impacts. The economic consequences of traffic congestion at Kalmandapam Junction are significant, especially consid ering its role as a major link in Palakkad's road network. Understanding these impacts in monetary terms is essential for planning effective traffic management solutions. This study aims to evaluate the congestion cost at Kalmandapam Junction by analyzing delay patterns, traffic volume, occupancy rates, and travel time values under prevailing traffic conditions.

A. AIM

The aim is to analyze traffic congestion issues in Palakkad town, evaluate existing traffic signal performance, and review modern techniques—particularly simulation-based signal op timization—for improving traffic flow, reducing delays, mini mizing fuel wastage,and enhancing overall urban mobility and road safety.and designing a traffic signal based on the data and simulating it using PTV Vissim software.

B. OBJECTIVES

1) To determine the opportunity cost of different modes of transport along the specified study area.
2) Designing a traffic signal.
3) To evaluate the performance of the traffic signal after simulating it using PTV vissim software.

II. LITERATURE REVIEW

The review of existing literature highlights significant advancements in the field of traffic signal optimization and congestion management. Singh and Sharma (2019) utilized the VISSIM simulation platform to analyze and optimize traffic signal timings at congested intersections. Their study demonstrated that simulation-based optimization can effectively reduce delays by more than 30 percentage,emphasizing the importance of traffic modeling tools in improving urban traffic f low. Similarly, Mishra and Jaiswal (2020) explored the role of Intelligent Transport Systems (ITS) in congestion control, fo cusing particularly on adaptive traffic signaling. Their research revealed that ITS-based adaptive solutions enable real-time traffic regulation at intersections, thereby improving safety, minimizing delays, and reducing fuel consumption, making it a sustainable approach to urban transport challenges. In another study, Ramesh and Rajkumar (2018) conducted a case study on traffic management in Tier-2 Indian cities. Their evaluation revealed several traffic bottlenecks and limitations in the existing system, and they proposed planning-based

signal improvements to address these issues, highlighting the necessity of context-specific management strategies suited to Indian infrastructure and travel behavior. Deshmukh and Kulkarni (2017) also contributed to this field by investigating traffic signal design and coordination across urban corridors. Using simulation tools, they found that synchronized and coordinated traffic signal control significantly improved corridor performance, reduced delays, enhanced safety, and decreased vehicular emissions caused by stop-and-go traffic. Taken together, these studies reinforce the importance of integrating simulation-based approaches, ITS solutions, and signal coordination in developing efficient traffic management systems. The collective outcomes underscore that while simulation tools like VISSIM provide measurable improvements in operational efficiency, ITS and adaptive signaling ensure real time responsiveness, and context-specific case studies reveal gaps that can be addressed through strategic planning. This integrated perspective provides a strong foundation for developing resilient, efficient, and sustainable traffic management solutions in rapidly urbanizing environment.

III. METHODOLOGY

A. General

The methodology adopted in this study provides a structured and scientific framework for analyzing urban traffic congestion and exploring effective signal optimization strategies. Urban intersections act as critical nodes in the transport network, where imbalances in demand and supply often lead to severe traffic delays, fuel wastage, and environmental pollution. Addressing these issues requires a stepwise approach that begins with understanding existing research and standards, continues with systematic site selection and data collection, and progresses through detailed analysis, redesign, and simulation of improved scenarios. The methodological process was designed not only to diagnose the current problems but also to provide realistic and implementable solutions that can be applied to the study area and extended to similar urban contexts. By combining traditional engineering methods such as traffic surveys and signal analysis with modern tools like VISSIM simulation, this study integrates both empirical evidence and advanced modeling techniques. Such an approach ensures that the results are technically sound, scientifically validated, and practically beneficial for traffic authorities and urban planners.Furthermore, the methodology places equal emphasis on theoretical grounding and practical application. A comprehensive literature review served to identify gaps in existing research and to provide guidance on best practices in traffic management, while the simulation-based validation allowed for a performance-based comparison between existing and proposed conditions. This combination strengthens the reliability of the outcomes and ensures that the recommended improvements are not only efficient but also feasible for real-world implementation .In summary, the methodology was structured to provide a holistic understanding of the traffic congestion problem, identify practical redesign strategies, and validate them through simulation, thereby ensuring that the findings contribute meaningfully to both academic knowledge and applied traffic engineering practices.

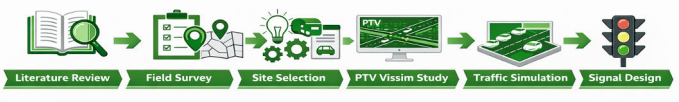


Fig. 1. methodology

IV. DATA ANALYSIS

A. KEY JUNCTION

A survey was conducted among road users to assess traffic conditions, signal efficiency, and waiting times at key junctions in Palakkad town. Data from about 192–198 commuters across Kalamandapam, IMA Junction (Fort Maidan), and Thanavu–Olavakkode showed that Kalamandapam had the highest traffic concern, with 119 respondents (62%). IMA Junction followed with 64 responses (33.3%), while Thanavu–Olavakkode recorded 48 responses (25%). The results indicate that congestion is more severe at central urban junctions than at peripheral intersections.

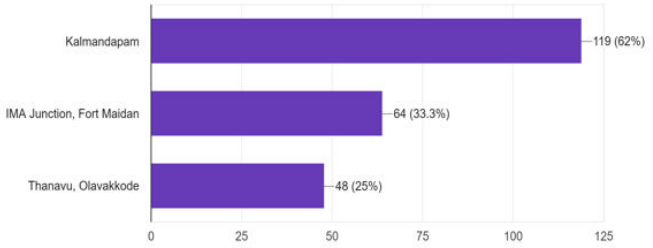


Fig. 2. Key junction

B. AVERAGE WAITING TIME AT SIGNALS/JUNCTION

The waiting time analysis showed that 52.5% of respondents face congestion at junctions without signal control. At signalized intersections, 20.7% reported waiting 1–2 minutes and 16.2% experienced 30–60 seconds of delay. Smaller groups reported less than 30 seconds (8.6%), more than 10 minutes (8.1%), and over 2 minutes even at signalized junctions (2%). These results indicate inconsistent traffic management across different junctions.

V. ROAD DETAILS AT KALMANDAPAM

Name of the road	Total width of the road	Shoulder width
Stadium - chandranagar	12.10m	1m
Kozhikode bypass	14.5m	1m
Cut road	7.10m	0.9m

According to Indian Roads Congress (IRC) specifications, the observed road widths at Kalmandapam generally comply with standard requirements for urban roads. The 12.10 m wide

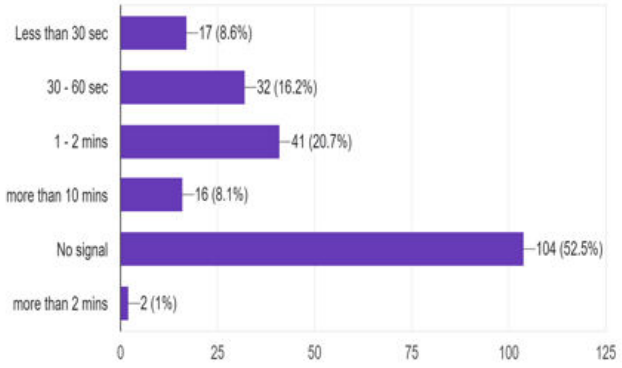


Fig. 3. Average waiting time at signal

road is adequate for a two-lane carriageway with shoulders, while the 14.5 m wide bypass road exceeds typical two-lane standards and is suitable for handling higher traffic volumes. The 7.10 m wide cut road meets the minimum requirement for a two-lane undivided road; however, its shoulder width of 0.9 m is slightly below the desirable IRC standards, which may impact safety and vehicle movement.

VI. VISSIM SOFTWARE

A. CONFLICT AREA

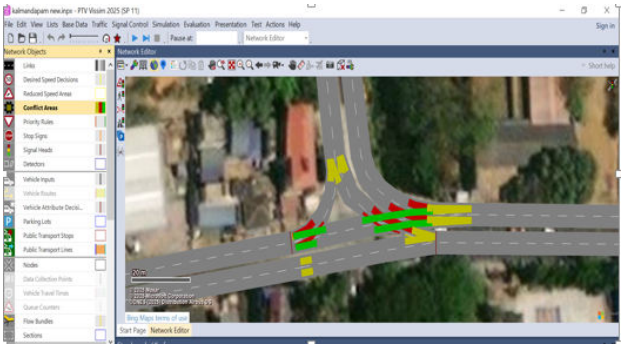


Fig. 4. Conflict area

The image shows the conflict areas defined in PTV Vissim for the Kalmandapam junction model, highlighted as yellow shaded regions. These areas are created where vehicle paths intersect, such as at merging, crossing, and turning points. They regulate vehicle interactions at unsignalized sections by allowing movements based on priority rules and gap acceptance, simulating realistic yielding and merging behavior.

VII. WEBSTER'S METHOD

The method determines the optimal signal timing by first estimating total lost time based on the number of phases and clearance intervals. It then evaluates traffic demand using flow ratios derived from vehicle volumes and saturation flow, which are summed to represent overall intersection load. Using these values, the optimum cycle length is calculated, and the effective green time is proportionally distributed among phases according to their traffic demand.

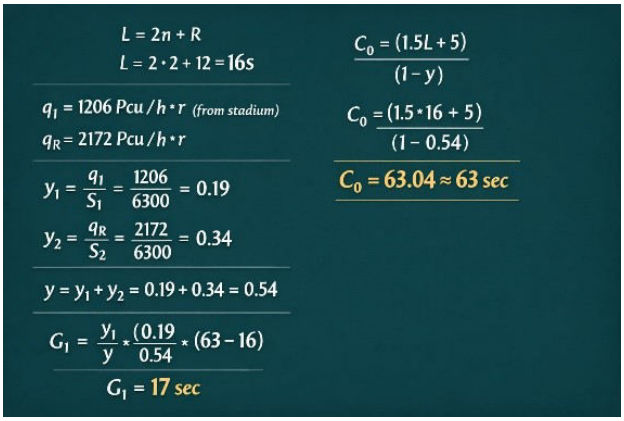


Fig. 5. Webster method

VIII. SIGNAL DESIGNING

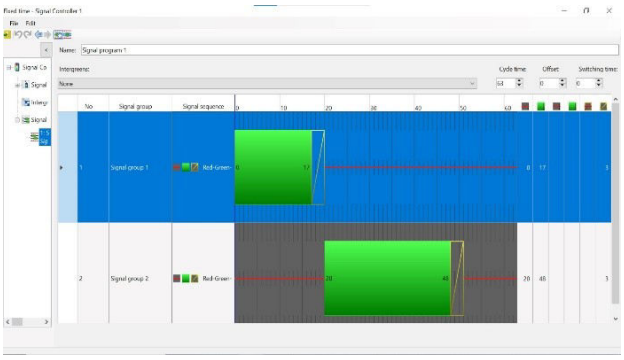


Fig. 6. Signal timimgs

Vissim with a 63-second cycle and two signal groups operating alternately. Signal Group 1 starts first, followed by Signal Group 2 in the next phase, each with green, yellow, and red intervals. This predefined sequence helps manage and analyze traffic flow at the junction.

IX. VEHICLE TRAVEL TIME RESULT

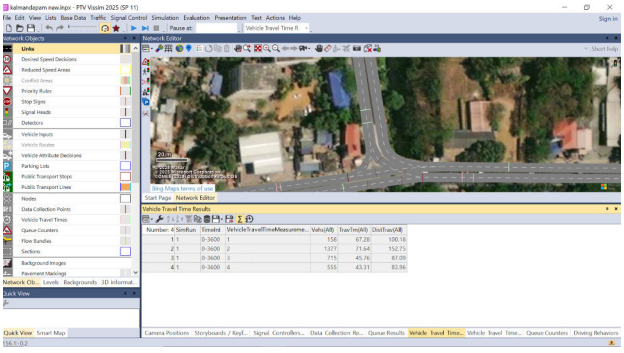


Fig. 7. Result analysis

A. Delay Analysis

The delay at the study junction was evaluated using simulation results obtained from PTV Vissim. Vehicle travel time data collected from the simulation was used to estimate the average delay experienced by vehicles under signalized conditions. A free flow travel time of 30 seconds was considered based on typical urban intersection conditions.

TABLE I
DELAY CALCULATION FROM VISSIM RESULTS

Route	Travel Time (sec)	Delay (sec)
1	67.28	37.28
2	71.64	41.64
3	45.76	15.76
4	43.31	13.31

$$\text{Average Delay} = \frac{37.28 + 41.64 + 15.76 + 13.31}{4} = 27\,\text{sec}$$

The delay at the junction was evaluated using simulation results obtained from PTV Vissim. Based on the vehicle travel time data, the average delay after signal optimization was calculated as 27 seconds by considering a free flow travel time of 30 seconds.

Under existing conditions, the junction experiences significantly higher delays ranging from 120 to 200 seconds during peak hours due to the absence of proper signal control and unorganized traffic flow. For analysis, the average existing delay is taken as 160 seconds.

A comparison of the results shows that the implementation of optimized signal timing has considerably improved the performance of the junction. The percentage reduction in delay is calculated as:

$$\text{Delay Reduction} = \frac{160 - 27}{160} \times 100 = 83.13\%$$

The implementation of optimized signal timing using VISSIM has led to a considerable reduction in delay, with an overall decrease of nearly 84%. This demonstrates the effectiveness of simulation-based signal design in improving traffic flow, reducing congestion, and enhancing the operational efficiency of the junction.

X. CONCLUSION

This project systematically evaluated traffic congestion at Kalmandapam Junction in Palakkad through detailed traffic volume surveys and peak hour analysis. The study identified that the absence of proper signal control, combined with high vehicular demand and mixed traffic conditions, significantly contributes to delays, queue formation, and operational inefficiencies, particularly during evening peak hours. The application of PTV Vissim enabled realistic modeling of existing traffic conditions and the testing of optimized signal timing strategies. Based on simulation results, the introduction of a fixed-time signal system led to a substantial reduction in average delay and improved overall traffic performance at the junction. The results clearly indicate that delay has been reduced by approximately 84%, demonstrating the effectiveness of signal optimization in minimizing congestion and improving travel time reliability. In addition, the proposed system is expected to reduce fuel consumption, lower vehicular emissions, and enhance road safety by minimizing conflict points and unregulated movements.Furthermore, this study highlights the importance of adopting simulation-based approaches in urban traffic management, as they allow planners to evaluate multiple scenarios without real-world implementation risks. The methodology used in this project can be extended to other congested junctions in urban areas for efficient traffic planning and control.

In conclusion, the implementation of optimized traffic signals at Kalmandapam Junction is a practical and effective solution for improving traffic flow, reducing delays, and promoting sustainable and safe transportation systems.

REFERENCES

[1] JAMAL, A., TAUHIDUR RAHMAN, M.. AL-AHMADI, H. M., ULLAH, I. ZAHID, M. 2020. Intelligent Intersection Control for Delay Optimization: Using Meta-Heuristic Search Algorithms. Sustainability, 12, 1896

[2] MAZAHERI, A. RAHIMI, A. M. Determining a suitable position for U-turns near signal-controlled road junctions. Proceedings of the Institution of Civil Engineers-Transport, 2017. Thomas Telford Ltd. 267-275

[3] JONI, H. H. 2017. Traffic performance for Intersection by Using Synchro Software of Traffic Flow in AdEn Square.

[4] UMALA, T., QUEZON, E. SHIFERAW, B. 2016. Effect of Cycle Time and Signal Phase on Average Time Delay, Congestion and Level of Service: A Case Study at Hager Astedader Signalized Intersection in Addis Ababa. International Journal of Scientific and Engineering Research, 7, 524-530

[5] JONI, H. HIKMATT, M. Assessment at Al-Ameer signalized intersection in Samawa city. MATEC Web of Conferences, 2018. EDP Sciences, 01039.

[6] GERESU, T. G. 2017. Optimization of Intersection Signal Timing -Case Study of Shola Gebeya and Estifanos Stadium Laghar Intersection. MSc, Addis Ababa University. Thesis.

[7] FHWA 2018. FHWA. Federal Highway Administration Traffic Data Computation Method-Pocket Guide. Department of Transportation, United States

[8] ABATE, T. G. 2018. Evaluating the Performance of Signalized Intersection and the Associated Economic Impact of Congestion: (A Case Study on Ras Mekonnen Street of Addis. Ababa, Ethiopia). Thesis.

[9] AL-MARAFI, M. N., AL-MARAFI, D. H., AL-JARABHA, M. S. AL-SHBEILAT, H. S. M.. AHED A 2021. Evaluation and Improvement of Traffic Flow on Main Intersections in Tafila City Using Synchro

[10] AL-OBAIDI .,AL-SADIK, S. AL-MANDELAWEEH. Optimizing the Traffic Operation System of 14th-Ramadhan. Signalized Intersection in Baghdad City using HCS Technique. IOP Conference Series: Materials Science and Engineering, 2020.

[11] BARA'W, A.-M., ALOMARI, A. H. AL ZOUBI, M. S. 2021. Investigation of saturation flow rate using video camera at signalized intersections in Jordan. Open Engineering.

[12] DOGAN, E., PAYIDAR AKGUNGOR A., ARSLAN, T. 2016. Estimation of delay and vehicle stops at signalized intersections using artificial neural network.

[13] DULEBENETS, M. A., RAHIMI, A. M. MAZAHERI, A. 2017. Assessing the effects of indirect left turn on a signalized intersection

[14] MUHAMMED, H. H. YOUSIF, M. A. 2021. Enhancement of Al-Rawad intersection by using synchro traffic simulation software: As case study. Materials Today: Proceedings

[15] MSALLAM, M., ABOJARADEH, M., JREW, B. AL-ALLAFF, R. A. 2016. Evaluation Of Traffic Flow And Traffic Network Management System In Jordan. International Journal of Civil Engineering. Construction and Estate Management, 4, 1-21

[16] OMER, B. R. KHALID, S. W. 2018. An Improvement of Three Leg Signalized Intersections in Duhok City Kurdistan-Iraq. Academic Journal of Nawroz University, 7, 87-92
[17] TAHA, M. ABDELFATAH, A. 2017. Impact of using indirect left-turns on signalized intersections' performance. Canadian Journal of Civil Engineering, 44, 462-471.
[18] ZHANG, X., FAN, X., YU, S., SHAN, A., FAN, S., XIAO, Y. DANG, F. 2022. Intersection Signal Timing Optimization: Multi-Objective Evolutionary Algorithm. Sustainability, 14, 1506.
[19] ZIBOON, A. -R. T., QASIM, Z. I. YOUSIF, M. A. 2019. Traffic Performance Evaluation and Analysis of Al-Fallah Intersection in Baghdad
[20] YAO, R., ZHOU, H. GE, Y.-E. 2018. Optimizing signal phase plan, green splits and lane length for isolated signalized intersections. Transport, 33, 520-535
[21] SOFIA, G., AL-HADDAD, A. AL-HAYDARI, I. S. Improvement of traffic performance at intersections in Karbala city. MATEC Web of Conferences, 2018.

Defluoridation Of Water Using Electrocoagulation

Sreeraghav S
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
sreeraghavmudra916@gmail.com

Bhavana Prahladh B
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
getbhavanapkd@gmail.com

Thejaswi S
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
thejaswisasikumar29@gmail.com

Neethu John
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
neethu.john@ahalia.ac.in

B V Mathew
Civil Engineering
Ahalia School of Engineering and Technology
Palakkad, India
hodce.aset@ahalia.ac.in

Abstract **- Excess fluoride in drinking water is an important environmental and public health concern in many parts of the world, and can lead to dental and skeletal fluorosis. The defluoridation of drinking water is sometimes necessary to prevent excessive fluoride exposure. This research work investigates the removal of fluoride ions from water using electro-coagulation with aluminium electrodes(plates). In this case, a synthetic fluoride stock solution was prepared by dissolving the required amount of fluoride in 1000 ml of distilled water to get a 5 ppm solution. The experiment was carried out in an electro-coagulation system composed of a reactor tank, two aluminum plate electrodes, and a direct current (DC) power supply. The applied voltage (15 V, 20 V, and 24 V), spacing between the electrode plates (1.5 cm, 2.0 cm, and 3.0 cm), and electrolysis time (30, 45, and 60 min) were varied. In the test of fluoride removal by the system, it was found that fluoride removal increased as the current voltage and electrolysis time increased. The efficiency of fluoride removal was greater when the distance between the electrodes was shorter, which was helpful for the treatment. The optimal electro-coagulation operating condition for maximum fluoride removal was 24 V applied voltage, 1.5 cm gap between the electrodes, and 60 min electrolysis time. This indicates that electro-coagulation is an effective and a simple defluoridation process for water treatment.**

***Keywords*— Defluoridation, Electro-coagulation, Aluminium plates, Fluoride removal, Applied voltage, Electrode spacing, Treatment time**.

I. INTRODUCTION

Water is an essential resource for human survival, and its quality plays a vital role in maintaining public health. However, contamination of drinking water with various pollutants has become a major environmental issue worldwide. One of the most common contaminants found in groundwater is fluoride. Fluoride occurs naturally in water due to the dissolution and leaching of fluoride-bearing minerals from rocks and soil. While low concentrations of fluoride are beneficial for preventing dental cavities, excessive fluoride intake can lead to serious health problems such as dental fluorosis and skeletal fluorosis. Fluoride is a natural element found in many groundwater sources because of the weathering and leaching of fluoride-bearing minerals According to the guidelines of the World Health Organization, the permissible limit of fluoride in drinking water is about 1.5 mg/L. When the fluoride concentration exceeds this limit, it becomes harmful to human health. Many regions around the world, especially areas dependent on groundwater sources, experience elevated fluoride levels, which necessitates the development of efficient defluoridation techniques.

There are several methods for removing fluoride, including adsorption, ion exchange, membrane filtration, and chemical precipitation. Among these, electrocoagulation has attracted attention due to its simplicity, effectiveness, and lower need for chemicals. In electrocoagulation, sacrificial metal electrodes, usually made of aluminum or iron, dissolve when an electric current passes through them. This process produces coagulant species that help remove contaminants from water.

In this study, the electrocoagulation process using aluminium electrodes was investigated for the removal of fluoride from water. A synthetic fluoride solution of 5 ppm was prepared using distilled water, and experiments were conducted by varying operational parameters such as applied voltage, electrode spacing, and electrolysis time. The objective of this study is to evaluate the efficiency of electrocoagulation in fluoride removal and to determine the optimum operating conditions for effective defluoridation.

II. MATERIALS

A. Chemicals and Reagents

In this study, a synthetic fluoride solution was prepared using Sodium Fluoride (NaF) as the fluoride source. Sodium fluoride is commonly used in laboratory experiments because it is highly soluble in water and provides a reliable source of fluoride ions for experimental studies. A fluoride concentration of 5 ppm was prepared by dissolving the required quantity of sodium fluoride in 1000 ml of distilled water**.** The solution was mixed thoroughly to ensure uniform distribution of fluoride ions throughout the water. Distilled water was used to eliminate the presence of other impurities that could influence the electrocoagulation process or affect the experimental results. No additional chemicals or additives were used in the preparation of the synthetic solution.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

B. Electrodes

Aluminium plates were used as electrodes in the electrocoagulation process. Two aluminium plates acted as the anode and cathode during the electrolysis process. Each plate had dimensions of 8 cm × 6 cm × 1 mm thickness**.** Aluminium is widely used in electrocoagulation because it has excellent coagulation properties and produces aluminium hydroxide species when subjected to an electric current. These hydroxides form flocs that help in removing contaminants from water through adsorption, precipitation, and aggregation mechanisms. Before each experiment, the electrode surfaces were cleaned with distilled water to remove any impurities and to ensure consistent experimental conditions.

C. Electrocoagulation setup

The electrocoagulation experiments were conducted in a laboratory-scale reactor tank made of non-reactive material. The reactor had dimensions of 20 cm × 15 cm × 20 cm, which provided sufficient volume to carry out the treatment process. The tank was filled with the prepared fluoride solution, and the aluminium electrodes were immersed vertically in the solution. Proper positioning of electrodes inside the tank was maintained to ensure uniform current distribution during the electrocoagulation process. A DC power supply was used to provide the required electrical energy for the electrocoagulation process. The power supply was connected to the aluminium electrodes using insulated connecting wires. Different voltages were applied during the experiments to evaluate the influence of electrical potential on fluoride removal efficiency.The voltages applied in this study were 15 V, 20 V, and 24 V. The power supply allowed control over the voltage applied across the electrodes during the electrolysis process. Basic laboratory measuring tools were used during the experimental work. A measuring scale was used to maintain accurate electrode spacing (1.5 cm, 2 cm, and 3 cm) between the aluminium plates. Measuring cylinders and containers were used to prepare and handle the fluoride solution. These materials and equipment together formed the electrocoagulation system used for the defluoridation experiments.

III. METHODOLOGY

A. Preparation of Synthetic Fluoride Solution

The synthetic fluoride solution used in this study was prepared according to the standard procedure described in IS 3025 (Part 60): 2008, which provides guidelines for the determination and preparation of fluoride solutions in water analysis. The preparation involved two stages: preparation of a fluoride stock solution and preparation of a working standard solution.

- *Preparation of Fluoride Stock Solution*

For the preparation of the fluoride stock solution, Sodium Fluoride (NaF) was used as the fluoride source. Initially, the sodium fluoride salt was dried in a hot air oven at 150 °C for 4 hours in order to remove any moisture present in the sample. After drying, the NaF was allowed to cool in a desiccator to prevent the absorption of moisture from the surrounding atmosphere.After cooling, 2.210 g ± 0.001 g of dried sodium fluoride was accurately weighed using an analytical balance. The weighed NaF was then dissolved in a small quantity of distilled water and transferred carefully into a 1000 mL volumetric flask. The flask was then filled up to the calibration mark using distilled water and the solution was mixed thoroughly to obtain a uniform fluoride stock solution.

- *Preparation of Fluoride Working Standard Solution (5 mg/L)*

From the prepared stock solution, a working standard solution was prepared for the electrocoagulation experiments. Using a pipette, 5 mL of the fluoride stock solution was measured accurately and transferred into another 1000 mL volumetric flask**.**The flask was then filled with distilled water up to the 1000 mL mark and mixed thoroughly to ensure uniform distribution of fluoride ions in the solution. This procedure resulted in the preparation of a fluoride working standard solution with a concentration of 5 mg/L (5 ppm)**.**The prepared 5 mg/L fluoride solution was used as the synthetic contaminated water for conducting the electrocoagulation treatment experiments. This solution ensured a controlled fluoride concentration for evaluating the efficiency of the Electrocoagulation process under different operating conditions.

B. Experimental Setup

The electrocoagulation experiment was conducted using a laboratory-scale treatment system designed to evaluate the removal of fluoride from water. The experimental setup mainly consisted of an electrocoagulation reactor tank, aluminium plate electrodes, a direct current (DC) power supply, and connecting wires. The setup was arranged in such a way that controlled electrical energy could be supplied to the electrodes immersed in the fluoride solution to initiate the electrocoagulation process. The electrocoagulation reactor used in this study was a rectangular tank with dimensions of 20 cm × 15 cm × 20 cm**,** which served as the treatment chamber for the fluoride solution. The reactor was filled with 1000 mL of the prepared 5 mg/L fluoride solution**.** The tank material was non-conductive and chemically stable to prevent interference with the electrochemical reactions occurring during the experiment.

Two aluminium plates were used as electrodes in the electrocoagulation cell. Each electrode had dimensions of 8 cm × 6 cm with a thickness of 1 mm. One plate functioned as the anode**,** while the other acted as the cathode**.** The electrodes were placed vertically inside the reactor and positioned parallel to each other to ensure uniform current distribution throughout the solution. The distance between the electrodes was adjusted according to the required electrode spacing during the experiments. The electrodes

were connected to a DC power supply using insulated copper wires. The power supply allowed the application of different voltages across the electrodes to study the effect of electrical potential on fluoride removal efficiency. When voltage was applied, an electric current passed through the solution, initiating electrochemical reactions at the electrode surfaces.

At the anode, aluminium metal undergoes oxidation and dissolves into the solution in the form of aluminium ions. These ions subsequently react with hydroxide ions present in the water to form gelatinous aluminium hydroxide species. These aluminium hydroxide flocs act as effective coagulants that can remove fluoride ions from water through mechanisms such as adsorption, charge neutralization, and precipitation. At the cathode, reduction reactions occur, resulting in the generation of hydrogen gas bubbles. These gas bubbles rise through the solution and assist in the flotation of the formed flocs, helping them to separate from the treated water. The combined action of coagulation, adsorption, precipitation, and flotation enhances the overall efficiency of the electrocoagulation process in removing fluoride ions from the solution.

The experimental setup enabled the systematic variation of important operational parameters such as applied voltage, electrode spacing, and electrolysis time. This allowed the evaluation of the performance of the Electrocoagulation process and the identification of optimum conditions for effective fluoride removal from water.

C. Operating Parameters

The efficiency of the Electrocoagulation process depends on several operating parameters that influence the rate of electrochemical reactions and the formation of coagulant species within the reactor. These parameters directly affect the dissolution of electrode material, generation of aluminium hydroxide flocs, and the overall removal efficiency of fluoride from water.In this study, three major operating parameters were investigated: **applied voltage, electrode spacing, and electrolysis time**. The experiments were carried out in a batch electrocoagulation reactor using aluminium electrodes**,** and the parameters were varied systematically to determine their influence on fluoride removal efficiency.

➢ Applied Voltage

Applied voltage plays a crucial role in controlling the current density and the rate of electrochemical reactions occurring at the electrode surfaces. When voltage is applied across the electrodes, aluminium dissolves from the anode and releases aluminium ions into the solution, which subsequently form aluminium hydroxide flocs responsible for contaminant removal.Experiments were conducted at three different applied voltages:

- **15 V**
- **20 V**
- **24 V**

Higher applied voltage generally increases the rate of aluminium dissolution and enhances the formation of coagulant species, thereby improving the efficiency of fluoride removal.

➢ Electrode Spacing

Electrode spacing refers to the distance between the anode and cathode plates inside the electrocoagulation reactor. The spacing between electrodes affects the electrical resistance of the solution and the distribution of electric current in the reactor.The following electrode spacings were investigated in this study:

- **1.5 cm**
- **2 cm**
- **3 cm**

Smaller electrode spacing reduces electrical resistance and allows better current flow through the solution, which improves the electrocoagulation process. However, appropriate spacing is required to maintain proper mixing and avoid operational issues.

➢ Electrolysis Time

Electrolysis time is the duration for which the electric current is applied during the electrocoagulation treatment. It determines the amount of aluminium ions released and the quantity of aluminium hydroxide flocs generated in the solution.The electrocoagulation process was carried out for the following treatment durations:

- **30 minutes**
- **45 minutes**
- **60 minutes**

Increasing the electrolysis time generally increases the production of coagulant species, which enhances the removal of fluoride ions from the water.

➢ Experimental Operating Conditions

To study the combined effect of these parameters, experiments were conducted by varying **voltage, electrode spacing, and electrolysis time simultaneously**. For each applied voltage, experiments were performed at all electrode spacings and treatment times.Thus, the electrocoagulation experiments were carried out under the following combinations:

- Voltage: **15 V, 20 V, 24 V**
- Electrode spacing: **1.5 cm, 2 cm, 3 cm**
- Electrolysis time: **30 min, 45 min, 60 min**

This resulted in multiple experimental runs to evaluate the performance of the electrocoagulation process under different operating conditions.

➢ Post-Treatment Procedure

After the completion of each electrocoagulation experiment, the treated solution was allowed to settle for a certain period to facilitate the separation of the coagulated flocs from the water. The supernatant liquid was then filtered using filter paper to remove the suspended particles and precipitated flocs. The filtered water sample was collected and used for further analysis to determine the remaining fluoride concentration. The results obtained under different operating conditions were compared in order to determine the optimum parameters for effective fluoride removal.

D. Experimental Procedure

The electrocoagulation experiments were conducted in a laboratory-scale batch reactor to evaluate the efficiency of fluoride removal from water. At the beginning of each experimental run**,** 1000 mL of the prepared 5 mg/L (5 ppm) synthetic fluoride solution was poured into the electrocoagulation reactor tank. Two aluminium plate electrodes with dimensions 8 cm × 6 cm × 1 mm thickness were placed vertically inside the reactor and positioned parallel to each other. The distance between the electrodes was adjusted according to the required electrode spacing (1.5 cm, 2 cm, or 3 cm) for the specific experimental condition. The electrodes were then connected to a DC power supply using insulated copper wires, ensuring proper electrical contact.

Once the reactor setup was prepared, the required voltage (15 V, 20 V, or 24 V) was applied across the electrodes to initiate the electrocoagulation process. When the electric current passed through the solution, electrochemical reactions occurred at the electrode surfaces. At the anode**,** aluminium metal was oxidized and dissolved into the solution in the form of aluminium ions (Al^{3+}). These ions reacted with hydroxide ions present in water to form gelatinous aluminium hydroxide [$Al(OH)_3$] flocs, which act as coagulants in the treatment process.

The formed aluminium hydroxide flocs have a large surface area and strong adsorption capacity. These flocs adsorb fluoride ions present in the water and facilitate their removal through mechanisms such as adsorption, charge neutralization, precipitation, and aggregation**.** Simultaneously, hydrogen gas bubbles were generated at the cathode due to the reduction of water molecules. These gas bubbles helped in the flotation of the formed flocs, allowing them to rise to the surface and separate from the treated water.Each experiment was conducted for a predetermined electrolysis time of 30 minutes, 45 minutes, or 60 minutes depending on the experimental condition being tested. During this period, continuous electrochemical reactions resulted in the formation of coagulant species and the progressive removal of fluoride ions from the solution.

After the completion of the treatment time, the power supply was switched off and the electrodes were removed from the reactor. The treated solution was then allowed to settle for a certain period so that the coagulated flocs could separate from the water. The supernatant liquid was carefully collected and filtered using filter paper to remove any remaining suspended particles and coagulated sludge.The filtered water sample was then collected and analyzed to determine the remaining fluoride concentration. The difference between the initial and final fluoride concentration was used to evaluate the efficiency of the electrocoagulation process.During the Electrocoagulation process, several electrochemical reactions occur at the electrode surfaces when an electric current is passed through the fluoride-containing solution. In this study, aluminium electrodes were used, which act as sacrificial electrodes and generate coagulant species directly within the solution.

➢ Anodic Reaction

At the anode, aluminium metal undergoes oxidation and dissolves into the solution as aluminium ions:

$$\mathbf{Al\ (s) \rightarrow Al^{3+} + 3e^-}$$

The released aluminium ions subsequently react with hydroxide ions present in water to form aluminium hydroxide:

$$\mathbf{Al^{3+} + 3OH^- \rightarrow Al(OH)_3\ (s)}$$

The aluminium hydroxide formed is a gelatinous precipitate with a large surface area. These flocs act as coagulants and adsorb fluoride ions present in the water.

➢ Cathodic Reaction

At the cathode, water molecules are reduced, producing hydrogen gas and hydroxide ions:

$$\mathbf{2H_2O + 2e^- \rightarrow H_2\ (g) + 2OH^-}$$

The hydrogen gas bubbles generated during this reaction help in the flotation of suspended particles and coagulated flocs, aiding their separation from the treated water.

➢ Overall Reaction

The aluminium hydroxide formed during the electrocoagulation process interacts with fluoride ions present in the solution, resulting in the removal of fluoride through adsorption and precipitation mechanisms:

$$\mathbf{Al(OH)_3 + F^- \rightarrow Al–F\ complex\ /\ adsorbed\ fluoride}$$

Thus, the combined electrochemical reactions lead to the formation of aluminium hydroxide flocs, which effectively remove fluoride ions from water through coagulation, adsorption, and precipitation processes.

E. Determination of Optimum Conditions

To determine the most effective conditions for fluoride removal, electrocoagulation experiments were performed using different combinations of applied voltage, electrode spacing, and electrolysis time. These parameters were varied systematically in order to study their individual and combined effects on the removal efficiency of fluoride ions from water.The experiments were carried out at three different applied voltages (15 V, 20 V, and 24 V), three electrode spacings (1.5 cm, 2 cm, and 3 cm), and three electrolysis times (30 minutes, 45 minutes, and 60 minutes). For each combination of operating parameters, the electrocoagulation process was performed and the treated water sample was analyzed to determine the remaining fluoride concentration. Higher applied voltage enhanced the rate of aluminium dissolution at the anode, resulting in the production of a larger quantity of aluminium hydroxide flocs. These flocs increased the adsorption and coagulation capacity of the system. Similarly, increasing the electrolysis time allowed more aluminium ions to be released into the solution, leading to greater formation of coagulant species and improved fluoride removal. In addition, smaller electrode spacing reduced electrical resistance and allowed better current flow through the solution, which improved the efficiency of the electrocoagulation process.

Among all the tested experimental conditions, the highest fluoride removal efficiency was observed under the following parameters:

- **Applied Voltage: 24 V**
- **Electrode Spacing: 1.5 cm**
- **Electrolysis Time: 60 minutes**

These conditions were therefore identified as the optimum operating parameters for the electrocoagulation process in this study. Under these conditions, the formation of aluminium hydroxide flocs was maximized, which significantly enhanced the removal of fluoride ions from the solution.The identification of these optimum parameters provides useful information for improving the efficiency of electrocoagulation systems used for defluoridation of drinking water and can serve as a basis for further research and practical applications.

Fig 1: Voltage: 15V, Ampere: 0.503

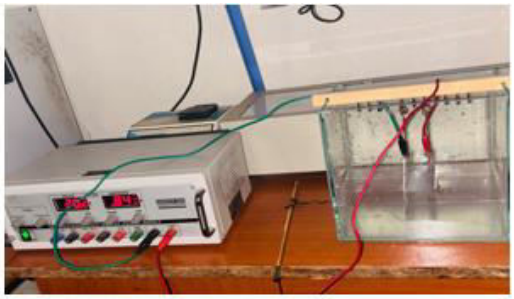


Fig:2 Voltage: 20V, Ampere: 0.843

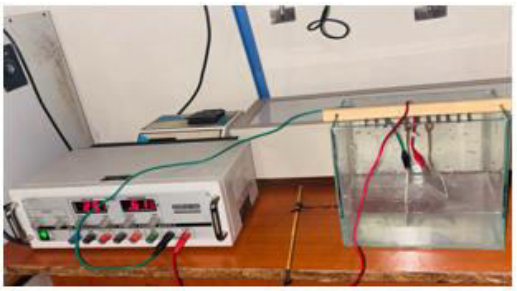


Fig 3: Voltage: 24V, Ampere: 0.613

IV. RESULT

The performance of the electrocoagulation process was evaluated under different operating conditions.

Effect of Voltage: It was observed that fluoride removal efficiency increased with increasing applied voltage. Higher voltage enhances the dissolution of aluminium electrodes and generates more coagulant species, which improves fluoride removal.

Effect of Electrode Spacing: Electrode spacing significantly influenced the treatment efficiency. Smaller spacing reduced electrical resistance and improved current flow, resulting in better coagulation. The highest removal efficiency was obtained at an electrode spacing of 1.5 cm.

Effect of Electrolysis Time: The fluoride removal efficiency increased with increasing treatment time. Longer electrolysis time allows more aluminium ions and hydroxides to form, which enhances the coagulation and adsorption of fluoride ions.

Optimum Conditions: The optimum operating parameters for maximum fluoride removal were found to be:

- Voltage: **24 V**
- Electrode spacing: **1.5 cm**
- Electrolysis time: **60 minutes**

From the preliminary experiments conducted at different voltages (15 V, 20 V, and 24 V), 24 V showed the highest fluoride removal efficiency**.** Therefore, further experiments were carried out by keeping the voltage constant at 24 V**,**

while varying the electrode spacing and electrolysis time. Electrode spacing influences the electrical resistance in the electrocoagulation reactor. Smaller spacing reduces resistance and increases current flow, which enhances the dissolution of aluminium from the anode and the formation of aluminium hydroxide flocs. These flocs adsorb fluoride ions and help in their removal from water. Electrolysis time also affects the removal efficiency. Longer treatment time allows more aluminium ions and flocs to form**,** which increases the adsorption of fluoride ions and improves the overall removal efficiency.

For these experiments, **a** synthetic fluoride solution of 5 ppm concentration was used as the sample. The electrocoagulation treatment was then performed at different electrode spacings (1.5 cm, 2 cm, and 3 cm) and electrolysis times (30, 45, and 60 minutes) to study their effect on fluoride removal efficiency.

The results obtained under these operating conditions are presented in **Table 1**

Voltage(V)	Spacing(cm)	Retention time(min)	Before treatment	After treatment	Efficiency%0
24	1.5	30	5ppm	1.12	77.6
24	1.5	45	5ppm	0.71	85.8
24	1.5	60	5ppm	0.59	88.2
24	2	30	5ppm	1.39	72.2
24	2	45	5ppm	1.20	76
24	2	60	5ppm	0.88	82.4
24	3	30	5ppm	2.02	59.6
24	3	45	5ppm	1.85	63
24	3	60	5ppm	1.13	77.4

Table 1: Fluoride Removal Efficiency at Different Electrode Spacing and Electrolysis Time at 24 V

In addition to the laboratory-prepared fluoride solution, real groundwater samples were collected from Elavancherry and Vannamada, which are known to have fluoride contamination issues.The collected samples were first tested to determine the initial fluoride concentration**.** After analysis, the samples were treated using the electrocoagulation process under the optimum operating conditions identified in this study (Voltage: 24 V, Electrode spacing: 1.5 cm, Electrolysis time: 60 minutes).After treatment, the samples were allowed to settle and then filtered using filter paper. The treated water was again analyzed for fluoride concentration to evaluate the effectiveness of the electrocoagulation process in removing fluoride from real groundwater samples.

The results of fluoride removal for the real water samples are presented in **Table 2**.

Location	Voltage(V)	Spacing(cm)	Retention time(min)	Before treatment	After treatment	Efficiency(%)
Elavancherry	24	1.5	60	1.82	0.22	87.9
Vannamada	24	1.5	60	1.7	0.21	87.6

Table 2: Operating Conditions and Fluoride Removal Efficiency

Apart from fluoride, other water quality parameters such as pH, turbidity, total hardness, calcium, magnesium, and total dissolved solids (TDS) were also analyzed before and after treatment. These parameters help in assessing the overall water quality and the effectiveness of the electrocoagulation process. The changes in these parameters indicate the improvement in water quality after treatment.

Parameters	Before treatment	After treatment	Acceptable limit as per IS 10500-2012	Permissible limit as per IS 10500-2012
Turbidity	0.3	2.1	1 NTU	5 NTU
Total Hardness	277.7	189.2	200	600
Calcium	14.65	18.50	75	200
Magnesium	51.45	35.32	30	100
pH	6.93	7.32	6.5-8.5	
Total dissolved solids	697	523	500	2000

Table 3: Analysis of Physicochemical Parameters of Elavancheery

Parameters	Before treatment	After treatment	Acceptable limit as per IS 10500-2012	Permissible limit as per IS 10500-2012
Turbidity	0.2	1.7	1 NTU	5 NTU
Total Hardness	330.7	234.6	200	600
Calcium	20.2	27.9	75	200
Magnesium	69.1	45.76	30	100
pH	6.51	7.7	6.5-8.5	
Total dissolved solids	673	520	500	2000

Table 4: Analysis of Physicochemical Parameters of Vannamada

V. CONCLUSION

The present study demonstrated that electrocoagulation using aluminium electrodes is an effective technique for the removal of fluoride from water. A synthetic fluoride solution of 5 ppm concentration was prepared and treated using a laboratory-scale electrocoagulation setup. The influence of important operating parameters such as applied voltage, electrode spacing, and electrolysis time was investigated to determine their effect on fluoride removal efficiency. The experimental results showed that the performance of the electrocoagulation process improves with increase in applied voltage and electrolysis time**,** while smaller electrode spacing enhances the current flow and coagulant formation. Among the tested conditions, the optimum operating parameters were found to be 24 V applied voltage, 1.5 cm electrode spacing, and 60 minutes electrolysis time**,** which provided the highest fluoride removal efficiency. In addition to the synthetic solution experiments, real water samples collected from Elavancherry and Vannamada were also analyzed to evaluate the applicability of the process for natural water sources. The results indicated that electrocoagulation can effectively improve water quality by reducing fluoride and other related parameters. Overall, the study confirms that electrocoagulation is a simple, efficient, and environmentally friendly technique for defluoridation of water**.** With further optimization and scale-up studies, this method has strong potential for practical application in drinking water treatment, particularly in fluoride-affected regions**.**

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my guide and supervisor for their continuous support, valuable guidance, and encouragement throughout the completion of this project titled “Defluoridation of Water Using Electrocoagulation.” Their insightful suggestions and

constructive feedback helped me successfully complete this work.

I would also like to thank the faculty members of the department for providing the necessary facilities, resources, and academic support required for this study.

REFERENCES

[1] Lakshmanan, D., Clifford, D. A., & Samanta, G. (2010). Comparative study of electrocoagulation and chemical coagulation for water treatment. *Environmental Engineering Science*, 27(3), 231–239.

[2] Yadav, K. K., Gupta, N., Kumar, V., et al. (2020). A review of emerging technologies for fluoride removal from water. *Environmental Technology & Innovation*, 20.

[3] Gupta, A., Singh, R., & Sharma, S. (2020). Electrocoagulation process for removal of fluoride from drinking water. *Water Science and Technology*.

[4] Kumar, P., & Singh, A. (2021). Electrocoagulation treatment for groundwater defluoridation using aluminum electrodes. *Environmental Engineering Research*.

[5] Sedou, M., Mande, S. L. A., Sanou, Y., & Arouna, K. (2022). Fluoride removal in drinking water by electrocoagulation using aluminum electrodes. *European Journal of Applied Sciences*, 10(4), 429–438.

[6] Kumar, R., & Singh, S. (2024). Defluoridation of water using aluminium electrode electrocoagulation. *Environmental Monitoring and Assessment*.

[7] Nair, R., & Krishnan, P. (2025). Electrocoagulation based defluoridation of groundwater: Experimental investigation. *Journal of Environmental Chemical Engineering*.

COMPOSITE GYPSUM BOARDS REINFORCED WITH AGRO-INDUSTRIAL WASTES

AKHIL M
CIVIL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
akhilmohandas034@gmail.com

AYANA S DEVAN
CIVIL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
ayanasdevan02@gmail.com

AKSHAYA KRISHNAN M
CIVIL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
chevikkalammu@gmail.com

ROHITH M
CIVIL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
rohith.m@ahalia.ac.in

AKHIL KRISHNAN M.K.
CIVIL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
akhilkrishnan.mk@ahalia.ac.in

***Abstract*—The construction industry is one of the largest consumers of natural resources and contributes significantly to environmental degradation due to excessive extraction of raw materials and generation of waste. In recent years, there has been a growing interest in developing sustainable and eco-friendly construction materials that reduce environmental impact while maintaining adequate mechanical performance. Gypsum boards are widely used in the construction industry for interior wall partitions, false ceilings, and decorative panels due to their lightweight nature, fire resistance, and ease of installation. However, conventional gypsum boards possess limitations such as brittleness, low mechanical strength, and sensitivity to moisture**.

Agro-industrial wastes such as rice husk ash, coir fibres, and waste paper are generated in large quantities in agricultural regions. Disposal of these wastes often leads to environmental problems such as land pollution and greenhouse gas emissions. Utilizing these wastes in building materials offers a sustainable solution by converting waste into valuable construction products.

This study focuses on the development of composite gypsum boards reinforced with agro-industrial wastes. Rice husk ash, coir fibres, and paper pulp were incorporated into gypsum boards to enhance mechanical properties and reduce the amount of gypsum required. The experimental procedure involved material collection, preparation, mix proportioning, moulding, and curing of composite boards. The results are expected to show improved flexural strength, better bonding, and enhanced durability compared to conventional gypsum boards. The proposed composite boards provide an eco-friendly alternative for non-structural construction applications.

***Keywords*—Gypsum board, Agro-industrial waste, Rice husk ash, Coir fibre reinforcement, Sustainable construction materials**

I. INTRODUCTION

In recent years, the construction industry has increasingly focused on the development of sustainable and high-performance materials to reduce environmental impact while enhancing structural efficiency. Gypsum-based composites have gained significant attention due to their lightweight nature, ease of fabrication, and inherent fire-resistant properties. Researchers have extensively explored the reinforcement of gypsum with agro-industrial wastes and natural fibers to improve its mechanical and durability characteristics. Studies conducted by **K. Ghavami** [1] highlighted the effectiveness of natural fibers in enhancing the strength and toughness of composite materials, particularly in low-cost construction applications. Similarly, **A. K. Bledzki** [2] demonstrated that agricultural residues can significantly improve the mechanical behavior and reduce the brittleness of composite systems.

Further investigations by **M. A. Mosallam** [3] revealed that fiber-reinforced gypsum composites exhibit improved flexural strength and crack resistance, making them more suitable for structural and semi-structural applications. In addition, **S. P. Singh** and **P. Kumar** [4] reported that the incorporation of agro-industrial waste materials not only enhances strength properties but also contributes to reduced density and improved thermal insulation. These characteristics are particularly beneficial for modern construction practices that demand lightweight and energy-efficient materials.

Moreover, research by **R. D. Toledo Filho** [5] emphasized the role of natural reinforcements in improving durability and resistance to cracking under different loading conditions. Other studies have also indicated that gypsum composites reinforced with suitable additives show enhanced fire resistance and sound absorption properties, making them ideal for building applications such as partition boards and panels. Despite these advancements, many of the existing studies primarily focus on individual property improvements, and there is a lack of comprehensive optimization of mix proportions to achieve a balanced combination of mechanical strength, durability, and functional performance.

Based on the insights obtained from the literature, it is evident that there is a need to develop an optimized composite mix that can simultaneously improve strength, reduce water absorption, enhance thermal properties, and maintain low density. Therefore, the present study focuses

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

on the development and evaluation of various composite mix ratios using gypsum reinforced with agro-industrial waste materials. Among the different mixes prepared, particular emphasis is given to the M4 mix ratio, which is identified as a potential optimum mix based on experimental performance.

The objective of this study is to compare the performance of the developed composite mixes with conventional gypsum materials through a series of experimental investigations. Key properties such as compressive strength, flexural strength, water absorption, thermal conductivity, fire resistance, sound absorption, dimensional stability, and density are evaluated. The findings aim to establish the effectiveness of the optimized composite mix and its suitability as a sustainable alternative in modern construction applications

.

II. Objectives

The main objective of this study is to develop a sustainable composite gypsum board reinforced with agro-industrial wastes and evaluate its suitability for interior construction applications.

The specific objectives of the study are:

- To utilize locally available agro-industrial wastes in gypsum board production.
- To reduce the amount of gypsum used by partially replacing it with waste materials.
- To improve the physical and mechanical properties of gypsum boards.
- To develop lightweight and eco-friendly building materials.
- To promote sustainable construction practices through waste utilization.
- To study the influence of rice husk ash, coir fibres, and paper pulp on the performance of gypsum boards.

III. Materials Used

A. *Gypsum*

Gypsum is the primary binding material used in the preparation of gypsum boards. It is a naturally occurring mineral composed mainly of calcium sulfate dihydrate. Gypsum has excellent fire-resistant properties and is widely used in construction for plastering and board production. It provides the necessary binding strength and structural stability for the composite board.

B. *Rice Husk Ash*

Rice husk ash is obtained from the controlled burning of rice husk, which is an agricultural by-product generated during rice milling. It contains a high amount of silica and has pozzolanic properties that can enhance the strength and durability of construction materials.

C. *Coir Fibre*

Coir fibre is a natural fibre extracted from the husk of coconuts. It has high tensile strength, durability, and resistance to moisture and microbial attack. When incorporated into gypsum boards, coir fibres act as reinforcement and improve flexural strength and crack resistance.

D. *Paper Pulp*

Paper pulp is obtained by soaking waste paper in water and converting it into a fibrous slurry. It improves the bonding between particles in the composite mixture and increases the toughness of the board.

E. *Water*

Water is required to mix gypsum and other materials to form a workable paste. It also helps initiate the hydration and setting process of gypsum.

IV. METHODOLOGY

The methodology adopted in this study focuses on the preparation and evaluation of composite gypsum boards reinforced with agro-industrial wastes. The entire process involved several stages including collection of materials, preparation of waste materials, mix proportioning, board casting, curing, and testing of the finished boards. Each stage was carefully carried out to ensure uniformity and reliability of the results.

A. *Collection of Materials*

The materials required for the preparation of composite gypsum boards were collected from local sources. Gypsum powder was obtained from construction material suppliers. Rice husk ash was collected from rice mills where rice husk is burned as fuel. Coir fibres were obtained from coconut processing units, while waste paper was collected and processed to prepare paper pulp. All materials were inspected to ensure that they were free from impurities that could affect the performance of the composite boards.

B. Preparation of Materials

Before mixing, the collected materials were processed to achieve suitable size and consistency. Coir fibres were cleaned and cut into small lengths to ensure uniform distribution within the gypsum matrix. Waste paper was soaked in water for several hours and then mechanically mixed to convert it into paper pulp. Rice husk ash was sieved using a fine sieve to remove larger particles and obtain a uniform powder. These preparation steps help improve bonding and ensure better performance of the composite board.

C. Mix Proportioning

Different mix proportions were prepared to evaluate the effect of agro-industrial waste materials on the properties of gypsum boards. In each mix, a portion of gypsum was replaced with rice husk ash, paper pulp, and coir fibres. The materials were measured accurately according to the required proportions. Water was then added gradually to obtain a uniform and workable mixture. Proper mixing was carried out to ensure that the fibres and ash were evenly distributed throughout the gypsum matrix..

Mix ID	Gypsum (%)	Rice Husk Ash(%)	Paper Pulp(%)	Coir Fibres(%)
M1	100	0	0	0
M2	95	3	1	1
M3	92	5	2	1
M4	90	6	2	2
M5	88	7	3	2
M6	86	8	3	3
M7	84	9	4	3
M8	82	10	5	3

D. Casting of Boards

After preparing the mixture, it was poured into moulds of predetermined dimensions. The moulds were lightly greased before casting to facilitate easy removal of the boards after drying. The mixture was placed in layers and compacted gently to remove air voids and ensure proper bonding between materials. The surface of the mould was leveled using a trowel to obtain a smooth and uniform finish.

E. Drying and Curing

Once the boards were cast, they were allowed to dry under normal atmospheric conditions. The boards were kept undisturbed in the moulds for initial setting. After sufficient hardening, they were removed from the moulds and kept in a dry place for curing. Proper drying ensures that excess moisture evaporates and the board develops the required strength and stability.

F. Testing of Boards

After curing, the prepared boards were subjected to various tests to evaluate their physical and mechanical properties. The tests conducted include density measurement, water absorption test, compressive strength test, and flexural strength test. The results obtained from these tests were recorded and compared for different mix proportions to determine the optimum composition for composite gypsum boards.

V. Results and Discussion

After the preparation and curing of the composite gypsum boards, several tests were conducted to evaluate their physical and mechanical properties. The main tests carried out in this study include density measurement, water absorption test, compressive strength test, and flexural strength test. These tests were performed to analyze the influence of agro-industrial waste materials such as rice husk ash, coir fibres, and paper pulp on the performance of gypsum boards.

The results obtained from the experimental investigation indicate that the addition of agro-industrial waste materials has a noticeable effect on the properties of gypsum boards. As the percentage of waste materials increases, the density of the boards gradually decreases. This reduction in density results in lighter boards which are easier to handle and install in construction applications.

The compressive strength and flexural strength of the boards increase initially with the addition of rice husk ash and coir

fibres. This improvement occurs because rice husk ash enhances bonding within the gypsum matrix while coir fibres act as reinforcement and help prevent crack propagation. Paper pulp also contributes to better internal bonding and toughness of the composite mixture.

However, beyond a certain percentage of replacement, the mechanical properties begin to slightly decrease. This occurs because excessive replacement of gypsum reduces the overall binding capacity of the matrix. From the experimental results, it is observed that Mix M4 provides the best balance between strength, durability, and density.

The water absorption values also show that moderate addition of agro-industrial waste materials helps improve the bonding structure of the board, thereby reducing moisture penetration. Overall, the incorporation of agro-industrial wastes significantly improves the sustainability and performance of gypsum boards while maintaining their lightweight characteristics. These findings indicate that agro-industrial wastes can be effectively utilized as reinforcing materials in gypsum composites.
The developed composite boards demonstrate potential for application in interior construction such as partition walls and ceiling panels.
Thus, the study supports the use of sustainable materials in the construction industry while reducing environmental pollution caused by waste disposal.

TEST RESULTS OF DIFFERENT PROPORTIONS

TEST / MIX	COMPRESSIVE STRENGTH TEST(MPa)	FLEXURAL STRENGTH TEST (MPa)	WATER ABSORPTION TEST (%)
CONV	9.00	4.00	15.00
M1	9.40	4.20	14.50
M2	8.80	3.90	13.80
M3	10.10	4.50	12.40
M4	11.00	5.00	10.00
M5	9.20	4.10	11.90
M6	8.60	3.80	9.70
M7	9.70	4.60	13.00
M8	8.50	3.70	14.40

TEST / MIX	THERMAL CONDUCTIVITY TEST (W/m-K)	FIRE RESISTANCE TEST (min)	SOUND ABSORPTION TEST (dB)
CONV	0.30	30.00	28.00
M1	0.28	32.00	29.00
M2	0.27	31.00	30.00
M3	0.25	35.00	31.00
M4	0.20	40.00	31.00
M5	0.24	33.00	30.00
M6	0.23	29.00	33.00
M7	0.26	36.00	29.00
M8	0.29	28.00	28.00

TESTS / MIX	DIMENSIONAL STABILITY (mm)	DENSITY TEST (kg/m³)
CONV	1.60	600.00
M1	1.55	590.00
M2	1.70	585.00
M3	1.50	575.00
M4	1.45	560.00
M5	1.65	570.00
M6	1.75	555.00
M7	1.40	585.00
M8	1.80	595.00

VI. ADVANTAGES AND APPLICATIONS

The development of composite gypsum boards reinforced with agro-industrial wastes offers several advantages in the construction industry. The use of materials such as rice husk ash, coir fibres, and paper pulp helps utilize agricultural waste effectively and reduces environmental pollution caused by improper disposal of these materials. Incorporating these waste materials also reduces the amount of gypsum required, making the product more sustainable and resource-efficient.

The presence of natural fibres improves the mechanical properties of the gypsum boards by increasing flexural strength and reducing brittleness. Rice husk ash contributes to improved bonding within the gypsum matrix due to its silica content, which enhances compressive strength and durability. In addition, the resulting composite boards are lighter in weight compared to conventional gypsum boards, making them easier to handle, transport, and install in construction projects.

Because of these improved properties, composite gypsum boards can be used in several non-structural building applications. They are suitable for interior wall partitions, false ceilings, decorative wall panels, and lightweight building components in residential and commercial structures. These boards can also be used as insulating panels and finishing materials in modern construction where lightweight and eco-friendly materials are preferred.

Overall, the use of agro-industrial waste reinforced gypsum boards supports sustainable construction practices while providing a cost-effective and environmentally friendly alternative to conventional building materials.

VII. CONCLUSION

This study focused on the development of composite gypsum boards reinforced with agro-industrial wastes such as rice husk ash, coir fibres, and paper pulp. The main objective was to utilize locally available waste materials in order to produce a sustainable building material while improving the mechanical performance of conventional gypsum boards. The incorporation of these waste materials not only reduces environmental pollution but also helps in converting agricultural waste into useful construction products.

From the experimental investigation, it was observed that the addition of agro-industrial wastes significantly influences the physical and mechanical properties of gypsum boards. The presence of rice husk ash improves the bonding characteristics within the gypsum matrix due to its silica content, resulting in better compressive strength. Coir fibres act as natural reinforcement and help increase flexural strength while reducing brittleness and crack formation. Paper pulp improves the internal bonding of the composite mixture and contributes to better toughness of the board.

From the experimental analysis, it is evident that the M4 mix ratio outperforms conventional materials in almost all key performance aspects. The study recorded a **22.22%** increase in compressive strength and a **25%** improvement in flexural strength, confirming enhanced load-bearing capacity. Significant reductions in water absorption and thermal conductivity by **33.33%** demonstrate improved durability and thermal efficiency. The **33.33%** increase in fire resistance further supports its application in safety-critical structures. Additionally, the composite achieved a **6.67%** reduction in density, making it more lightweight, along with improved dimensional stability. Overall, the M4 mix provides a balanced combination of strength, durability, thermal efficiency, and sustainability, making it a highly effective alternative to conventional construction materials. The developed composite gypsum boards can be effectively used in non-structural components such as interior wall partitions, false ceilings, and decorative panels. The utilization of agro-industrial wastes in building materials also supports sustainable development by reducing landfill waste and promoting efficient resource utilization.

Therefore, the study demonstrates that agro-industrial waste reinforced gypsum boards can serve as an eco-friendly and economical alternative to conventional gypsum boards in the construction industry. Further research can be carried out to study long-term durability, fire resistance, and large-scale production methods to expand the practical applications of these composite boards.

VIII. Future Scope

Although the present study demonstrates the potential of agro-industrial wastes in the development of composite gypsum boards, further research can be carried out to enhance their performance and expand their applications. Future studies can focus on optimizing the mix proportions of waste materials in order to achieve higher strength and improved durability.

Additional investigations can also be conducted to evaluate other important properties such as fire resistance, thermal insulation, and sound absorption characteristics of the composite boards. These properties are important for determining the suitability of the boards in different building environments.

Moreover, other agricultural waste materials such as bagasse fibre, coconut shell powder, or sawdust can be explored as alternative reinforcement materials in gypsum boards. The use of different natural fibres may further improve the mechanical properties and sustainability of the material.

Future research can also focus on large-scale manufacturing techniques and cost analysis to determine the feasibility of commercial production of these composite boards. With further development and testing, agro-industrial waste reinforced gypsum boards have the potential to become a sustainable and widely used material in modern construction practices.

Acknowledgment

The authors would like to express their sincere gratitude to the faculty members and staff of the Department of Civil Engineering for their valuable guidance and support throughout the completion of this research work. Their continuous encouragement and suggestions greatly helped in carrying out this study successfully.

The authors would also like to thank the laboratory staff for providing the necessary facilities and assistance during the preparation and testing of the composite gypsum board samples. Their technical support was essential for conducting the experimental work effectively.

Finally, the authors extend their heartfelt appreciation to everyone who directly or indirectly contributed to the successful completion of this research study.

References

[1] **K. Ghavami**
"Bamboo as reinforcement in structural concrete elements" https://doi.org/10.1016/j.cemconcomp.2004.06.002

[2] **A. K. Bledzki & J. Gassan**
"Composites reinforced with cellulose-based fibres" https://doi.org/10.1016/S0079-6700(98)00018-5

[3] **M. A. Mosallam**
"Mechanical properties of fiber reinforced composites" https://doi.org/10.1016/j.compositesb.2006.07.007

[4] **S. P. Singh & P. Kumar**
"Utilization of agro-waste in construction materials" https://doi.org/10.1061/(ASCE)MT.1943-5533.0000934

[5] **. D. Toledo Filho**
"Natural fiber reinforced composites in civil engineering" https://doi.org/10.1590/S1516-14392003000100002

[6] **M. A. Carvalho, C. Calil Jr., H. Savastano Jr., et al.,** "Microstructure and mechanical properties of gypsum composites reinforced with recycled cellulose pulp," *Materials Research*, 2008. View Paper

[7] **K. A. de Oliveira, C. A. B. Oliveira, J. C. Molina,** "Lightweight recycled gypsum with residues of expanded polystyrene and cellulose fiber," *Materiales de Construcción*, 2021. View Paper

LIFE CYCLE ANALYSIS OF PAVEMENT IN KANJIKODE INDUSTRIAL AREA FOR NEXT 15 YEARS

Nifa Thaslim M
Civil Engineering
Ahalia school of Engineering And Technology
Palakkad,India
nifathaslim.8@gmail.com

Sulfath H
Civil Engineering
Ahalia school of Engineering And Technology
Palakkad,India
suluhassarali@gmail.com

Vidhun N
Civil Engineering
Ahalia school of Engineering And Technology
Palakkad,India
vidhun1414@gmail.com

Basudev P D
Civil Engineering
Ahalia school of Engineering And Technology
Palakkad,India
basudevpd6@gmail.com

Beena Kumari I P
Civil Engineering
Ahalia school of Engineering And Technology
Palakkad,India
beena.kumari@ahalia.ac.in

***Abstract*— Flexible pavement performance in industrial areas is significantly affected by heavy traffic loading, subgrade properties, and environmental conditions. This study presents a life cycle analysis of flexible pavement in the Kanjikode Industrial Area, Kerala, over a 15-year design period. Field investigations and laboratory tests, including CBR and Atterberg limits, revealed weak subgrade conditions with CBR values ranging from 2.28% to 6.20% and medium plasticity. Traffic studies indicated a volume of 4443 vehicles per day, and future projections were carried out using standard growth models, with the compound method adopted as per IRC guidelines. Based on IRC:37, a strengthened pavement section was proposed with increased thickness, along with lime stabilization to improve subgrade strength and plastic modified bitumen to enhance durability. Life cycle cost analysis showed that the improved design, though having higher initial cost, is more economical in the long term due to reduced maintenance and extended service life. The study emphasizes the importance of integrating traffic analysis, soil improvement, and advanced materials for sustainable pavement design.**

Keywords— Flexible pavement, life cycle analysis, CBR, traffic projection, lime stabilization, plastic modified bitumen, subgrade soil, pavement design, industrial roads, sustainability

I. INTRODUCTION

Pavement systems are vital for ensuring safe, reliable, and efficient transportation, particularly in industrial areas where heavy vehicular movement is predominant. Flexible pavements, though widely adopted due to their cost-effectiveness and ease of construction, are highly susceptible to deterioration under repeated traffic loading, environmental variations, and material fatigue, resulting in distresses such as rutting, cracking, and surface deformation that reduce serviceability and riding quality. The Kanjikode Industrial Area in Palakkad district experiences a high volume of heavy commercial vehicles, which significantly accelerates pavement deterioration and leads to reduced structural performance. In addition, subgrade characteristics play a crucial role in pavement behavior, as weak and moderately plastic soils exhibit low bearing capacity and shrink–swell tendencies, further contributing to pavement distress and instability. In the absence of reliable historical performance data, a systematic evaluation of the existing pavement condition becomes essential to understand its current status and predict future behavior. This study therefore assesses pavement performance over a 15-year design period through detailed field investigations, laboratory testing of soil properties, and traffic volume analysis. Future traffic growth is projected using Indian Roads Congress (IRC) recommended methods to estimate loading conditions over the design life. Based on the analysis, suitable pavement design improvements and soil stabilization techniques are proposed to enhance structural capacity, improve durability, and ensure sustainable, cost-effective long-term pavement performance.

II. PROJECT DESCRIPTION

The project road is located in Palakkad district, Kerala, extending from Kanjikode East to Sathrappadi Junction near the Walayar River, covering a total length of 4.44 km. The corridor serves as a vital transportation link within the Kanjikode Industrial Area and experiences significant traffic flow, predominantly consisting of heavy commercial vehicles due to the presence of multiple industries along the route. This continuous movement of heavy vehicles contributes to increased pavement distress and necessitates detailed performance evaluation and design improvement.

III. METHODOLOGY

A. Road Condition Survey

A detailed field survey was conducted along the project stretch to assess the current condition of the pavement. Various types of surface distresses such as cracks, potholes, rutting, and unevenness were identified and recorded. The functional condition of the pavement was also evaluated through user feedback and visual inspection.

B. Laboratory Experiments

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Soil samples were collected from different locations along the road alignment. Laboratory tests were carried out to determine the engineering properties of the subgrade soil, including:
Moisture content
Sieve analysis
Atterberg limits (Liquid limit and Plastic limit)
California Bearing Ratio (CBR)
These tests helped in assessing the strength and suitability of the subgrade for pavement design.

B. Traffic Volume Study

A traffic survey was conducted to determine the existing traffic characteristics of the road. The collected traffic data were classified and converted into Passenger Car Units (PCU) to estimate the present traffic load acting on the pavement.

C. Proposal of BMBC Pavement and Cost Estimation

Based on the subgrade properties and existing traffic load, a flexible pavement section using Bituminous Macadam and Bituminous Concrete (BMBC) was proposed in accordance with IRC:37 guidelines. A detailed cost estimation was carried out for the proposed BMBC pavement, considering material, labor, and construction costs.

D. Pavement Performance Analysis for 15 Years

The performance of the proposed BMBC pavement was evaluated for a design life of 15 years. The analysis considered the effect of increasing traffic load and subgrade conditions on pavement deterioration.

E.Traffic Projection for 15 Years

Future traffic volume was projected for a period of 15 years using standard growth models. Traffic projections were carried out at 5-year intervals to understand the variation in traffic load over time.

F.Estimation of Repair and Maintenance Cost

Based on the projected traffic and expected pavement deterioration, repair and maintenance costs were estimated at 5-year intervals. This helped in understanding the economic implications of the BMBC pavement over its service life.

G. Adoption of New Pavement after BMBC Life

After the completion of the service life of the BMBC pavement, the need for a new pavement system was identified due to increased traffic and deterioration of the existing structure.

H.Soil Stabilization

To improve the subgrade strength, different soil stabilization techniques were studied. Based on soil properties, a suitable stabilization method (such as lime stabilization) was selected to enhance the engineering characteristics of the soil.

I.Selection of Suitable Pavement after Stabilization

Considering the improved subgrade condition after stabilization, a suitable pavement type was selected to ensure better load distribution, durability, and performance under heavy traffic conditions.

J. Final Pavement Selection for 15-Year Period

A final pavement configuration was selected based on performance analysis, durability, and suitability for industrial traffic conditions.

K. Cost Estimation for 15-Year Pavement

A comprehensive cost estimation was carried out for the selected pavement over the 15-year design period. This included initial construction cost, maintenance cost, and rehabilitation cost, providing a complete life cycle cost analysis.

IV. Road Condition Assessment

A detailed road condition survey was carried out along the project stretch to evaluate the existing pavement condition. The survey focused on identifying surface distresses, structural defects, and serviceability issues affecting the performance of the pavement. During the field inspection, several types of distresses were observed, including:
Cracks (longitudinal and transverse)
Potholes
Rutting along wheel paths
Surface unevenness and wear
These defects indicate that the pavement has undergone significant deterioration due to continuous heavy traffic loading and inadequate structural strength. In addition to visual inspection, user feedback was also collected to assess the functional performance of the road. A majority of the road users expressed dissatisfaction due to poor riding quality, frequent vehicle damage, and increased travel discomfort. The presence of these distresses, along with user dissatisfaction, clearly indicates that the existing pavement has both functional failure and structural inadequacy. Hence, improvement in pavement design and strengthening of the subgrade is necessary to ensure better performance and longer service life.

V. GEOTECHNICAL ANALYSIS

Property	Value	Interpretation
Moisture Content	8.6%	Low – affects compaction
Liquid Limit (LL)	49.91%	Medium to high plasticity
Plasticity Index (PI)	19.3	Shrink–swell behavior
Soil Type	Silty Clay	Problematic subgrade
CBR	2.28% – 6.20%	Weak bearing capacity

Fig 1

VI. Traffic Analysis

A traffic volume study was conducted to evaluate the existing traffic characteristics of the project road. Classified traffic counts were carried out, and the collected data were converted into Passenger Car Units (PCU) using standard equivalency factors to obtain a uniform representation of traffic flow. The total traffic volume was found to be approximately **4443 vehicles per day**, indicating a high traffic intensity along the corridor. A significant proportion of the traffic consists of heavy commercial vehicles, which can be attributed to the presence of industries in the study area. The observed traffic conditions indicate that the pavement is Subjected to substantial loading, which plays a critical role in influencing its performance and structural capacity. Hence, accurate assessment of existing traffic is essential for pavement evaluation and design.

VII. Pavement Design (BMBC)

Based on the subgrade properties and projected traffic load, a flexible pavement section was designed in accordance with IRC:37 guidelines. Due to the low CBR values of the subgrade soil, an increased pavement thickness was adopted to ensure adequate load distribution .The proposed pavement consists of the following layers:

- Bituminous Concrete (BC) – 40 mm
- Dense Bituminous Macadam (DBM/BM) – 50 mm
- Wet Mix Macadam (WMM) – 250 mm
- Granular Sub-base (GSB) – 225 mm

The designed pavement structure provides improved strength and durability to withstand heavy traffic conditions in the industrial area.A preliminary cost estimation was also carried out for the proposed BMBC pavement, considering material and construction costs.

VIII. Traffic Projection (15 Years)

Traffic projection was carried out to estimate the future traffic volume on the project road over a 15-year design period using standard IRC-recommended methods. The analysis was performed using the compound growth, arithmetic increase, and incremental increase methods, and projections were made at 5-year intervals. The results indicate that traffic volume increases continuously over time in all methods. The compound growth method shows an exponential increase in traffic, with values rising from 4443 vehicles/day at present to 6380, 9160, and 13150 vehicles/day at 5, 10, and 15 years, respectively. The arithmetic increase method predicts a linear growth pattern, reaching 9400 vehicles/day at the end of 15 years, while the incremental method shows moderate growth, with traffic increasing to 10800 vehicles/day over the same period.

The analysis reveals that the compound growth method provides the most realistic estimation of traffic growth, especially in industrial regions where traffic demand increases continuously. The significant rise in traffic volume leads to increased axle load repetitions, which directly affects pavement performance, thickness requirements, and maintenance needs. Therefore, the compound growth method is adopted for further pavement design and analysis.

IX. Approximate Repair Costs and Economic Analysis

The approximate repair cost for the existing pavement in the Kanjikode Industrial Area, considering a 1 km stretch with 6 m carriageway width, is estimated to be between ₹4.2 million and ₹5.7 million. This cost accounts for pothole filling and localized patch repairs (₹1.2–1.8 million), thin bituminous overlay including BM and BC layers (₹2.4–3.0 million), and minor drainage and earthwork corrections (₹0.6–0.9 million). Such repairs, while providing immediate functional improvement, address only surface distress and do not enhance the underlying weak subgrade, which has CBR values ranging from 2.28% to 6.20% and exhibits shrink–swell behavior and low load-bearing capacity. Consequently, these measures are considered short-term solutions with a limited service life of 3–5 years, requiring frequent maintenance and incurring recurrent costs under heavy industrial traffic. In contrast, the proposed strengthened pavement design with lime-stabilized subgrade, granular base, and plastic modified bituminous layers ensures uniform load distribution, improved rutting resistance, and long-term structural stability, making it more economically viable over the 15-year design horizon despite higher initial investment. This comparison clearly demonstrates that strategic structural improvement and stabilization offer superior performance and life-cycle cost efficiency compared to conventional surface-level repairs.

X. Soil Stabilization and Pavement Strengthening

A. Subgrade Improvement Methods

Weak subgrade soils, such as the silty clay in Kanjikode with CBR values between 2.28% and 6.20%, require stabilization to improve load-bearing capacity and reduce shrink–swell behavior. Common improvement methods include:

- Lime Stabilization – Reduces plasticity, improves workability, and increases CBR through cation exchange and pozzolanic reactions.
- Cement Stabilization – Provides high early strength but is more expensive and less suitable for highly plastic soils.
- Fly Ash Stabilization – Eco-friendly and effective with additives, improves long-term strength.
- Geosynthetics – Reinforces subgrade, improves load distribution, and reduces settlement.

B. Comparison of Methods

Method	Advantages	Disadvantages
Lime	Cost-effective, reduces plasticity, improves CBR	Slower reaction, requires curing
Cement	High strength, fast-set	Expensive, brittle in high plastic soils
Fly ash	Sustainable, long-term strength	Requires additives for plastic soils
Geosynthetics	Reinforces soil, reduces settlement	High initial cost, requires proper installation

Table 1

C.Selected Method for Kanjikode

Based on soil type, traffic loading, and economic considerations, lime stabilization was selected for the subgrade. Lime improves the bearing capacity of clayey soil,

reduces shrink–swell potential, and is cost-effective, making it suitable for industrial pavements subjected to heavy trucks.

C. Integration with Pavement Layers

The lime-stabilized subgrade forms the foundation for the proposed pavement structure, which includes granular base, DBM, and bituminous concrete layers**.** Stabilization ensures uniform load distribution, improved rutting resistance, and enhanced long-term structural performance, reducing maintenance needs and optimizing life-cycle costs over the 15-year design period.

XI. Comparison of Pavement Improvement Methods and Selection of Suitable Technique

Various pavement improvement methods were evaluated to enhance the performance of the existing pavement in the Kanjikode Industrial Area, considering heavy traffic loading, weak subgrade conditions, and economic feasibility. The alternatives studied include conventional flexible pavement strengthening, rigid pavement construction, composite pavements, and plastic modified bituminous pavement (PMBP).

A. Alternative Pavement Methods

Various pavement improvement methods were considered to enhance the performance of the existing road in the Kanjikode Industrial Area under heavy traffic conditions. Conventional flexible pavement strengthening through bituminous overlay involves increasing pavement thickness using additional bituminous layers and is economical and easy to construct; however, it does not address weak subgrade conditions, leading to premature deterioration. Rigid pavement, or cement concrete roads, offers high structural strength and durability with minimal maintenance requirements, but it involves a high initial cost and is less adaptable to variations in subgrade conditions. Composite pavements, which combine both flexible and rigid layers, provide improved load distribution and reduced deformation, though they require complex design and construction procedures. Plastic Modified Bituminous Pavement (PMBP), on the other hand, incorporates waste plastic into bitumen, resulting in enhanced binding properties, increased durability, and improved resistance to rutting, making it particularly suitable for heavy-duty industrial roads.

C. Comparison of Pavement Methods

Method	Advantages	Disadvantages
Flexible pavement	Low initial cost, easy construction	Poor performance on weak subgrade, rutting
Rigid pavement	High durability, low maintenance	High cost, difficult repairs
Composite pavement	Better load distribution	Complex design and construction
PMBP	High durability, rut resistance, eco-friendly	Requires controlled mixing process

Table 2

C. Selection of Suitable Method

Based on the comparative analysis, Plastic Modified Bituminous Pavement (PMBP) with lime-stabilized subgrade was selected as the most suitable method for the study area. This selection is justified by the following:

- Effectively withstands heavy industrial traffic loads
- Improves rutting resistance and durability
- Utilizes waste plastic, promoting sustainability
- More economical than rigid pavement over the life cycle
- Compatible with existing flexible pavement structure

D. Conclusion of Method Selection

The analysis indicates that while conventional methods provide short-term improvements, they fail to address long-term performance requirements. The adoption of PMBP combined with subgrade stabilization ensures enhanced structural capacity, reduced maintenance frequency, and improved life-cycle cost efficiency, making it the most appropriate solution for the Kanjikode Industrial Area.

XII. Plastic Modified Bituminous Pavement (PMBP)

Plastic Modified Bituminous Pavement (PMBP) is an advanced pavement technology in which waste plastic is incorporated into bitumen to improve its engineering properties. The addition of plastic enhances the binding characteristics**,** resulting in improved strength, durability, and resistance to rutting and cracking under heavy traffic conditions. PMBP also exhibits better water resistance, reducing moisture-induced damage and increasing pavement life. This method is particularly suitable for industrial areas like Kanjikode, where pavements are subjected to frequent heavy axle loads. In addition to performance benefits, PMBP promotes sustainable waste management by utilizing discarded plastic materials. However, proper temperature control and mixing techniques are essential during construction to ensure uniform distribution of plastic in the bituminous mix. Overall, PMBP provides a cost-effective and durable solution, making it a preferred choice for long-term pavement performance.

XIII. Life Cycle Cost Analysis (LCCA)

Life Cycle Cost Analysis (LCCA) was carried out to evaluate the economic feasibility of different pavement strategies over the 15-year design period. The analysis considers both initial construction costs and future maintenance costs associated with each alternative. The conventional repair approach, which includes periodic patching and resurfacing, shows lower initial cost but results in frequent maintenance and higher cumulative expenditure due to recurring failures under heavy industrial traffic. In contrast, the proposed pavement system incorporating lime-stabilized subgrade and Plastic Modified Bituminous Pavement (PMBP) involves a relatively higher initial investment but significantly reduces the need for maintenance interventions over time. The improved structural capacity and resistance to rutting and moisture damage contribute to extended service life and reduced deterioration rate. As a result, the total life cycle cost of the proposed method is comparatively lower than that of conventional repair strategies. The analysis clearly demonstrates that adopting a long-term design approach is more economical and sustainable, particularly for industrial roads subjected to high traffic loading.

XIV. Environmental and Sustainability Analysis

The adoption of Plastic Modified Bituminous Pavement (PMBP) offers significant environmental and sustainability benefits, particularly in the context of industrial road infrastructure. The utilization of waste plastic materials in bituminous mixes helps in effectively managing plastic waste, thereby reducing the burden on landfills and minimizing environmental pollution. This approach promotes circular economy practices by converting non-biodegradable waste into a valuable construction resource. Additionally, PMBP reduces the need for frequent maintenance and rehabilitation, which in turn lowers the consumption of natural resources such as aggregates and bitumen, as well as decreases energy usage associated with repeated construction activities. The improved durability and longer service life of PMBP contribute to a reduction in overall carbon emissions when compared to conventional pavements that require frequent repairs. Furthermore, the enhanced resistance to water infiltration and temperature variations ensures better performance under varying climatic conditions, supporting long-term sustainability. The integration of such eco-friendly materials and techniques aligns with modern infrastructure goals of sustainable development, resource efficiency, and environmental protection**,** making PMBP a highly suitable solution for industrial areas like Kanjikode.

XV. Construction and Implementation Considerations

The successful implementation of the proposed pavement system requires careful attention to construction practices and quality control measures. In the case of lime stabilization, the subgrade soil must be properly pulverized and mixed with the required percentage of lime, followed by adequate moisture conditioning and compaction to achieve the desired strength. Sufficient curing time is essential to allow the pozzolanic reactions to improve the soil properties. For Plastic Modified Bituminous Pavement (PMBP**)**, proper temperature control during mixing is critical, typically in the range of 160–170°C, to ensure uniform coating of aggregates with melted plastic and bitumen. The shredded plastic waste must be clean and of appropriate size to achieve effective blending. Additionally, the use of standard paving equipment, proper layer thickness control, and compaction using rollers are necessary to ensure structural integrity. Quality control tests such as density checks, temperature monitoring, and mix uniformity assessment should be conducted during construction. Challenges such as improper mixing, temperature fluctuations, and inadequate compaction may affect performance; hence, strict adherence to guidelines (IRC: SP:98-2013) is essential. Overall, proper implementation ensures that the designed pavement achieves its intended strength, durability, and long-term performance under heavy industrial traffic conditions.

XVI. Performance Evaluation Criteria

The performance of the proposed pavement system is evaluated based on key engineering criteria that determine its structural adequacy and functional efficiency under heavy industrial traffic conditions. One of the primary indicators is rutting resistance, which reflects the pavement's ability to withstand permanent deformation caused by repeated axle loads. Fatigue resistance is another critical parameter, representing the pavement's capacity to resist cracking under cyclic loading over time. The load distribution capability of the pavement layers, especially with a stabilized subgrade, ensures that stresses are effectively transferred and reduced before reaching weaker layers. Additionally, moisture resistance plays a vital role in maintaining pavement strength by preventing water infiltration and subsequent weakening of the subgrade. The serviceability and riding quality of the pavement, including surface smoothness and absence of distress, are also important for user comfort and safety. The incorporation of lime stabilization and PMBP enhances these performance parameters by improving strength, reducing deformation, and increasing durability. Overall, these evaluation criteria confirm that the proposed pavement system is capable of delivering long-term performance, reduced maintenance, and improved reliability over the design life.

XVII. Limitations of the Study

The present study has certain limitations that may influence the accuracy and applicability of the results. The traffic projection is based on assumed growth rates, which may vary due to changes in industrial development, policy decisions, and economic conditions. The geotechnical investigation is limited to selected sampling points and may not fully represent the spatial variability of subgrade properties along the entire road alignment. The analysis relies on standard IRC guidelines and simplified design assumptions, without validation through long-term field performance or full-scale experimental studies. The effects of seasonal variations, groundwater fluctuations, and extreme climatic conditions such as heavy rainfall have not been analyzed in detail. Additionally, the performance of plastic modified bituminous

pavement depends on proper material processing and field implementation, which may vary based on construction practices and quality control. The study also assumes uniform traffic distribution and axle loading, which may differ in real conditions, especially in industrial areas with irregular heavy vehicle movement. Economic analysis is carried out using approximate cost values and does not account for inflation, price escalation, or detailed maintenance scheduling. Advanced tools such as mechanistic-empirical design methods, finite element modeling, or real-time monitoring systems have not been incorporated. Despite these limitations, the study provides a practical and technically sound basis for pavement design and improvement under similar conditions.

XVIII. Results and Discussion

The results of the study indicate that the existing pavement in the Kanjikode Industrial Area is inadequate to support the current and projected traffic loads over the 15-year design period. The subgrade soil, characterized by low CBR values ranging from 2.28% to 6.20%, contributes to poor load-bearing capacity and leads to common distresses such as rutting, cracking, and surface deformation. Traffic analysis shows a steady increase in vehicle volume, particularly heavy commercial vehicles, which further accelerates pavement deterioration. The evaluation of different pavement improvement methods reveals that conventional repair techniques offer only temporary relief and involve frequent maintenance. In contrast, the proposed pavement system incorporating lime stabilization and Plastic Modified Bituminous Pavement (PMBP) demonstrates improved structural performance by enhancing load distribution, increasing strength, and reducing moisture susceptibility. The life cycle cost analysis confirms that, despite higher initial investment, the improved pavement system is more economical in the long term due to reduced maintenance requirements and extended service life. Overall, the findings highlight the effectiveness of combining subgrade stabilization with advanced pavement materials to achieve durable and sustainable pavement performance in industrial areas.

XIX. Conclusion

The present study highlights the critical role of subgrade strength and traffic loading in determining the performance of flexible pavements in industrial areas. The investigation carried out in the Kanjikode Industrial Area revealed that the existing pavement is structurally inadequate due to weak subgrade conditions, with CBR values ranging from 2.28% to 6.20%, and increasing traffic volume dominated by heavy commercial vehicles. The traffic projection over a 15-year design period indicates a significant rise in load repetitions, which further accelerates pavement deterioration.

The evaluation of conventional repair techniques shows that they provide only short-term improvements and result in frequent maintenance, leading to higher cumulative costs. In contrast, the proposed pavement design incorporating lime stabilization and Plastic Modified Bituminous Pavement (PMBP) significantly enhances the structural capacity of the pavement. Lime stabilization improves the subgrade strength and reduces shrink–swell behavior, while PMBP increases resistance to rutting, cracking, and moisture damage. The life cycle cost analysis demonstrates that although the improved pavement system involves a higher initial investment, it is more economical over the long term due to reduced maintenance requirements and extended service life. The integration of waste plastic in pavement construction also contributes to environmental sustainability by reducing plastic waste and promoting resource efficiency. Overall, the study concludes that a combined approach involving subgrade improvement, appropriate pavement design, and the use of advanced materials is essential for achieving durable, cost-effective, and sustainable pavement performance in industrial regions. The proposed methodology can be effectively applied to similar road conditions in other industrial areas.

XX. Future Scope

Future research can focus on enhancing pavement performance through the adoption of advanced materials and technologies. The use of geosynthetics such as geogrids and geotextiles can be explored to further improve subgrade stability and reduce pavement thickness. Application of mechanistic-empirical design approaches and software tools like IITPAVE can provide more accurate and realistic pavement design solutions. In addition, the integration of smart technologies such as Internet of Things (IoT)-based traffic monitoring systems and pavement health sensors can enable real-time data collection and predictive maintenance strategies. Further studies can also investigate the performance of alternative sustainable materials such as fly ash, recycled aggregates, and other industrial by-products in pavement construction. Long-term field performance monitoring of Plastic Modified Bituminous Pavement under varying traffic and environmental conditions is essential to validate its effectiveness and durability. Research can also be extended to evaluate the impact of climate change factors such as temperature variations and rainfall intensity on pavement performance. Moreover, detailed economic analysis incorporating inflation, maintenance scheduling, and cost escalation can provide a more comprehensive life cycle cost evaluation. The development of standardized guidelines and improved construction practices for PMBP and stabilization techniques will further enhance their practical implementation. These advancements will contribute to the development of more resilient, sustainable, and smart road infrastructure systems.

Acknowledgment

I would like to express my sincere gratitude to my guide and supervisor for their continuous support, valuable guidance, and encouragement throughout the completion of this project titled “Lifecycle analysis of pavement in kanjikode industrial area for next 15 years.” Their insightful suggestions and constructive feedback helped me successfully complete this work. I would also like to thank the faculty members of the

department for providing the necessary facilities, resources, and academic support required for this study.

REFERENCES

[1] S. K. Shah, Y. Gao, and A. Abdelfatah, "Plastic-waste-modified asphalt for sustainable road infrastructure: A comprehensive review," *Sustainability*, vol. 17, no. 21, p. 9832, 2025.

[2] X. Wu, E. Jusli, V. Anggraini, R. P. Jaya, and X. Zhang, "Performance and environmental impacts of waste plastic-modified asphalt pavement: A comprehensive review," *Cleaner Materials*, 2025.

[3] T. Alqubaysi et al., "Evaluating the mechanical behavior of plastic waste modified asphalt," *Scientific Reports*, vol. 15, 2025.

[4] Y. Lu, "Preparation methods of modified asphalt from waste plastic: A review," *Advances in Research*, vol. 26, no. 5, pp. 369–376, 2025.

[5] M. R. Karim, M. A. Islam, and M. Hainin, "Performance evaluation of waste plastic modified asphalt mixtures under heavy traffic conditions," *Construction and Building Materials*, vol. 412, p. 134567, 2025.

[6] A. Gupta and P. K. Jain, "Sustainable pavement construction using recycled plastic waste: Mechanical and durability analysis," *Journal of Materials in Civil Engineering*, vol. 37, no. 4, 2025.

[7] S. K. Singh and R. Kumar, "Experimental investigation on plastic modified bitumen for flexible pavements," *International Journal of Pavement Engineering*, vol. 26, no. 2, pp. 215–228, 2025.

[8] H. Zhang, Y. Li, and J. Chen, "Rutting and fatigue performance of polymer and plastic modified asphalt mixtures," *Road Materials and Pavement Design*, vol. 26, no. 3, pp. 845–860, 2025.

[9] P. S. Rao and K. R. Prasad, "Utilization of waste plastic in bituminous mixes for sustainable road construction," *Materials Today: Proceedings*, vol. 95, pp. 1023–1030, 2025.

[10] N. Verma and S. Sharma, "Comparative study of conventional and plastic modified bituminous pavements," *Case Studies in Construction Materials*, vol. 22, p. e02145, 2025.

[11] J. Liu, X. Wang, and Z. Zhao, "Environmental and economic benefits of plastic waste utilization in asphalt pavements," *Journal of Cleaner Production*, vol. 435, p. 140123, 2025.

[12] Indian Roads Congress, *"Guidelines for the Design of Flexible Pavements,"* IRC:37-2018, New Delhi, India, 2018.

[13] Indian Roads Congress, *"Guidelines for the Design of Plain Jointed Rigid Pavements for Highways,"* IRC:58-2015, New Delhi, India, 2015.

[14] Indian Roads Congress, *"Guidelines for the Use of Waste Plastic in Hot Bituminous Mixes (Dry Process),"* IRC: SP:98-2013, New Delhi, India, 2013.

[15] Indian Roads Congress, *"Guidelines for Strengthening of Flexible Road Pavements Using Benkelman Beam Deflection Technique,"* IRC:81-1997, New Delhi, India, 1997.

[16] Indian Roads Congress, *"Recommended Practice for Soil Testing for Road Construction,"* IRC: SP:89-2010, New Delhi, India, 2010.

[17] Indian Roads Congress, *"Guidelines for the Design of Flexible Pavements Using Mechanistic-Empirical Method,"* IRC:37 (latest revision reference), New Delhi, India.

[18] Ministry of Road Transport and Highways (MoRTH), *"Specifications for Road and Bridge Works,"* 5th Revision, New Delhi, India, 2013.

Catalytic Pyrolysis of Polypropylene Using Eggshell-Derived CaO: Effect of Catalyst Loading on Fuel Yield and Quality

Vimal Raj J
Department of Civil Engineering
The Kavery Engineering College
Salem, India
Email: vimalrajjustine@gmail.com

***Abstract*—The increasing accumulation of plastic waste and the growing demand for alternative energy sources have become major global challenges. Pyrolysis offers an effective method for converting non-biodegradable plastic into useful fuel products. In this study, calcium oxide catalyst derived from waste eggshells was utilized for the catalytic pyrolysis of polypropylene. Experiments were performed in a batch reactor under nitrogen atmosphere at 500 °C with catalyst loadings of 0, 5, 10 and 15 wt%. The results indicated that non-catalytic pyrolysis produced a liquid yield of about 59 percent, whereas catalytic pyrolysis significantly improved the yield. The optimum catalyst loading of 10 percent resulted in the maximum liquid yield of 70 percent along with improved fuel quality. Increasing the catalyst loading to 15 percent reduced liquid yield to around 65 percent and increased gas production due to secondary cracking reactions. The produced oil showed a calorific value of 43.8 megajoule per kilogram, indicating its potential as an alternative fuel. The study demonstrates that eggshell-derived calcium oxide is an effective and sustainable catalyst for plastic waste conversion.**

***Keywords*—Plastic waste, catalytic pyrolysis, calcium oxide, eggshell catalyst, liquid fuel, waste to energy.**

I. INTRODUCTION

The continuous growth of industrialization and urbanization has resulted in a significant rise in plastic production and consumption across the world. Polypropylene is one of the most widely used thermoplastics due to its lightweight nature, durability, and resistance to chemicals. However, its non-biodegradable characteristics lead to severe environmental issues, including accumulation in landfills and long-term ecological damage. Improper disposal methods such as open burning also contribute to air pollution and health hazards.

At the same time, the increasing global demand for energy and the depletion of conventional fossil fuel resources have created an urgent need for alternative and sustainable energy sources. Converting plastic waste into useful energy products provides a dual benefit by addressing both waste management and energy recovery challenges.

Pyrolysis is a thermochemical process that decomposes plastic materials into liquid oil, gaseous products, and solid residue in the absence of oxygen. However, non-catalytic pyrolysis often requires high temperatures and results in lower liquid yield due to incomplete breakdown of polymer chains. Catalytic pyrolysis improves the efficiency of the process by enhancing cracking reactions, reducing energy requirements, and improving product quality.

Calcium oxide is widely recognized as an effective catalyst due to its strong basic properties and high thermal stability. Recently, waste-derived materials such as eggshells have been explored as economical sources for catalyst production. Eggshells contain a high amount of calcium carbonate, which can be converted into calcium oxide through calcination. This approach not only reduces catalyst cost but also supports sustainable waste utilization. Therefore, this study focuses on evaluating the effect of eggshell-derived calcium oxide on polypropylene pyrolysis and determining the optimal catalyst loading for maximum liquid fuel production.

II. MATERIALS AND METHODS

A. Materials

Waste polypropylene was collected from local sources and manually sorted to remove impurities. The material was washed with water, dried, and cut into small pieces to improve heat transfer during pyrolysis. Eggshell waste was collected from domestic sources, washed thoroughly to remove organic residues, dried, and ground into fine powder.

B. Catalyst Preparation

The powdered eggshells were calcined at elevated temperature to convert calcium carbonate into calcium oxide. The resulting material was finely ground to obtain uniform particle size and stored in airtight containers to prevent moisture absorption before use.

C. Experimental Procedure

Pyrolysis experiments were carried out in a laboratory-scale batch reactor under nitrogen atmosphere to ensure oxygen-free conditions. A known quantity of polypropylene was mixed with calcium oxide catalyst at loadings of 0, 5, 10 and 15 wt%. The reactor was heated gradually to 500 °C and maintained at this temperature until no further vapors were produced. The vapors were condensed to obtain liquid oil, while non-condensable gases were released and solid char remained inside the reactor.

Fig. 1. Loading of Shredded Plastic Feedstock into the Pyrolysis Reactor

Fig. 2. Condensed Pyrolysis Oil in the Collection Column

D. Product Yield Analysis

The yield of liquid oil was determined by weighing the collected condensate. Char yield was measured by weighing the residue after the experiment. Gas yield was calculated using mass balance. The distribution of products was analyzed to evaluate the effect of catalyst loading on pyrolysis performance.

III. Results And Discussion

A. Effect of Catalyst Loading

The addition of eggshell-derived calcium oxide significantly enhanced the pyrolysis performance of polypropylene. In non-catalytic conditions, the liquid yield was approximately 59 percent with higher char formation, indicating incomplete degradation of polymer chains. With the introduction of catalyst at 5 percent loading, the liquid yield increased to around 64 percent due to improved cracking activity. The optimum catalyst loading of 10 percent produced the highest liquid yield of about 70 percent and reduced char formation to nearly 12 percent. This indicates efficient conversion of polypropylene into liquid hydrocarbons. However, increasing the catalyst loading to 15 percent resulted in a decrease in liquid yield to approximately 65 percent and an increase in gas yield to about 24 percent. This behavior is attributed to secondary cracking reactions, where liquid hydrocarbons are further broken down into lighter gaseous compounds. This confirms that an optimum catalyst loading provides a balance between primary cracking and secondary reactions.

TABLE I: Product Yield Distribution

Catalyst Loading	Liquid Oil (wt%)	Gas (wt%)	Char (wt%)
0% CaO	59	19	22
5% CaO	64	18	18
10% CaO	70	18	12
15% CaO	65	24	11

B. Product Distribution

The product distribution clearly shows the influence of catalyst loading. Char yield decreased from about 22 percent in non-catalytic conditions to nearly 12 percent at optimal catalyst loading, confirming improved conversion efficiency. Gas yield remained relatively low at moderate catalyst loading but increased significantly at higher catalyst loading due to intensified secondary cracking reactions.

C. Fuel Properties

The liquid oil produced at optimal catalyst loading exhibited properties comparable to conventional fuels. The density of the oil was approximately 840 kilogram per cubic meter and viscosity was around 3.9 centistokes, which are close to diesel fuel values. The calorific value of 43.8 megajoule per kilogram indicates high energy content. These results suggest that the produced oil can be used as an alternative fuel with minimal upgrading.

TABLE II: Comparison of Fuel Properties

Property	Pyrolytic Oil (10% CaO)	Diesel Fuel
Density (kg/m³)	840	830
Viscosity (cSt)	3.9	3.0
Calorific Value (MJ/kg)	43.8	45.0
Flash Point (°C)	42	52

IV. Sustainability And Environmental Implications

The use of eggshell-derived calcium oxide as a catalyst provides a sustainable solution for both plastic waste management and biowaste utilization. This approach reduces environmental pollution caused by plastic disposal and minimizes dependence on fossil fuels. The conversion of waste materials into valuable fuel supports circular economy principles and resource recovery. Additionally, the use of low-cost and naturally available materials makes the process economically viable for large-scale applications.

V. Conclusion

This study demonstrates that eggshell-derived calcium oxide is an effective catalyst for polypropylene pyrolysis. Catalyst loading plays a crucial role in determining product yield and distribution. The optimum catalyst loading of 10 percent produced the maximum liquid yield of 70 percent along with improved fuel quality. Higher catalyst loading resulted in increased gas production due to secondary cracking reactions. The use of waste-derived catalyst enhances the sustainability and cost-effectiveness of the process. Overall, the study highlights the potential of converting plastic waste into valuable fuel using environmentally friendly catalytic methods.

References

[1] M. S. Abbas-Abadi, M. H. R. Ghoreishy, and H. Yeganeh, “Kinetic study of LDPE pyrolysis in the presence of eggshell-derived calcium oxide catalyst,” *J. Anal. Appl. Pyrolysis*, vol. 165, p. 105560, 2022.

[2] S. Ahmed and F. Khan, “Statistical optimization of catalytic plastic pyrolysis using factorial design methodology,” *Fuel Process. Technol.*, vol. 231, p. 107215, 2022.

[3] J. Almeida and P. Costa, “Techno-economic assessment of decentralized plastic pyrolysis systems using eggshell-derived CaO catalysts,” *Renew. Energy*, vol. 219, p. 119534, 2024.

[4] S. Balaji and P. Kannan, “Development of CaO–biochar hybrid catalysts for enhanced plastic pyrolysis performance,” *Energy Convers. Manag.*, vol. 262, p. 115695, 2022.

[5] D. Chandrasekaran and K. Murugan, “Fouling mechanisms and regeneration behavior of CaO catalysts in plastic pyrolysis,” *Chem. Eng. J.*, vol. 418, p. 129376, 2021.

[6] Q. Chen, Y. Li, H. Zhang, and X. Liu, “Catalytic upgrading of plastic-derived pyrolysis oil using calcium oxide,” *Fuel*, vol. 304, p. 121353, 2021.

[7] Y. Chen and Z. Lin, “Nano-sized calcium oxide catalysts derived from industrial waste for polypropylene pyrolysis,” *J. Clean. Prod.*, vol. 389, p. 136136, 2023.

[8] S. Das and P. Ghosh, “Life cycle assessment of catalytic plastic pyrolysis: environmental benefits and emission reduction potential,” *Resour. Conserv. Recycl.*, vol. 162, p. 105041, 2020.

[9] K. Devi and R. Srinivasan, “Synthesis of eggshell-derived CaO catalyst for sustainable plastic pyrolysis,” *Waste Manag.*, vol. 146, pp. 27–36, 2022.

[10] R. Evans, T. Smith, and L. Brown, “Industrial challenges and scalability of CaO-based catalytic plastic pyrolysis,” *Prog. Energy Combust. Sci.*, vol. 81, p. 100870, 2020.

[11] H. Fatima and M. Ali, “Effect of calcium oxide particle size on oil yield and hydrocarbon distribution during plastic pyrolysis,” *Fuel*, vol. 340, p. 127439, 2023.

[12] L. Fernandez and A. Romero, “Catalyst aging and regeneration strategies for CaO in plastic pyrolysis,” *Ind. Eng. Chem. Res.*, vol. 61, no. 18, pp. 6465–6476, 2022.

[13] N. Gupta and V. Sharma, “Hydrogen-rich gas production during CaO-assisted co-pyrolysis of plastic waste and biomass,” *Int. J. Hydrogen Energy*, vol. 45, no. 54, pp. 30591–30602, 2020.

[14] T. Huang, S. Wang, and K. Zhou, “Alkali-modified calcium oxide catalysts for selective plastic pyrolysis,” *Appl. Catal. B Environ.*, vol. 327, p. 122472, 2023.

[15] N. Iqbal, R. Ahmad, and M. Hassan, “Comparative evaluation of fixed-bed and fluidized-bed CaO catalytic pyrolysis systems,” *Energy*, vol. 272, p. 127138, 2023.

[16] R. Karthik, S. Praveen, and R. Kumar, “Influence of moisture content on eggshell-derived CaO catalyst during polyethylene pyrolysis,” *J. Environ. Chem. Eng.*, vol. 12, no. 2, p. 110225, 2024.

[17] A. Kumar and R. Singh, “Comparative study of CaO and silica–alumina catalysts for plastic pyrolysis,” *Fuel Process. Technol.*, vol. 229, p. 107179, 2022.

[18] S. H. Lee and J. Park, “Catalytic conversion of mixed plastic waste into high-quality fuel using calcium oxide,” *Energy Convers. Manag.*, vol. 236, p. 114040, 2021.

[19] J. Li and H. Xu, “Modified calcium oxide catalysts for improved selectivity and fuel stability in plastic pyrolysis,” *Fuel*, vol. 336, p. 126875, 2023.

[20] G. Lopez, J. Alvarez, and J. Bilbao, “Effect of CaO catalyst loading on energy efficiency of plastic pyrolysis,” *Energy*, vol. 226, p. 120404, 2021.

[21] A. Menon, S. Krishnan, and R. Nair, “Effect of heating rate on CaO catalytic performance during plastic pyrolysis,” *J. Therm. Anal. Calorim.*, vol. 146, pp. 1831–1843, 2021.

[22] N. Miskolczi, A. Angyal, and L. Bartha, “Comparative study of metal oxide catalysts in plastic pyrolysis,” *Fuel Process. Technol.*, vol. 197, p. 106191, 2020.

[23] P. Nair and B. Joseph, “Reusability and regeneration potential of eggshell-derived calcium oxide catalysts in plastic pyrolysis systems,” *J. Clean. Prod.*, vol. 418, p. 138019, 2024.

[24] S. Nair and A. Joseph, “Biowaste-derived catalysts for thermochemical plastic waste conversion: a comprehensive review,” *Renew. Sustain. Energy Rev.*, vol. 186, p. 113657, 2024.

[25] V. Nandakumar and J. Prakash, “Life cycle assessment of eggshell-derived CaO versus limestone-based CaO catalysts,” *J. Environ. Manag.*, vol. 336, p. 117588, 2023.

[26] H. Omar and N. Yusof, “Vapour-phase reactions and gas quality enhancement in CaO-catalyzed plastic pyrolysis,” *Fuel*, vol. 322, p. 124179, 2022.

[27] D. R. Patel and P. Mehta, “Comparative performance of limestone- and eggshell-derived CaO catalysts in plastic pyrolysis,” *Waste Biomass Valorization*, vol. 12, pp. 6139–6152, 2021.

[28] N. Patni, A. Verma, and R. Joshi, “Thermochemical plastic waste management: a review of sustainable conversion pathways,” *J. Environ. Chem. Eng.*, vol. 8, no. 6, p. 104561, 2020.

[29] M. M. Rahman, S. Ahmed, and M. Islam, “Combined effects of temperature and catalyst loading on polyolefin pyrolysis,” *J. Anal. Appl. Pyrolysis*, vol. 168, p. 105731, 2023.

[30] P. Rajesh and S. Kumar, “Enhancement of oil yield in plastic pyrolysis using calcium oxide catalyst,” *Fuel*, vol. 275, p. 117987, 2020.

Performance Evaluation of an Integrated Ultrafiltration–Reverse Osmosis System for Municipal Wastewater Reuse in Industrial Applications

Vignesh S[1]
Department of Environmental Engg,
The Kavery Engineering College,
Salem,Tamil Nadu, India.
s.vigneshvicky18@gmail.com

***Abstract*— This study evaluates the effectiveness of a full-scale integrated ultrafiltration (UF) and reverse osmosis (RO) system for wastewater reuse in industrial processes. Wastewater samples were collected from different stages of treatment, namely raw influent, outlet of the moving bed biofilm reactor (MBBR), ultrafiltration permeate, and reverse osmosis permeate. Important parameters like turbidity, total suspended solids (TSS), biochemical oxygen demand (BOD), chemical oxygen demand (COD), total dissolved solids (TDS), and total coliforms were analyzed. High removal efficiency was observed from the UF-RO system, resulting in permeate turbidity less than 0.4 NTU, BOD less than 2 mg/L, COD less than 7 mg/L, and TDS less than 60 mg/L. Water quality parameters satisfied the guidelines of the Central Pollution Control Board (CPCB) for industrial purposes. However, the study lacks detailed variability and statistical analysis. It appears that UF-RO systems have the potential for a reliable solution for the reduction of fresh water consumption in industrial processes.**

***Keywords*—** *Wastewater recycling, Sustainable water management, Advanced water treatment, Water conservation, UF-RO system, Treated effluent quality.*

I. INTRODUCTION

Along with rapid urbanization, the wide expansion of industries results in the high generation of sewage in urban areas, thus seriously challenging the municipal administrations for the collection, treatment, and disposal of wastewater. At the same time, freshwater resources are becoming increasingly scarce, while urban water demand continues to rise. Therefore, under such circumstances, the reuse of treated sewage water has emerged as a vital strategy for sustainable urban water management rather than an optional practice, according to (Soni et al. 2020; Van Rossum, 2020; Global Water Intelligence, 2019). Wastewater reutilization compensates for the deficit between the amount of available water and the quantity of its need. It further helps in sustaining the water resources for a longer period. Industries are crucial for the development of cities and require huge water quantities for cooling purposes, washing, feeding boilers, and production purposes. At the same time, houses produce enormous amounts of waste water that is disposed into city sewer systems. The utilization of too much fresh water supply for industries will lead to the depletion of underground as well as superficial water resources. Similarly, improper disposal of sewage results in pollution and harms public health, according to (Duckett and Stoker, 2014; Lahnsteiner and Mittal, 2010). Utilizing treated sewage water for industries is a practical and feasible approach that would save freshwater supplies and ensure that clean water is primarily utilized for drinking purposes, states (Soni et al., 2020).

For industrial purposes, the feasibility of reusing treated wastewater was evaluated based on important parameters such as turbidity, BOD, and COD. High cloudy indicates high silty in the water, while high BOD and COD indicate organic and chemical pollution, respectively, that can harm industrial equipment as well as the environment (Stenstrom & Song, 2010). It is important to reduce these values to safe thresholds for the use of treated sewage for uses other than the original one. This, according to (Cailean et al., 2015), calls for sewage treatment to establish safe thresholds for treated sewage to be used for other uses safely. Sewage treatment comes in four forms, Specifically, the initial stage of wastewater management is basic treatment., where screens and grit chambers eliminate large objects such as plastics, rags, and other debris that could cause abrasion to equipment, thus protecting other subsequent processes. The primary treatment uses sedimentation tanks that facilitate the settling of heavy matter in the form of sludge, thus reducing the organic compound load for other stages (Stenstrom & Song, 2010). The primary treatment alone cannot provide satisfactory treatment for industrial use. The secondary treatment uses microorganisms that degrade dissolved and colloidal organic matter. Techniques used here include activated sludge, filters, and oxidation ponds that significantly reduce BOD and COD (Stenstrom & Song, 2010). Despite secondary treatment, fine solids and other contaminants remain, hence a need for more advanced treatment before reuse (Lahnsteiner & Mittal, 2010).
Advanced solutions include MBBR, ultrafiltration, and reverse osmosis, etc. are now commonly used to obtain high-quality water that is usable in industries. Membrane technologies are preferred because they require less space, are efficient, and produce clean water on a consistent basis (Azmi et al., 2025; Birniwa et al., 2024; Intini & Ceglia, 2016). Technologies such as membrane bioreactors (MBRs) have been developed that combine biological treatment with membrane filtration to efficiently remove solids as well as

organic components (Qin et al., 2009). UF and RO technologies have the potential to eliminate 95-99% turbidity, BOD, and COD, along with dissolved salt, and other pollutants, resulting in clear, odor-free water. Such treated water can be used for cooling towers, boiler feedwater after additional treatment, process wash, and irrigation purposes of landscaping. The use of UF, RO technology reduces dependence on fresh water, thus minimizing the discharge of wastewaters into rivers, lakes, and the environment, which aids in the reduction of pollution and conserving the environment (Doederer and Farre, 2014; Zhao and Song, 2010).

Additionally, reuse of treated wastewaters has economic advantages for industries, which are useful in the long run. The cost of recycling wastewaters is generally lower than that of procuring, treating, and transporting fresh water, which enhances water security and minimizes expenses (Zawadzki and Kolczak, 2023; Soni et al., 2020). Reuse of treated sewage is, therefore, a contributory factor toward saving water, reducing pollution, and offering assistance in the sustainable development of cities, industries, as well as growth. This research explores in detail the treatment process of sewage, thereby identifying modern technologies that efficiently reduce turbidity, BOD, and COD to satisfactory levels for industries.

II. Materials and Methods

The climate of the study region is semi-arid and tropical with moderate to hot temperatures, with seasonal rains that are impacted by the southwest and northeast monsoons. However, owing to the growth and urbanization of the population, along with the increase in industrial and factory activities, there is a rise in wastewater and pressure on the freshwater resources. The STP also collects domestic sewage produced from residential, commercial, and institutional areas lying within the urban boundaries. The characteristics of the influent wastewater are dominated by high levels of organic and microbial contamination, and hence it is representative of representative of typical domestic sewage found in an Indian environment.

A. Study Area

The treatment plant is expected to operate under varying hydraulic and organic loading rates and work under continuously operating mode treating wastewater collected from urban areas. In this current study, sewage water samples are collected at various points in the wastewater treatment process, including raw sewage intake and post-processing units, in order to determine the effectiveness of the integrated treatment process adopted in this study. The treated sewage from the advanced UF-RO process is planned for re-use in industrial applications such as cooling towers, boilers (with a minor polishing process), chillers, and many other non-potables uses in different industries. The choice of this location in this study gives a practical determination of the viability of membrane sewage recycling as a useful means of addressing water scarcity in India. (Especially in Tamilnadu).

B. Sampling of Wastewater samples

- Wastewater samples were collected to determine the effectiveness of the overall sewage treatment and reuse system. This involved sampling at five points along the treatment system to test varying levels of water quality based on treatment level, and these stages include raw sewage entry to treatment, secondary biological treatment (MBBR stage), tertiary treatment of ultrafiltration permeate, and reverse osmosis permeate stages.
- The physico-chemical parameters were observed on a weekly basis in order to determine treatment efficiency. The operational parameters for the UF and RO processes were recorded on a daily basis. The heavy metals and microbiological parameters were observed on a monthly basis to check if the process is within regulation. The whole experimental setup was carried out for a period of six months. This includes both steady state and fluctuating conditions.
- The analysis of both wastewaters and treated water is done following the guidelines laid by CPCB, a body governed by the Government of India. The guidelines followed by CPCB are mainly APHA Standard Methods for Examination of Water & wastewaters. These methods are mainly developed according to Indian Standards (IS codes) for precision and duplicability.
- The samples for wastewater and treated water were analysed for pH values, turbidity, TSS, total TDS, BOD, COD, and total coliform.

C. Principle of Treatment

A Sewage Treatment Plant employs methods such as biological, chemical, ultra-filtration, and RO for cleaning the water. The process removes pollution from the water and reduces the TDS. In the next step, the plant makes the water fit for recycling. The primary motto of a Sewage Treatment Plant is to accelerate and replicate the natural purification process that takes place in nature, thereby converting sewage into clean water.

Another significant aim is to promote the reuse of purified water in industries. The research employed a direct membrane-based tertiary treatment process, whereby the aeration tank effluent is directly passed through ultrafiltration followed by reverse osmosis, bypassing other conventional treatment processes such as pressure sand filters and activated carbon filters. The rationale for this is based on the superior separation characteristics of membrane filtration processes, which provide superior removal of turbidity, colloidal, and dissolved impurities. The process, however, increases the rate of membrane fouling, which requires more frequent backwashing and cleaning. This is a major drawback associated with the process. The process, however, provides permeate water suitable for reuse, thus demonstrating the process's ability to provide advanced treatment.

The Sewage treatment system has three steps for handling wastewater, namely: mechanical treatment, biological treatment, and advanced treatment. In the mechanical treatment, the STP removes large solid particles, grit, fat, oil, grease, and sludge from the waste by screening and percolating the waste, thus enabling the heavy particles to settle at the bottom. The secondary treatment at the STP employs technologies such as Moving Bed Biofilm Reactor (MBBR), which helps to reduce BOD, COD, and other pollutants with the use of living organisms. The tertiary

treatment, also known as tertiary purification, employs disinfection, ultrafiltration, and reverse osmosis.

D. Advanced Membrane Treatment

The treated Sewage get polished through ultrafilter and reverse osmosis process and the same is reused at Industrial, Agriculture, Construction works, etc.… in the place of raw water. Sewage treatment plants are not only protecting our environment by removing contaminants, treated sewage is reused in above mentioned areas to save raw water for our future generations.

The performance of the membrane systems was analysed on the basis of the most important operational and fouling factors. While using Polypropylene (PP) capillary membranes are highly effective for sewage treatment plant (STP) processes, especially when MBBR technology is employed owing to their excellent chemical stability, porosity, and mechanical strength. The porous structure of these membranes (>65% porosity) enables a high water flux rate.

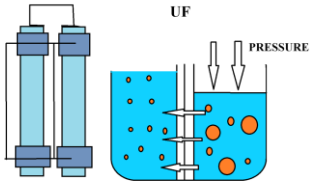


Fig 1- Ultra filter

The UF system was running at a Trans Membrane Pressure (TMP) range of 0.7-1.3 bars, which indicated the UF system had an effective antifouling strategy. The automatic hydraulic backwash cycle was conducted every 30-45 minutes to remove accumulated suspended solids. To ensure membrane hydraulic flux was maintained, the UF system employed a technique called Chemical Enhanced Backwash (CEB), conducted weekly with sodium hypochlorite at a concentration of NaOCl at 200 mg/l for antifouling purposes, as well as HCl acid at a 3% concentration carried out monthly for anti-scaling purposes.

The integrated UF–RO treatment system achieves substantial pollutant removal. Typical treated sewage characteristics include, Pressure sand filter which contains dual media (sand and anthracite) is designed based on the following,

TABLE I. DESIGN BASIS OF UF

Sl. No	Parameters	Unit	Value
1.	Number of UF Skids	Nos	7 (6W +1S)
2.	Nominal recovery rate	%	~89.5
3.	Filtration Flux	GFD	50.7 LMH
4.	Permeate Flow	m^3/hr	390
5.	Silt Density Index (UF Outlet)	-	≤ 2

RO The RO process was functioning at a feed pressure varying from 11 to 15 bar with a salt rejection efficiency of 97-99%. This is an indicator of successful salt rejection. Also, the decrease in the normalized permeate flowrate did not exceed 10% during a period of six months, which is an indication of properly managed fouling.

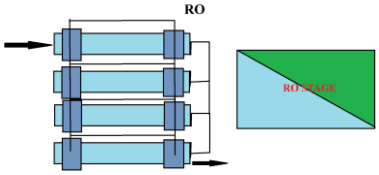


Fig 2- Reverse osmosis

Additionally, the Clean-in-Place (CIP) process using chemicals was necessary only once every 3-4 months. The ultrafiltration (UF) permeate is collected in the UF permeate tank, serving as a storage buffer before further treatment. RO feed pumps are used to transfer the UF permeate through a cartridge filter, which removes any remaining fine particles, before it enters the RO skids for final treatment and purification. RO system is designed based on Following,

TABLE II. DESIGN BASIS OF RO

Sl. No	Parameters	Unit	Value
1	Number of trains	Nos	5 (4 duty and 1 standby)
2	Nominal flow per train	m³/day	580
3	Average flux (l/m²/hr.)	l/m²/hr.	15.8
4	Nominal recovery rate	%	85
5	Conductivity	ppm	< 75
6	BOD	ppm	BDL
7	COD	ppm	BDL

The disinfection is the destruction, inactivation, or removal of pathogenic microorganisms that are present in the treated wastewater. This treatment ensures that the bacteria, viruses, and microorganisms that are harmful are destroyed or made incapable of growing and multiplying, which helps to ensure that the treated effluent is safe for release. This is because it would not affect the health of people who come into contact with it.

Types of Disinfection
1. Chlorination
2. Ultra Violet Radiation
3. Ozonation

III. RESULTS AND DISCUSSION

This study makes a contribution to the existing literature by evaluating a full-scale UF–RO system, which incorporates real municipal wastewater conditions in India. Unlike the existing literature, which has used UF–RO

systems in a laboratory and pilot scale, the current study evaluates the long-term operational performance of the UF–RO system, including the fouling characteristics and cleaning mechanisms. In addition, the study evaluates the CPCB reuse criteria and the reuse applications, which bridge the gap between the laboratory and industrial applications.

TABLE III. WATER QUALITY RESULTS

Parameter	Raw Sewage	MBBR Outlet	UF Permeate	RO Permeate
PH	8.5	8	6.5	6.5
Turbidity (NTU)	230	90	<1	BDL
TSS (mg/l)	250	28	<5	BDL
BOD (mg/l)	260	30	<5	<2
COD (mg/l)	500	80	<13	<7
TDS (mg/l)	1800	1880	1900	~60
Total Coliform (MPN/100 ml)	$>10^6$	10^4	<10	BDL

Membrane separators make a highly efficient two-stage advanced treatment process commonly employed in industries for preparing high-quality recycled water. UF is a most essential part of the treatment process acting as the primary barrier that removes suspended matter, turbidity, colloidal matter, and microbial agents such as bacteria, viruses, etc. This helps prevent the potential for fouling of the RO membrane, which essentially leads to better efficiency of the Reverse Osmosis plant, increased lifespan of the membrane, reduced need for cleaning, as well as reduced operating costs. After pre-treatment, when the water passes from the UF stage, it is further purified by the Reverse Osmosis plant, which strips dissolved salts, TDS, organic compounds, and other pollutants from the water. The Reverse Osmosis plant works under a high degree of pressure, letting only the molecules of water diffuse through a semipermeable membrane, thus rejecting a major component of ions, as well as various impurities.

The UF-RO hybrid treatment system helps produce water that is useful for a variety of industrial purposes. In Heating, Ventilation, and Air conditioning (HVAC) systems, treated water helps facilitate efficient thermal exchange, inhibits scaling, corrosion, and bacterial growth in chilled water systems. In process water, a uniform quality is needed to maintain the quality of the product, and UF-RO-treated water is helpful in this regard. For cooling towers, hardness reduction and low TDS are highly favorable, which helps reduce scaling, increasing efficiency and reducing the need for blow-down cycles. Further polishing steps in demineralization such as Mixed Bed (MB) can further make the RO permeate water useful as high-purity feed water in high-pressure boilers. Along with meeting process conditions, UF-RO systems contribute to making industries sustainable on a long-run basis. The re-use of treated wastewaters significantly cuts down on the reliance on fresh water sources, which plays a very important role in saving natural resources, thereby meeting ecological regulations. It is supported by the fact that the total quantum of wastes discharged back into the environment is reduced, which alternatively helps in increasing protectionist mechanisms on an ecological front. In a nutshell, the combined technologies of UF and RO achieve an efficient, reliable, and sustainable manner of producing high-quality water suitable for modern industries; hence, it is a contribution to protecting the environment.

A. *Environmental and Managerial Implications*

The implemented integrated UF–RO-based sewage reuse system reduced freshwater withdrawal for industrial use by 70%, reducing the demand on surface and groundwater resources. In addition, this system significantly reduced the discharge of partially treated wastewater into natural water courses by reusing treated municipal sewage for industrial purposes, thus improving the local environmental quality and reducing the potential for deterioration of natural aquatic ecosystems. The high quality of the treated water consistently met or exceeded CPCB guidelines for industrial reuse, and it was reliably employed with cooling towers, boilers-once minimal polishing was done-and other process utilities. Water management-from a system redesign that was based on water reuse-increased industrial water security by supplying a non-seasonal, reliable alternative source of water decoupled from freshwater supply limitations. Managerially, the system reduced dependency on groundwater abstraction, supported compliance with environmental regulations, and contributed to long-term operational resilience. Besides, advanced membrane-based reuse technology allowed better planning of water resources, reduced vulnerability to regulatory tightening, and aligned industrial operations with national sustainability.

B. *Comparative Performance and Implications for Sustainable Water Reuse*

The findings achieved through the current study clearly illustrate a set of advantages that cannot be achieved or have not been achieved through previous research works endeavoring wastewater treatment and reuse. In fact, most previous research works carried out on wastewater treatment and reuse have been focused on laboratory-scale works, pilot plants, or membrane processes, and, hence, the current work attempts to investigate a comprehensive, full-scale wastewater reuse scheme that is actually working under actual, dynamic conditions.

Compared to the traditional secondary and tertiary treatment processes reviewed by (Teodosiu et al. 2015 and Duckett and Stoker 2014), which mostly focus on reducing organic matter and eliminating pathogens, this study showed the ability to produce high-quality reusable water. The UF-RO combination showed >99% removal of turbidity and TSS,

>95% removal of TDS, BOD <2 mg/l, COD <5 mg/l, and the removal of coliform bacteria to below detectable levels, compared to the reliability of water reuse mentioned by non-membrane tertiary treatment methods. Various studies by (Qin et al. in 2006, 2009, Aydiner and Demircioglu in 2014, and others by Al-Shaeli and Al-Juboori in 2022) have found the treatment efficiencies to be high in the UF-RO or the MBR-RO combined processes. However, the focus was on the treatment efficiencies or the techno-economic viability instead of the long-term regulatory requirements.

This study, on the other hand, takes pains to emphasize the CPCB standards required for the reuse of the treated water in industry, such as cooling towers or boilers. Moreover, as compared to most other studies, which reported a higher treatment cost without a specific contextual explanation, it has been proven that despite a higher treatment cost, a 70% reduction in water withdrawal, reduced reliance on groundwater, smaller plant size, and improved water security are achieved by the UF RO system. Additionally, a stable performance level of the membrane, as measured by a normalized permeate flow decline value of <10% and a reduced frequency of CIP, imply higher robustness levels, which are generally unreported.

In the context of these findings, it can be concluded that the present study not only focuses on proving the efficiency of the treatment in its study but also offers a detailed investigation of wastewater reuse in the field by covering all the aspects of efficiency in a broader sense. This multi-faceted approach appears to be an innovative feature of the present study in comparison to previous literature.

C. Conclusion

Sewage recycling systems hold paramount importance with respect to environmentally responsible management of water, especially in the context of rising challenges to fresh water availability associated with population expansion and industrialization, as well as increased variability in the weather. In this context, this study proves that the use of high-tech sewage recycling involving ultrafiltration and reverse osmosis technologies is quite successful in converting sewage into an alternative source for industrial use and eliminating environmental hazards. Results of quantitative performance analysis indicated that the combined UF-RO system removed turbidity by around 99% and attained a level of <0.5 NTU in the RO permeate, starting from 230 NTU in raw sewage. Removal efficiency was excellent in case of organic matter, and BOD was reduced from 260 mg/l to ~2 mg/l, while COD was reduced from 500 mg/l to ~7 mg/l. Removal of dissolved salts was also satisfactory, and TDS in the RO permeate was ~70 mg/l, which is as per CPCB norms for highly specialized industries. Decimation of pathogenic microbes was also attained. The project allowed for the reuse of about 4000 m3/day of municipal sewage, thus reducing the usage of freshwater by 70% for industrial processing purposes. This significantly reduced the dependency on the normal freshwater source and, thus, the discharge of the wastewater into natural water resources. Moreover, the project of reusing the sewage helped in relieving the current water supply infrastructure, thus postponing the expansion of the costly network.

In conclusion, UF-RO based sewage recycling systems are found to offer a pragmatic, scalable, and sustainable solution to global water scarcity. The system is capable of providing a constant supply of high-quality recycled water, thereby enabling water security, sustainability, and protection of the environment.

REFERENCES

[1] Al-Shaeli, M., Al-Juboori, R.A. 2022. Recent improvements and future potential for applying renewable and reclaimed resources for responsible membrane development. *Science of the Total Environment* 838: 156014. https://doi.org/10.1016/j.scitotenv.2022.156014

[2] Azmi, L.S., Ain Jabit, N. 2025. Evaluation of membrane separation techniques for effective processing of hazardous heavy-metal-enhanced industrial wastewater. *Desalination and Water Treatment* 323: 101321. https://doi.org/10.1016/j.dwt.2025.101321

[3] Birniwa, A.H., Habibu, S. 2023. Membrane-based systems for removing contaminants from freshwater and industrial wastewater. Case Studies in *Chemical and Environmental Engineering* 9: 100538. https://doi.org/10.1016/j.cscee.2023.100538

[4] Doederer, K., Farré, M.J., Keller, J. 2014. Removal of disinfection by-product residues by RO/NF membranes: Influence of operational conditions. *Journal of Membrane Science* 467: 195-205. https://doi.org/10.1016/j.memsci.2014.05.019

[5] Duckett, D., Stoker, P. 2014. Creating a wave: Supporting freshwater recycling instead of limiting it. *Water Policy* 16(5): 912-926. https://doi.org/10.2166/wp.2014.135

[6] Goh, Pei Sean, Ismail, Ahmad Fauzi, Chong, Ng Boon. 2016. Nanocomposite membranes for advanced water reuse applications. *Desalination* 391: 42-51. https://doi.org/10.1016/j.desal.2016.03.012

[7] Grant, Steven B., Saphores, Jean-Daniel, Feldman, David L., Hamilton, Andrew J., Fletcher, Tim D., Cook, Paul L.M., Stewardson, Michael J., Sanders, Brian F., Levin, Lisa A., Ambrose, Richard F., Deletic, Ana. 2012. Taking the "waste" out of wastewater for human water security and ecosystem sustainability. *Science* 337(6095): 681-686. https://doi.org/10.1126/science.1216852

[8] Hocaoglu, Sait M., Demir, Burcu, Ozkan, Mustafa. 2018. Reverse osmosis concentrate management strategies: A review. *Desalination* 439: 123-135. https://doi.org/10.1016/j.desal.2018.04.001

[9] International Water Intelligence. 2019. Portugal develops an integrated plan for wastewater reuse. GWI Brief 367.

[10] Intini, G., Ceglia, A. 2016. Wastewater recycling in the petroleum sector: UF followed by RO membrane corrosion issues. *Desalination* 392: 89-98. https://doi.org/10.1016/j.desal.2016.01.013

[11] Jafari, Mohammad, Shirazi, Saeed, Ramezani, Reza. 2020. Strategies for controlling fouling in

membrane-based wastewater reuse systems. *Journal of Water Process Engineering* 36: 101262. https://doi.org/10.1016/j.jwpe.2020.101262

[12] Jeanmayer, J.P., Suti, H. 2007. Expandable microfiber filter components for seawater purification. Technological Hydrology 56: 157-165.

[13] Kim, J.J., Hong, J., Yoon, H. 2013. Lightweight pretreatment combining microfiber media and ultrafiltration for highly turbid saline water. *Desalination* 313: 28-35. https://doi.org/10.1016/j.desal.2012.12.014

[14] Kim, Sangho, Kim, Jaehoon. 2019. Industrial water reuse using advanced membrane technologies. *Journal of Cleaner Production* 231: 121-135. https://doi.org/10.1016/j.jclepro.2019.05.195

[15] Kumar, A., Soni, M. 2020. Applications and impacts of sewage remediation and reclamation. *Water, Air, & Soil Pollution* 231: 228. https://doi.org/10.1007/s11270-020-04577-1

[16] Lazarova, Valentina, Brissaud, Florent, Angelakis, Aristide N. 2012. Water reuse for sustainable water management in water-stressed areas. *Water Science and Technology* 65(11): 199-206. https://doi.org/10.2166/wst.2012.907

[17] Lee, J.J., Chan, J.H., Benam, R., Han, K.B., Kim, C.W. 2008. Microfiber filters as alternatives to flocculation and sedimentation. *Desalination* 231: 323-331. https://doi.org/10.1016/j.desal.2007.12.020

[18] Lee, J.J., Johir, M.A.H., Cheenu, K.H., Shawn, H.K. 2009. Mixed pretreatment method for seawater reverse osmosis. *Desalination* 247: 15-24. https://doi.org/10.1016/j.desal.2008.12.019

[19] Lee, Seungkwan, Kim, Youngjin, Park, Hyungseok. 2015. Application of low-pressure membrane systems for municipal wastewater reuse. *Water Research* 73: 17-26. https://doi.org/10.1016/j.watres.2015.01.021

[20] Liu, Yong, Zhang, Hui, Wang, Zhenyu. 2019. Sustainability assessment of wastewater reuse systems using life cycle analysis. *Journal of Environmental Management* 231: 293-303. https://doi.org/10.1016/j.jenvman.2018.10.065

[21] Mi, B., Elimelech, M. 2010. Organic fouling of reverse osmosis membranes. *Environmental Science & Technology* 44: 2022-2028. https://doi.org/10.1021/es903459j

[22] Mi, Baoxia, Elimelech, Menachem. 2010. Organic fouling of reverse osmosis membranes: Mechanisms and mitigation. *Environmental Science & Technology* 44(6): 2022-2028. https://doi.org/10.1021/es903459j

[23] Mittal, R., Lansteiner, J. 2010. Industrial wastewater recycling using layered treatment systems. *Water Science and Technology* 62(8): 1813-1820. https://doi.org/10.2166/wst.2010.382

[24] Natarajan, P., Jayavel, K., Somasundaram, S., Ganesan, S. 2020. Treatment of caustic industrial effluents using biological and catalytic methods. Environment, *Development and Sustainability* 22: 4975-4990. https://doi.org/10.1007/s40974-020-00178-y

[25] Qin, J.J., Kekre, K.A., Tao, G., Seah, H. 2009. MBR-RO process for municipal wastewater reuse. *Journal of Membrane Science* 272: 70-77. https://doi.org/10.1016/j.memsci.2005.07.041

[26] Qin, J.J., Wai, M.N., Kekre, K.A. 2006. Dual membrane system for industrial wastewater reclamation. *Separation and Purification Technology* 50: 380-387. https://doi.org/10.1016/j.seppur.2005.12.008

[27] Rekhate, C.V. 2020. Ozone-based advanced oxidation processes for sewage treatment. *Environmental Technology Reviews* 9: 35-52. https://doi.org/10.1080/21622515.2020.1723794

[28] Shannon, M.A., Bohn, P.W., Elimelech, M., Georgiadis, J.G., Mariñas, B.J., Mayes, A.M. 2008. *Science and technology for water purification*. Nature 452: 301-310. https://doi.org/10.1038/nature06599

[29] Shannon, Mark A., Bohn, Paul W., Elimelech, Menachem, Georgiadis, John G., Mariñas, Benito J., Mayes, Anne M. 2008. Science and technology for water purification in the coming decades. *Nature* 452: 301-310. https://doi.org/10.1038/nature06599

[30] Stenstrom, M.K., Song, S.S. 2010. Oxygen transfer limitations in activated sludge systems. *Water Research* 44: 208-219. https://doi.org/10.1016/j.watres.2009.09.020

[31] Tang, Chuyang Y., Yang, Zhenyu, Wang, Zhiwei. 2011. Membrane fouling mechanisms and control strategies in water reuse applications. *Journal of Membrane Science* 287(2): 146-156. https://doi.org/10.1016/j.memsci.2006.10.032

[32] Teodosiu, C., Barjoveanu, G., Cailean, D. 2015. Role of membranes in improving ultrafiltration efficiency. *Desalination* 356: 147-158. https://doi.org/10.1016/j.desal.2014.10.031

[33] Van der Bruggen, Bart. 2018. Sustainable membrane technology for water and wastewater treatment. *Chemical Engineering Journal* 334: 1-10. https://doi.org/10.1016/j.cej.2017.10.104

[34] Van Rossum, T. 2020. Water conservation in Canada: Past, present, and future perspectives. *Water Policy* 22: 110-125. https://doi.org/10.2166/wp.2020.023

[35] Werber, Jeffrey R., Osuji, Christopher O., Elimelech, Menachem. 2016. Materials for next-generation desalination and water reuse membranes. *Nature Reviews Materials* 1: 16018. https://doi.org/10.1038/natrevmats.2016.18

[36] Wu, J.H., Chan, S.S. 2024. Fiber filters and ultrafiltration supporting RO for sewage reuse. *Sustainability* 14(9): 5443. https://doi.org/10.3390/su14095443

[37] Zawadzki, P., Konczak, B. 2023. Municipal sewage recycling for electrolysis-based gas generation. *Measurement* 216: 112928. https://doi.org/10.1016/j.measurement.2023.11292

Flow Modelling of Aquifer for tracing out of contaminants in karst topography

Pradeepa.F
Department of Civil Engineering
The Kavery Engineering College
Salem,Tamilnadu,India
email:pradeepa400724@gmail.com

Kavisri Manikannan.M
Department of Civil Engineering
The Kavery Engineering College
Salem,TamilNadu,India
email:
kavisrimanikannan1985@gmail.com

Abstract

Aquifer flow modelling is a mathematical modelling represents the water flow through the subsurface formations of rocks called aquifers. Modelling helps to predict the water level, flow directions and water quality parameters. Groundwater modelling estimates the availability of water sources and also the Contamination transport mechanisms within the formations. Hydraulic parameters such as permeability, Transmissivity and storativity decides the transporting mechanisms of the contaminants. Flow prediction complexity in karst topography is due to the highly heterogeneous and non-linear nature of its underground pathways mostly difficult to predict, which includes a network of fractures, and large conduits sometimes vein like formations within the soft rock formations. This creates multiple, complex flow paths and storage systems with rapid, unpredictable responses to rainfall, making it difficult to model accurately. Challenges are further exacerbated by the interpolating multidimensional geological, drainage and satellite GIS data for ensuring its accuracy. The study area is Tirunelveli and Tuticorin district which shows exceptional flow and drainage pattern studied previously. It was identified that the unusual drainage pattern associated with karst topography based on the geophysical studies called VES (Vertical Electrical sounding). For detailed modelling of subsurface formation and selected contaminant Calcium is selected and traced out through the karst to determine the flow mechanisms and discharge prediction of freshwater through the caverns by using MODFLOW MODELMUSE software and GIS. The origin of a selected contaminant from the geochemical data of the study region can be traced out and suitable remediation methods for treating was suggested.

Keywords:MODFLOW MODELMUSE,GIS, transmissivity, storativity, karst, Calcium

INTRODUCTION

Groundwater systems in karst terrains play a crucial role in sustaining ecosystems and human water demand, particularly in regions dominated by limestone and dolomite formations. However, the hydrogeological behaviour of karst aquifers differs significantly from that of granular porous media due to the presence of dual or triple porosity domains—matrix, fractures and conduits that result in non- Darcian and turbulent flow conditions**.** The rapid transport of water and solutes through conduits makes karst aquifers highly susceptible to contamination, posing serious challenges to water resource management.

By integrating field data, geophysical methods, and numerical simulations**,** this research contributes to improved understanding of the hydrogeological processes that govern contaminant dynamics in karst aquifers and supports the formulation of effective groundwater protection strategies.Tracing out of contaminants and flow modelling using three-dimensional mathematical modelling is actually a combined method that represent natural dynamic system into simplified assumed system for convenient mapping and prediction. Model's validity and accuracy are based on the available data from various sources and dimensions such as geophysical, geochemical, isotopic and geographical data from GIS. Chemical analysis of groundwater samples for specific selected pollutants to be traced combined with numerical modelling to determine the possible geohazard in the study region. Aayankulam village of Tirunelveli district previously got attention from researchers because of its geological importance. Exceptional flow and drainage pattern in the hard granulitic terrain showed strange signs of unstable geology. Hydrologically the magical well was previously used as the groundwater recharge well without even knowing about the geology. Numerical models can predict its complex vein formation within the sedimentary basin and its dissection with coastal region can be point out by tracing a selected pollutant.

Many researches were extensively studied about the dissolution of calcareous rocks and karstification across the globe and because of its hydrological and constructional importance. Karst is considered as the effective source of water storage and supplies 40 to 50% of the World's drinking water needs. Karst aquifers are characterized by solutionally enlarged conduits, fractures, and voids that develop in soluble rocks such as limestone and dolomite(1). Field investigations in European karst terrains (2,3) and North American systems (4) have demonstrated that contaminants can travel several kilometres within

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

hours following heavy rainfall events. According to (5) there should be significant bending, fractures, or faults in the limestones. Water may permeate porous limestones, causing the entire mass to become unstable and prone to collapse. MODFLOW6 currently includes the Groundwater Flow (GWF) Model and the Groundwater Transport (GWT) Model each with packages to represent surface water processes, groundwater extraction, external boundaries, mass sources and sinks, and mass sorption and reactions. Hydraulic studies have shown that recharge variability and conduit geometry strongly influence flow velocity and storage within karst systems (6,7). They focus (8) on the preliminary assessment of potential aquifer pollution from contaminated sites. A common approach to aquifer risk assessment considers a groundwater body to be contaminated if the simulated pollutant concentrations at a predefined control Point of Compliance (POC). The groundwater contamination transport mechanism has concealment and irreversible in nature which makes difficult to predict the sources in time once contamination occurred. Spreading of contaminants involved various mechanisms in the subsurface flow in aquifer(9). Therefore for the effective management to reduce the deeper and wider range inclusion tracing out of contaminant source is much important(10,11)

LOCATION MAP

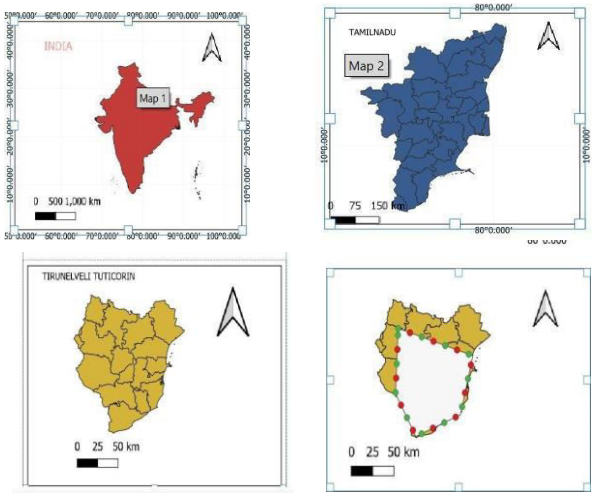


Fig 1 Location map of the study area

GEOLOGICAL AND GEOPHYSICAL STUDIES

The gravity anomalies over the Southern granulitic terrain were studied thoroughly from the Bouguer Anomaly map produced by NGRI (National Geophysical Research Institute) as a part of Geophysical mapping program. The map is coarsely spaced and the variation in Bouguer gravity ranges from -30 to -10mGal along the coast towards the inland. The clear variation in gravity and gravity breaks are visible and marked as top and bottom boundaries of Achankovil shear zone. The gravity high at the western boundary of shear zone is clearly visible. The NW-SE pattern is seen along the entire southern portion of Achankovil shear zone and the khondalite rock also shows the NW-SE directional alignment towards the shear zone. Based on a study done by seismic method (Laxmidhar et al.,2011) an anomalous zone of exhumation near AKSZ having transpression of exhumed rocks at mid-to-lower crustal level (20– 30 km) with significant underplating and mantle upwelling forming a complex metamorphic province. The presence of shear zones with high-grade charnockite massifs in the upper-crust exposed in several places indicate large scale exhumation of granulites during the Pan-African rifting (~ 550 Ma) and explains important processes involved in plume-continental lithosphere interaction with reconstruction of the Gondwanaland.

MAGNETIC MAP STUDIES

Aeromagnetic anomalies over the southern granulitic terrain were analysed by focusing the Achankovil shear zone as the Karst topography of the study area closely lying near to the shear zone. As the groundwater flow controlled by the tectonic boundaries detailed study is needed about the structural boundary and its extension. The complete south of the Achankovil shear zone in aeromagnetic data shows magnetic high susceptible rocks as the aeromagnetic focuses and enhances the deeper magnetic sources. The clear magnetic highs on north as well as south of the Achankovil shear zone demarcated the clear boundaries and width of the shear zone with magnetic lows in-between. The trend of NW-SE of the Achankovil shear zone follows towards the south due to the presence of khondalites belts and its alignment follows the same trend. The tectonic boundary between the Trivandrum block and the Nagercoil block can also be delineated by the trend change of NW-SE toward N-S at the middle of the two blocks as marked in map shown in below. In aeromagnetic map the khondalite belts shows magnetic high and at the eastern and southern ends shows higher magnetic values as tectonically these are the regions with emplacement of magmatic intrusions.

ANALYTIC SIGNAL MAP ANALYSIS

The analytical map was analysed from the previously published work (12) for the delineation of Achenkoil shear zone to predict the origin of limestone formation commonly observed in Northern Tirunelveli and border of Tuticorin districts. As per the analytic signal map our study area crossed by the shear zone particularly near the region of magical wells of Aayankulam-muthumuthanmozhi region.so the shear zone can act as the barrier to arrest the towards south and act boundary between two contrasting susceptible rocks at north and south.

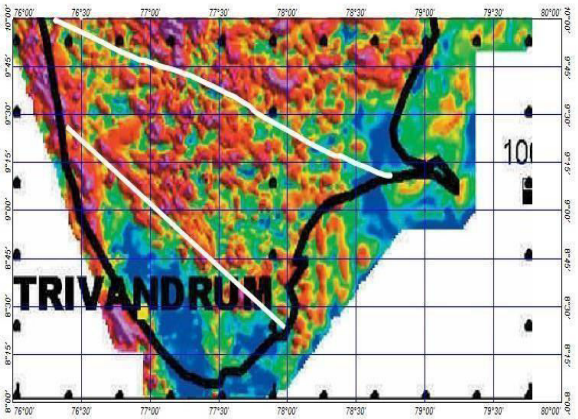


Fig 2 Analytic map of the area of interest (Anand S.P et al.,)

Electrical resistivity mapping

The resistivity of the region of interest was surveyed and mapped to find out the resistivity variations at deep to predict the aquifer characteristics. The available lithology data combined with electrical resistivity variations for the higher accuracy prediction. The 2D resistivity mapping was performed using IPI2WIN software after computing the subsurface resistivity of different layers. The Caverns was spotted in the 2D image of resistivity gridding due to sudden break in resistivity due to the caverns. The array used for VES data collection was Schlumberger array and profiling was done at nine data points with 3 points in single array.

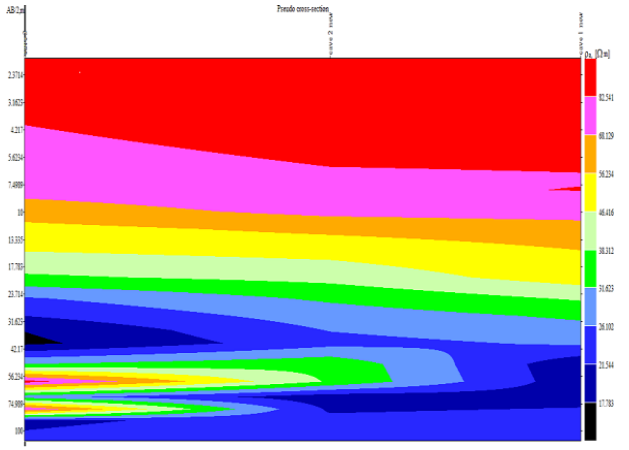


Fig 3 2D electrical resistivity mapping for 3 VES points

Water level maps

The groundwater level data was obtained from the TWAD board of Tirunelveli district. The data was mapped to various timeframe for five consecutive years starting from 2020- 2024.The water level was monitored by TWAD board for two times per year at March and June months.

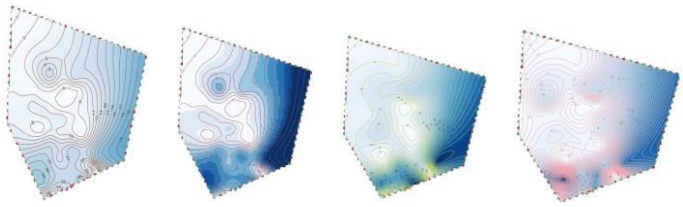


Fig 4 water level for March and June months 2020,2021

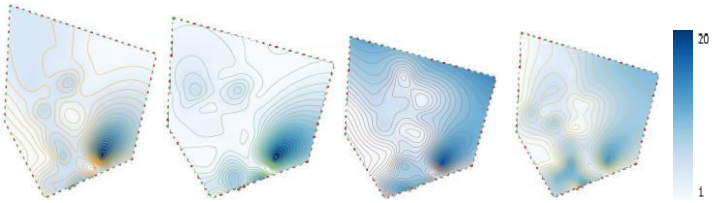


Fig 5 water level map using GQIS for march and June of 2022,2023

As the water level variation contouring, gives possible insights towards the variations of water level at various time scale and the short variations in water level suggests the tectonic boundary within the region based on the contour map (due to sharp elevational changes) and its diversified characteristics. Most of the contour maps have common water level depth that is deeper water level along the eastern side of the contour nearer to the coast. Southern deeper water level was separated with the eastern side deep water level by a tectonic feature we can termed it as a lineament.

AQUIFER MODELLING USING MODELMUSE

In aquifer modelling the study region (polygon) first infused in MODELMUSE with suitable elevation contours to get flow details. The study area polygon was prepared from QGIS software from the base map of Tirunelveli and Tuticorin districts. Initially the modelling of aquifer was done with three aquifer layers below the model top. In order to get focused region for mapping karst features a separate modelling was done with the region closely associated with the magical well and nearer by wells. The drain which supplies the excess water from the large pond storage is also included as surface drain at model top. In detailed model of the area of interest which was previously studied geophysically by electrical resistivity method and available lithology data and surface mining data of the surrounding region. Boundary conditions are fixed by two different heads boundary of the model at left and right sides. The initial head is given as 5m and the final boundary with the head as 0m as the coastal layer of zero elevation. Apparently, the flow will be from higher head 5m to the lower head towards the coast. As the study region lies in the Achankoil shear zone which acts as the structural boundary arrests the flow away from it.

The linear extension towards the coast was well studied by the available geophysical data sources such as gravity, magnetic data sources. As this tectonic boundary plays significant role with the limestone formation and its extension towards the south was studied carefully. After the geophysical analysis this limestone formation may not have the southward extension as it was bounded by the shear zone. The associated limestone formation was formed as a result of thrusting of ocean basin along the shear zone particularly at the northern Tirunelveli district which has granulitic limestone formations. Significantly by considering the coastal features associated with the eastern coastal and inland region marine deposits suggests the possible high coastal aggradation from the past leads to the shallow water level limestone deposits along the Eastern coast. This marine originated limestone formations might have developed the karst topography along the Tirunelveli and Tuticorin district coastal region towards the inland for more than 5km span from the current coastal line. More oceanic evidences of shallow marine environment in Tirunelveli district were previously recorded and studied.

.

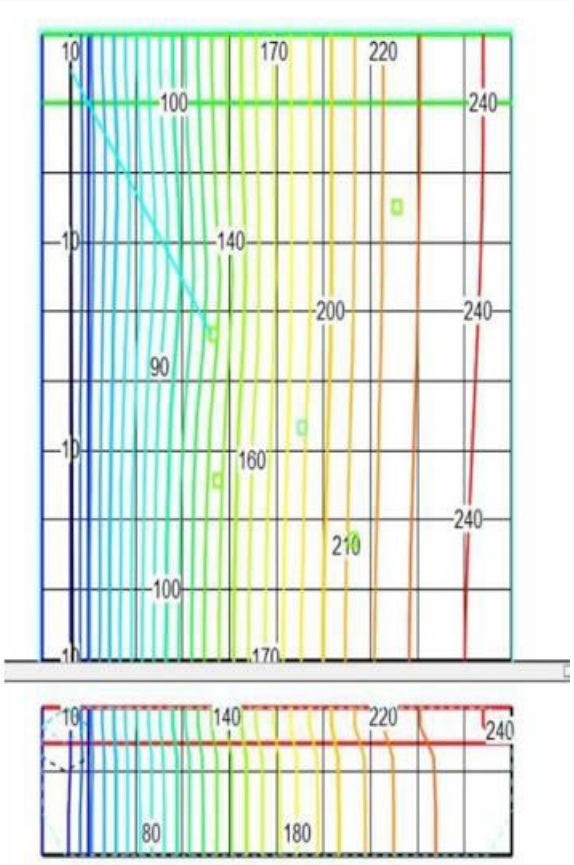


Fig 6 Result of aquifer modelling

The depositional limestone formation may have the depth of formation extends deeper when moving closer to the coast .So based on the study of remote sensing images the limestone formation was controlled by the shear zone and ends that formation to the region closely found at near to the village Uvari which lies along the district boundary of Tirunelveli and Tuticorin.so the possible discharge of karst transported water may be dumped here to the ocean by the subsurface pathways. The veins made by the erosion of limestone possibly gets wider as the thickness of limestone deposits will be high towards the coast. It Is like a subsurface river dumps thousands of m^3 of water when it was discharged towards the magical well by the surface drain. This karst well make recharge the wells in its paths towards the coast follows the same EW alignment of the Achankovil shear zone.

Entering into the modelling procedure another 4 wells other than magical well are modelled with 3 layers of aquifer. Initial aquifer of layer 2 below the model top was taken as 10m depth and the next aquifer layer will be at 20m depth. The aquifer 1 is considered as the karst associated sandstone layer with veins of limestone intercalations. The second aquifer will be the sandstone formation without limestone and the final layer will be the fractured rock zone of deep-water source for the bore wells. The horizontal hydraulic conductivity was taken as 1.5 to 3m/s due to the fastest flow of water through the caverns. The secondary sandstone layer was taken as having hydraulic conductivity as 0.001m/sec

TOTAL ITERATION	OUTER ITERATION	INNER ITERATION	MAXIMUM CHANGE	MAXIMUM CHANGE MODEL-(CELLID)	MAXIMUM RESIDUAL	MAXIMUM RESIDUAL MODEL-(CELLID)
1	1	1	32.150725	1_GWF-(1,10,10)	225.98899	1_GWF-(1,1,2)
2	1	2	131.67014	1_GWF-(3,9,10)	352.20834	1_GWF-(2,1,2)
3	1	3	87.662587	1_GWF-(3,8,8)	-109.33418	1_GWF-(1,3,9)
4	1	4	47.376840	1_GWF-(3,9,10)	148.72089	1_GWF-(1,1,2)
5	1	5	25.415094	1_GWF-(3,9,9)	-86.286591	1_GWF-(2,9,10)
6	1	6	14.831939	1_GWF-(3,9,10)	-36.564461	1_GWF-(1,10,10)
7	1	7	3.0448374	1_GWF-(1,10,10)	12.242839	1_GWF-(1,1,2)
8	1	8	1.0823346	1_GWF-(3,1,10)	-2.3945604	1_GWF-(1,1,7)
9	1	9	0.54568407	1_GWF-(3,1,6)	-1.9622534	1_GWF-(1,2,8)
10	1	10	0.41578866	1_GWF-(3,1,7)	1.2910734	1_GWF-(1,1,2)
11	1	11	-0.21706022	1_GWF-(3,10,9)	-0.91758730	1_GWF-(1,1,2)
12	1	12	-0.13009522	1_GWF-(3,10,7)	-0.31297244	1_GWF-(2,1,2)
13	1	13	2.17732900E-02	1_GWF-(2,1,10)	6.23772256E-02	1_GWF-(1,1,2)
14	1	14	-6.17416021E-03	1_GWF-(3,8,9)	-2.77099584E-02	1_GWF-(1,1,2)
15	1	15	-3.06468006E-03	1_GWF-(3,9,10)	-1.97190663E-02	1_GWF-(1,2,10)
16	1	16	1.61896443E-03	1_GWF-(3,1,9)	-4.71375942E-03	1_GWF-(1,2,2)
17	1	17	5.83436018E-04	1_GWF-(3,1,10)	-7.60877869E-04	1_GWF-(1,2,7)
18	1	18	7.76817947E-05	1_GWF-(3,1,7)	2.99075582E-04	1_GWF-(1,2,2)
19	2	1	-9.4464050	1_GWF-(1,3,3)	-45.160487	1_GWF-(1,3,3)
20	2	2	-5.9119283	1_GWF-(2,3,3)	-26.347637	1_GWF-(1,4,2)
21	2	3	-5.2562030	1_GWF-(3,1,4)	19.320256	1_GWF-(1,2,4)
22	2	4	-4.6441225	1_GWF-(3,1,6)	-23.157080	1_GWF-(1,2,4)
23	2	5	-5.6079482	1_GWF-(3,1,8)	-27.753387	1_GWF-(1,2,6)
24	2	6	-5.2209287	1_GWF-(3,1,10)	-18.086070	1_GWF-(1,2,2)
25	2	7	-2.0377912	1_GWF-(3,9,10)	10.608647	1_GWF-(1,2,2)
26	2	8	-1.6221321	1_GWF-(3,9,10)	-8.3666240	1_GWF-(1,2,2)
27	2	9	-0.58105593	1_GWF-(3,9,10)	-3.0461892	1_GWF-(1,4,4)
28	2	10	-0.51019980	1_GWF-(3,10,6)	-1.9046592	1_GWF-(2,2,4)
29	2	11	-0.18528541	1_GWF-(3,10,6)	-1.0397511	1_GWF-(1,2,2)
30	2	12	-6.55472111E-02	1_GWF-(3,7,4)	0.30702195	1_GWF-(1,2,2)
31	2	13	2.06474098E-02	1_GWF-(3,1,3)	-8.51776315E-02	1_GWF-(2,1,4)
32	2	14	9.37837980E-03	1_GWF-(3,1,4)	-3.84392471E-02	1_GWF-(1,2,4)
33	2	15	2.73152026E-03	1_GWF-(3,1,9)	1.21907118E-02	1_GWF-(1,2,4)
34	2	16	9.94099460E-04	1_GWF-(3,1,10)	-3.47883692E-03	1_GWF-(1,2,10)
35	2	17	2.28091519E-04	1_GWF-(3,1,6)	8.26493766E-04	1_GWF-(1,2,10)
36	2	18	4.27446902E-05	1_GWF-(3,1,10)	-1.56111468E-04	1_GWF-(1,2,5)
37	3	1	-1.35294527E-05	1_GWF-(3,1,10)	5.85057898E-05	1_GWF-(1,2,10)

Fig 7 Head values obtained from aquifer modelling

MODELMUSE MT3DMS

MODFLOW coupled with MT3DMS is a widely used numerical modelling framework for simulating groundwater flow and contaminant transport in saturated aquifer systems. In this coupled approach, MODFLOW is first used to calculate the groundwater flow field by solving the three-dimensional groundwater flow equation, while MT3DMS uses the computed hydraulic heads and cell-by-cell flow terms from MODFLOW to simulate the movement and spreading of dissolved solutes.

Boundary conditions in MT3DMS define how dissolved contaminants enter, leave, or interact with the model domain and are essential for controlling solute mass balance in groundwater transport simulations. Unlike MODFLOW, where boundary conditions primarily govern water flow, boundary conditions in MT3DMS govern the transport of solutes associated with

those flow boundaries. MT3DMS does not independently create transport boundaries; instead, it inherits all hydraulic boundary conditions from the MODFLOW flow model and assigns solute concentrations to them.Constant-head boundaries in MODFLOW act as either solute sources or sinks in MT3DMS depending on the direction of groundwater flow. When water flows into the model domain through a constant-head boundary, MT3DMS requires a specified solute concentration for the inflowing water. This concentration represents the quality of groundwater entering the system. When flow is outward, solute mass is removed from the system in proportion to the simulated groundwater concentration at the boundary.

SELECTION OF CONTAMINANT FOR TRACING

From the water quality data collected from TWAD Board Tirunelveli, a prominent contaminant calcium is selected to trace due to its geological origin.

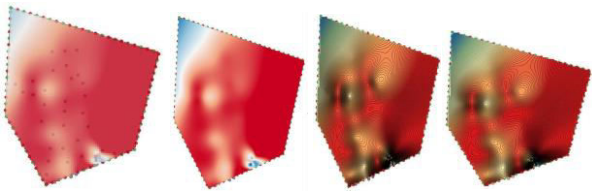


Fig 8 Calcium concentration variation for March,June of 2020,2021

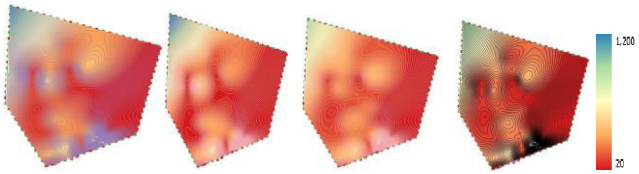


Fig 9 Calcium concentration for March June of 2022,2023

CONTAMINANT TRACING USING MT3DMS

Contamination tracing was modelled using the MODELMUSE MT3DMS with the source of well with the Calcium concentration of 700mg/lit based on the calcium concentration studied among the surrounding the wells. Due to the fast flow through the erosional cavern the water leached the limestone formations and thus the concentration of Ca in the water gets gradually increased towards the coast. The flow is towards the coast and the formation depth was also gradually increased based on considering the depositional origin of limestone. Eventually Calcium concentration peaks nearer to the coast. Total modelling has the time steps 365 for analysing for the complete one year to know the propagation of the contaminant from the

well source. The magical well was taken as the source as it transfers high volume of water through it unlike the other wells had increased the possibility of more erosion due to dumping of water through it for recharge purpose.

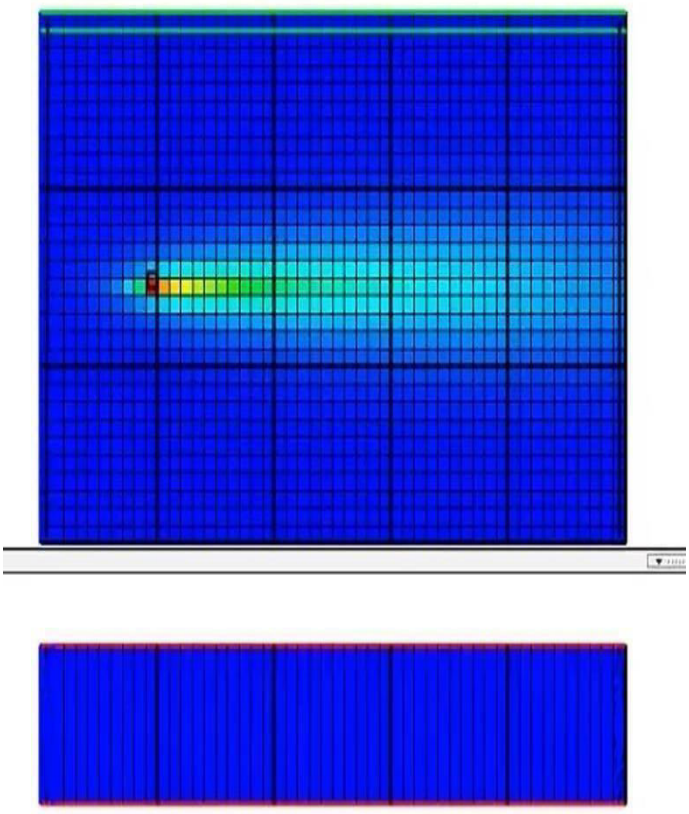


Fig 10 Contaminant tracing of Calcium as natural tracer

```
UBDSV1 SAVING     FLOW-JA-FACE ON UNIT   1011 AT TIME STEP    364, STRESS PERIOD     1
UBDSV06 SAVING       DATA-SPDIS IN MODEL MODFLOW          PACKAGE NPF           CONNECTED TO MODEL MODFLOW          PACKAGE NPF
ON UNIT   1011 AT TIME STEP    364, STRESS PERIOD     1

HEAD WILL BE SAVED ON UNIT 1012 AT END OF TIME STEP  364, STRESS PERIOD    1



      TIME SUMMARY AT END OF TIME STEP  364 IN STRESS PERIOD    1
                SECONDS     MINUTES      HOURS       DAYS        YEARS
                -----------------------------------------------------------
  TIME STEP LENGTH  1.0000    1.66667E-02 2.77778E-04 1.15741E-05 3.16881E-08
STRESS PERIOD TIME  364.00      6.0667     0.10111     4.21296E-03 1.15345E-05
      TOTAL TIME  364.00      6.0667     0.10111     4.21296E-03 1.15345E-05
```

Fig 11 Result of contaminant tracing and its concentration variation obtained after 1 year.

RESULTS AND INTERPRETATION

The Geophysical analysis suggest the limestone may be originated by the past geological depositional environment of shallow coast and later the coastline might be degraded and

this was proved by the shallow oceanic inland sediments and formations along the coast of Tirunelveli district. The Achenkoil shear zone act as the tectonic boundary of the study region to arrest the groundwater flow towards the southern region towards Koodankulam area. The water level contour map and the calcium tracer concentration maps using QGIS suggest the sharp tectonic boundary divides the study region by the sudden drops an abrupt calcium concentration variations. The groundwater flow direction was predicted along the shear zone and ended up towards Uvari coast and discharged thousands of m^3 of fresh water directly to the sea during the rainy season when the excess runoff water diverted to the magical well and the interconnected wells due to the erosional veins of limestone formations gets recharged but most of the fresh water quantity is wasted. So possible alternative method should be required to store the excess water for the future use such as constructing massive tanks or storage facilities. Contaminant tracing suggest the calcium concentration will get increases if the continuous dissolution is allowed by supplying more runoff water for recharge purpose and possibly will make the terrain unstable for construction purpose. This discharging of water towards magical well poses threads when its dimension gets increases towards the coast as limestone formation thickened. This caverns probably looks shallow origin and possibly deeper aquifers may gets recharged based on the aquifer modelling results.

REFERENCES

[1] Ford D, Williams P "Karst Hydrogeology and Geomorphology".Chichester, England,2007 Wiley. P: 562.

[2] Goldscheider N, Drew D Methods in Karst Hydrogeology. London Taylor & Francis. International Contributions to Hydrogeology 2007 26: 264.

[3] Goldscheider and David, 7 Goldscheider, D. David (Eds.), Methods in Karst Hydrogeology. 2007 12

[4] White WB Geomorphology and Hydrology of Karst Terrains: New York. Oxford University Press,1988 P: 464.

[5] Belay Zerga.,Karst topography: Formation, processes, characteristics, landforms, degradation and restoration: A systematic review Watershed eco and environment volume 2024

[6] Bakalowicz M Karst groundwater: a challenge for new resources. Hydrogeology Journal.2005. 13: 148-160.

[7] Ford D, Williams P Karst Hydrogeology and Geomorphology.Chichester, England, Wiley.2007 P: 562.

[8] Bailly-Comte, V., Jourde, H., Roesch, A., Pistre, S., & Batiot-Guilhe, C. Time series analyses for karst–aquifer characterization and modeling: A review. Journal of Hydrology,2012. 425– 426, 1–13

[9] Bashi-Azghadi, S.N.,Kerachian.,Bazargan-Lari,M.R.,and Solouki,K. Characterizing an unknown pollution source in groundwater resources systems using PSVM and PNN.Expert syst application 37(10)2010.,7154-7161.

[10] Mangold, D. C., and Tsang, C. F. A summary of subsurface hydrological and hydrochemical models. 1991 *Rev. Geophys.* 29 (1), 51–79. doi:10.1029/90rg01715

[11] Lapworth,Ward.R 'Emerging Organic contaminants in groundwater:A review of sources,fate and occurances Environ.pollution.163 Apr,287-303

[12] Mita Rajaram,S.P.Anand "Crustal structure of south India by Aeromagnetic data"Journal of Virtual explorer 2003 12:72-82

[13] Mathiazhagan Mookiah[1*], Srinivas Yasala[2], Saravanan Sakthivel[3], Chandrasekar Nainarpandian[4],Ramji Vaidhyanathan[5] and Swaster Robert[6]"Identification of Artificial Groundwater Recharge System near Ayyankulam Village of Tirunelveli, Tamilnadu, India Using Electrical and Magnetic Survey" DOI:10.24966/AES-8780/100042

[14] F.Pradeepa,Jerin.P.Shaji,Raj Kumar,Dr.Y.Srinivas "Delineation of aquifer using electrical geophysical method in Aayankulam-Radhapuram which shows exceptional drainage and flow pattern" International Journal of basic and applied research 2018 vol 8 pg:1228-1236

TRACK-2

COMPUTER SCIENCE & ENGINEERING

"PreviDict: An AI-Based Framework for Pre-Upload View Prediction and Content Optimization of YouTube Shorts"

JAVHARA SHERIN V
Department of Information Technology
Institute of Engineering and Technology
University of Calicut
Malappuram, Kerala, India
javharasherin@gmail.com

ANUPAMA K T
Department of Information Technology
Institute of Engineering and Technology
University of Calicut
Malappuram, Kerala, India
ktanupama2005@gmail.com

NAVARANG S MANOJ
Department of Information Technology
Institute of Engineering and Technology
University of Calicut
Malappuram, Kerala, India
navarangsm@gmail.com

YASH KRISHNA K P
Department of Information Technology
Institute of Engineering and Technology
University of Calicut
Malappuram, Kerala, India
yashvalsaraj@gmail.com

ASWANI N M
Department of Computer Science
Institute of Engineering and Technology
University of Calicut
Malappuram, Kerala, India
aswaninm@gmail.com

ARYA A SURYA
Department of Computer Science
Institute of Engineering and Technology
University of Calicut
Malappuram, Kerala, India
arayasurya92@gmail.com

***Abstract*—The rapid proliferation of short-form video platforms has created an ultra-competitive environment, where creators find it extremely difficult to assess the potential performance of videos until they are actually published. Most of the currently available tools only provide analytics about how a video performed after the content is published, thus providing little or no support to the creator at the time of content creation. In this paper, we present a system named PreviDict, which is based on artificial intelligence (AI) technologies and will be able to predict the expected view count for a post on a short-form video platform before the content is published. The proposed technique will use machine learning to analyze the metadata of the video (i.e., title, description, hashtags, thumbnail attributes) to give an estimate of the likely been popularity of the video after it is actually posted. In addition to providing a prediction of expected number of views for the video, PreviDict will generate content optimization suggestions that would help to maximize the video's reach and engagement. In addition to the above functionality, a societal impact analysis module is included that will allow PreviDict to categorize the class of each video based on its potential impact, which will encourage responsible video production. Experimental tests of PreviDict have shown that the system provides reliable prediction accuracy and actionable insights, which will allow video creators to refine their video prior to posting it, as well as reduce uncertainty about the performance of their short-form video prior to posting them.**

***Keywords*—Short-form videos, machine learning, view prediction, content optimization, societal impact.**

I. Introduction

Video Platforms like TikTok, Instagram Reels and YouTube Shorts are changing the way we consume and create content online. These services have allowed content creators a way to reach a large number of people in a relatively short time due in part to the reach these platforms provide and their algorithmically determined recommendations for videos. Because of these factors and the competitive environment found on these platforms, video creators find it difficult to consistently get a reasonable amount of visibility and engagement.

One of the problems video creators face is they are unable to predict how well their desired video will perform prior to its release. Other than intuition or analytical data once a video is released, creators cannot accurately estimate how well their video will perform based on various factors including title, thumbnail design, hashtags and description, which all fundamentally affect an audience's reach. Most of the tools that exist today can only report metric data after a video has been posted, specifically focusing on views, watch time and engagement; therefore these tools have limited functionality in the content creation process and as a result, video creators often end up relying on trial-and-error when creating content resulting in wasted time and/or resources with little to no consistency in results.

Recent research has reviewed video popularity predictions through the use of machine learning and deep learning methods utilizing video performance data and other post-release metrics based on this data. Even though these approaches are informative, they are primarily used.

II. Related Works

The application of artificial intelligence to forecast digital content's performance is gaining attention in academic literature, mostly in the context of social networks and other digital platforms. Researchers have begun to evaluate the

potential of using machine learning algorithms that analyze data about uploaded videos (metadata), visual aspects (visual content features), and user behaviors (user interaction patterns) in order to predict the likelihood that a given video will be popular.

Short video platforms like YouTube Shorts, TikTok and Instagram Reels have rapidly gained in popularity, and predicting the reach of uploaded videos is an emerging area of research.

Studies in this area have historically focused on statistical and regression-based models that use engagement metrics such as views, likes, comments and watch time (post-upload metrics) to estimate video popularity [1], [2]. Results from early studies were able to use engagement features to accurately predict video popularity over time. A major limitation of using post-upload metrics is that they do not help provide video creators with information on whether or not their video will perform well prior to uploading it.

Recently, researchers have applied new deep learning techniques such as convolutional neural networks and recurrent neural networks to analyze multi-modal feature sets (the relationship between textual data related to the video (metadata), visual data related to the video's thumbnail (visual feature), and the temporal relationship of the video (temporal features)) [3], [4]. Predictive accuracy using deep learning models is superior to that of traditional regression-based models. However, these systems have all been developed to help optimize the use of video recommendation algorithms and for predicting post-upload performance data.

In addition, there are a new breed of content optimization tools/models available that help video creators with the development of video titles, hashtags and descriptions.

III. System Architecture

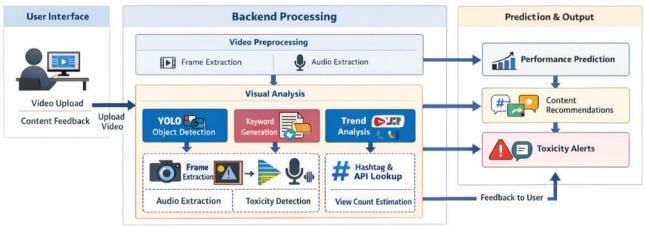


Fig. 1. System Architecture of the Proposed PreviDict Framework

A modular web-based approach to integrated task management, behavioral insights and tracking, collaborative workspaces, visual analytics and artificial intelligence-powered assistance will provide one unified solution for users. The proposed system will have a client/server architecture, ensuring that the system can grow easily, be flexible, and be able to handle large amounts of data processed efficiently.

A web-based (standard web browsers) front-end interface will allow users to perform tasks such as entering tasks, working with others on completing tasks, participating in focus sessions, and viewing insights about their productivity using interactive dashboards. All user actions such as creating or updating tasks and collaborating on tasks will be stored and sent to the back-end application in real-time.

The back-end application serves as the core of the system, providing the application's logic and coordination of overall system components. The back-end application provides authentication, task scheduling, management of focus sessions, and coordination of community users. The back-end application also manages multiple applications supplying data to the back-end for further aggregation, and communications between the multiple applications of the system.

To enable users to achieve productive collaboration, a persistent connection between users will provide them with live access to focus rooms and will synchronize the collaboration between users, ensuring the same outcome when multiple users are participating in a particular task simultaneously.

The system uses one central database that will store a database of all structured data (user profiles, tasks, session duration, timestamps of user actions and all productivity metrics). The integrated task management solution's data will provide clear evidence of how effective and productive all users are with their respective tasks.

IV. Proposed Methodology

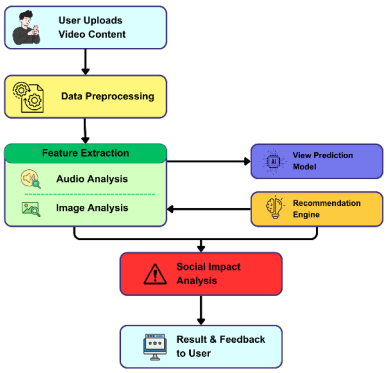


Fig. 2. Workflow of the proposed PreviDict methodology

The PreviDict's method for forecasting how successful short form videos can become before posting is defined in detail within this document. Using a variety of methods including; machine learning and content analysis as well as other recommendation engines, creators can receive assistance evaluating their content at the point of time when it will be uploaded.

To begin with, either the mobile or web interface is used by a user. Once they have successfully registered and logged into their account, users can then upload their short form video

and provide any metadata needed to describe their submission such as a title, description, hashtags, and thumbnail image.

Once submitted, this data is validated and pre-processed to ensure that each data point is reliable for analysis so as to not affect the end result.

During the extraction of features, both textual/visual attributes will be extracted from the input data. Textual features will include those taken from the title, description, and hashtags through natural language processing techniques to provide a list of keyword correlations between the short form video submission, and the prior correlation between other short form video submissions. The second category of features will consist of those related to a video's thumbnail, which have been identified as providing major influences for securing viewer engagement. These would include, but not be limited to; how color contrasts in an image, how clear the image is, the usefulness of any text displayed, and the presence of any faces or individuals in the image.

At this time, all feature will be passed to the machine learning model associated with the system used for making predictions on how many views users can expect their short form video will gain, based upon using historical data from previously uploaded short form video data. Once the model has established visual / amount of views relationship, the amount of views that each new submission to the PreviDict system will gain can also be established with a measure of accuracy.

V. Technologies Used for Feature Extraction and Content Analysis

The PreviDict process uses a combination of deep learning, natural language processing, and analytics from external platforms in order to anticipate the performance of short-form video content before publication. All the streams' pipes are structured in an automated way to analyze the visual, textual, and audio content of a given video after it is uploaded..

A. Video Frame Extraction Using Transformers

The video is uploaded and the representative pictures are extracted using transformer-based video processing methods for frame extraction and the selection of key moments throughout the entire timeline of the length of the video by providing a higher degree of understanding of how to obtain temporal relationships in video without manual selection of frames.

B. Sensitive Content Detection Using YOLO

The extracted pictures in the video are processed through the YOLO model (You Only Look Once, also known as real-time object detection) to identify sensitive content or restricted visuals (non-annotated) automatically during the framing process. In addition, the processing required for the YOLO detection process will be automatically created in order to ensure that the input is always recognized consistently by the model for quicker detection. After the detection of sensitive or restricted object and scene locations, the model generates relevant visual keywords based on those locations so that each video can be described according to identified video characteristics.

C. Keyword Generation and Hashtag View Analysis

The YOLO generated motion detection will yield keywords that will generate relevant hashtags. After the hashtags are created, the views collected for those hashtags on the YouTube API will be used to get history of views for similar or trending hashtags. By monitoring the viewcounts of the marked hashtags, we can predict the reach of the video uploaded. An average view count will be calculated from the collected views of all marked hashtags.

D. Audio Toxicity Detection

As well as being evaluated visually, the audio of the video uploaded will also undergo analysis to ascertain if any harmful or obscene words are present. The words identified in the audio will undergo a speech to text conversion which will then be analysed using audio toxicity detection methods to identify any words are present that could be considered offensive or could restrict the ability for the content produced to be viewed due to being harmful content or not compliant with the platform.

E. AI-Based Content Improvement Suggestions

From the observation of the visual aspects of the video for sensitivity, performance for hashtags, and toxicity of the audio we are able to use all this data to provide suggestions to the creator on how to improve their video or content. Using a separate API that uses artificial intelligence, we are able to provide actionable improvements such as modifying the hashtags being used, enhancement the visual quality of the content, or improving the manner in which the audio can be understood. These suggested improvements are intended to increase the engagement opportunities for the creator and ensure the content being created meets the requirements of the platform prior to posting.

VI. Implementation Details

The PreviDict system employs a modular architecture that integrates front-end interfaces, back-end services, deep learning models, and components for data processing. The system was designed for scalability and efficiency allowing for the automated analysis of uploaded short-form videos and making instantaneous predictions regarding content performance.

A. System Development Environment

The back-end of the PreviDict System has been created using Python due to its extensive support of machine learning and data processing functions. The back-end services support video uploads; coordinate the inference of the models; manage interactions with the APIs; and aggregate results of predictions that are made about the uploads. The front-end of the system has been developed using standard web technologies, which allow the user to upload a video, view predictions made for their videos, and receive feedback on those predictions through a responsive design for use on multiple devices.

B. Video Data Handling and Preprocessing

As one uploads short-form video content into the PreviDict system, the system then extracts both audio and video streams from the uploaded video for further possible analysis. Additionally, the uploaded video's video frames are sampled regularly throughout the duration of the uploaded video for reducing computational overhead by recording only those video frames that could provide relevant information. The preprocessing of both the audio and video being employed by the assigned deep learning models takes care of preprocessing. Additionally, the audio data is separated from the video stream so that both could independently be analyzed with respect to speech.

C. Feature Extraction and Representation

Extracting visual characteristics is achieved by using a YOLO (You Only Look Once)-based detection system that detects certain objects and visual elements within a video frame that may represent sensitive content or visual indications of restricted/substantive (or adult-rated) content. Descriptive keywords are created based on the visual characteristics found to help represent and/or describe the visual context of a video.

Temporal (frame) video information is processed by using transformer architectures to allow for consistent frame (temporal) processing and to assist with the efficient use of temporal-based video processing.

The keywords that were created by the visual analysis will also be used to assess the video in relation to its trending, referrals, and overall content relevance.

Also, the audio of the video will be processed to extract spoken language. The audio will be converted from speech to written language. The written language will then be analyzed for negative or harmful speech to determine whether the video contains negative/harmful content that might prohibit the video from reaching its full potential with an audience.

D. Trend Analysis and View Estimation

The keywords formed through the process of visual analysis will be attached/mapped to the related hashtags that correspond with the trending content of the derived keywords. The view counts for videos related to the hashtags created through visual analysis will then be retrieved using the YouTube Data API. The mean view count of the collected (retrieved) view counts will be used to estimate the engagement that the uploaded video might receive from the general audience.

E. Prediction and Output Generation

A single representation is created by combining the visual sensitivity detection options with the trend analysis by keywords, average views by hashtag, and audio toxicity detection into one predictive template of the uploaded video. This predictive template is then used to create predictive intelligence of the likely success of the video, and the output produced will be formatted for delivery to the user.

F. User Interface and Feedback Loop

The user interface will show predicted engagement metrics, along with content-related insights, in a clear and easily interpretable format. Creators will be able to change their content according to the feedback provided, and then re-submit their content for assessment. This iterative process allows for creators to continue to improve their content and will enable the system to continue to help improve the quality and performance of all content submitted.

VII. Results and Performance Analysis

The User Interface has been designed to allow users to see the projected performance metrics as well as provide suggestions for optimisation in an easy to read and use display format. Users will be able to use the feedback loop to make changes to their content then reanalyze their content again using the same method to allow for iterative improvement. The feedback loop will further increase the system's ability to assist video creators.

A. Experimental Setup

Our testing dataset has been separated into both a training dataset and a testing dataset for purposes of identifying bias while testing our model. The training dataset will allow us to identify how different features that describe our content correlate with the number of views (on a video) while the testing set will only assist us in determining how well our model performs once we have built our model. All experiments will take place under the same conditions for both datasets in order to achieve unbiased results.

B. Prediction Accuracy

We will evaluate the accuracy of predictions made by PreviDict Model by performing various regression analysis metrics such as Mean Absolute Error (MAE) as well as the percentage of those predictions that fall within an acceptable margin of error. Based on the findings of the analysis; we have concluded that PreviDict is able to provide an accurate prediction of potential views for each of the inputted features to their corresponding media items prior to upload.

C. Comparative Analysis

In order to further confirm the validity and relevance of the various machine learning algorithms used within our PreviDict Model; we will contrast our predictions of potential views with other baseline machine learning models which utilize a lesser number of features in making their predictions. As such, it is expected that the estimates produced by the PreviDict model will produce a significantly greater level of accuracy than estimates produced from other machine learning models because we incorporate text and visual feature analysis into our model framework. Collectively, the results obtained from the comparative performance evaluation demonstrate that utilising multiple forms of data extraction is crucial for providing accurate estimates of success for short-form video content.

D. Recommendation Effectiveness

The performance of the recommendation module is evaluated through a qualitative analysis. System-generated recommendations improve the content using optimised content; both promoted keywords received more favourable placement against trending keywords, clearer thumbnails, and more relevant hashtags. The use of system-generated recommendations enhances the visibility and engagement that can be established with the content.

E. Societal Impact Evaluation

The societal impact analysis module provides an accurate evaluation of all content based on the sentiment and keyword evaluation of three categories: positive, neutral and negative. This evaluation provides an opportunity to identify potentially harmful or misleading content during the earlier stages of production to encourage the responsible creation of content.

F. Discussion of Results

The experimental results demonstrate that PreviDict produces accurate performance predictions, providing support for creators to optimise their content prior to publishing, thereby reducing uncertainty in their decision-making process. It is recognised that the minor prediction difference could be attributed to the rapidly evolving nature of the platform that shows an ongoing need to continually update the prediction model.

VIII. Conclusion

PreviDict is an intelligent and automated system for predicting the way short-form videos will perform before they are published. It uses computer vision, natural language processing, audio analysis and trend-based engagement estimates to help creators understand how large an audience their videos may reach and if the content is suitable for the audience.

The model uses YOLO-based visual detection to identify sensitive or impactful objects in video frames, transformer-based processing to create contextual keywords and audio toxicity detection to evaluate spoken content. The results of each video's feature extraction are then combined with real-time hashtag trend data from the YouTube Data API to produce an estimated average view count. As such, multimodal methods will not rely on one type of feature to make predictions; rather, they'll rely on a complete analysis of the video's visual, textual, and audio components.

Our experimental evaluation shows that by combining visual sensitivity detection, keyword-driven trending analysis, and toxicity assessment, the reliability of estimating performance is much greater than using only basic predictive models. In addition, this model provides a means to encourage safe and responsible content creation by identifying potentially harmful or inappropriate elements in advance of publishing.

Ultimately, PreviDict helps provide more certainty for the creator in determining how to publish and what type of content will be most successful and have the greatest social impact. Future enhancements could include continually retraining the model.

Acknowledgment

We appreciate all the support from the Department of Information Technology, Institute of Engineering and Technology, University of Calicut for their contributions to this project. Specifically, we appreciate their support, guidance, and encouragement during the course of this project.

References

[1] T. Trzciński and P. Rokita, "Predicting popularity of online videos using support vector regression," *IEEE Transactions on Multimedia*, 2017.

[2] N. Razali et al., "A regression approach for prediction of YouTube views," *Journal of Engineering Science*, 2019.

[3] V. Bhatnagar and N. Sangwan, "Multi-branch CNN-LSTM model for YouTube popularity prediction," *IEEE Access*, 2025.

[4] D. Li et al., "Short video popularity prediction using deep learning," *IEEE Internet of Things Journal*, 2024.

[5] A. Agarwal et al., "SEO-based optimization tools for social media content," *International Journal of Digital Media*, 2023.

Integrating Human Activity Recognition and Affective Computing for Elderly Well-Being: A Comprehensive Review

Dona Maria Mani,
Assistant Professor
Department of Computer Science
Rajagiri College of Management and Applied Sciences, Kakkanad
donamarianov25@gmail.com

***Abstract*—This review explores the integration of Human Activity Recognition (HAR) and affective computing to improve elderly care. Affective computing evaluates emotional states using behavioral and physiological cues, while HAR systems use sensors and machine learning to track behaviors such as walking and sleeping. When these technologies are combined, both mental and physical safety are ensured. This review examines the state-of-the-art HAR techniques, including sensor-based and vision-based approaches, and advances in deep learning for emotion recognition. Multimodal data fusion strategies are discussed to enhance robustness and accuracy. Key applications such as fall detection, stress monitoring, and personalized interventions are highlighted. The paper also addresses critical challenges, including data privacy, ethical considerations, and the need for interpretable models suitable for real-world deployment.**

Keywords—HAR, affective computing, Machine learning(ML), deep learning(DL), facial expression recognition(FER), convolutional neural networks (CNN), recurrent neural networks (RNN), mixed reality (MR)

I. INTRODUCTION

The world's aging population needs creative ways to protect senior citizens' health and wellbeing. Technologies like affective computing and human activity recognition (HAR) have enormous potential for tracking the mental and physical well-being of senior citizens.

Human Activity Recognition (HAR) uses a variety of sensors, such as cameras, gyroscopes, and accelerometers, to track the physical activity of the elderly. HAR helps us comprehend the everyday habits of older people by precisely recognizing actions including walking, sitting, and sleeping.

The Fig. 1. shows how daily activity data is turned into useful health insights.

A person wears sensors that collect raw data like movement and posture. This data is cleaned and organized by removing noise, filtering useful signals, normalizing values, and breaking it into meaningful segments while extracting key features.

Machine learning models then analyze the processed data to recognize patterns such as walking or sitting.

Finally, the system outputs simple insights like activity level, number of activities, duration, and intensity of behavior.

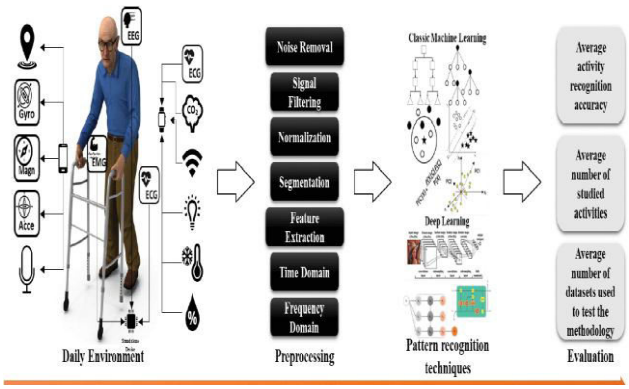


Fig. 1.

In contrast, affective computing uses physiological (heart rate, galvanic skin response) and behavioral (facial expressions, speech) clues to identify emotions. It enables healthcare systems to evaluate emotional states that are frequently problematic in the treatment of the elderly, such as stress, worry, and depression.

This paper explores how HAR and affective computing might be combined to enhance care for the elderly. The goal is to present a thorough analysis of current methods, important applications, difficulties, and potential future paths in this nascent field of study.

.

II. EXISTING APPROACHES IN HAR AND AFFECTIVE COMPUTING

A. Human Activity Recognition

There are two primary categories under which HAR falls: Sensor-based Approaches: These use wearable technology, including wristbands, accelerometers, and smartwatches, to track an elderly person's posture and movement. To identify and categorize actions, algorithms are used to process the data from these sensors. As an illustration, consider the use of accelerometers to identify standing, sitting, and walking.

Vision-based Approaches: These employ cameras or depth sensors (like Kinect) to keep an eye on senior citizens without making physical contact. To identify activities, deep learning models (CNNs, RNNs) and sophisticated computer vision techniques are used. For instance, employing a video system and body position analysis to identify falls.

B. Affective Computing

To identify emotions, affective computing uses a variety of data sources:
Facial Expression Recognition: This method uses computer vision algorithms to recognize facial expressions in pictures or videos. For instance, identifying facial expressions of happiness, sadness, or anger.

Speech Emotion Recognition: Identifying emotions by analyzing speech characteristics like pitch, tone, and speech rate.
For instance, using speech patterns to determine whether someone is stressed or anxious.
Physiological Emotion Detection: This technique uses sensors to track changes in body temperature, skin conductance, and heart rate that may be signs of emotions. Analyzing heart rate variability (HRV) to monitor stress is one example.

C. Multimodal Data Fusion

The accuracy and resilience of both HAR and emotion recognition systems are improved by combining data from many sources, such as wearable sensors, cameras, and physiological sensors. In dynamic, real-world settings, data fusion techniques increase the accuracy of activity and emotion detection.
Multimodal data fusion integrates these heterogeneous data sources using different fusion techniques:

Feature-Level Fusion: Raw sensor data from different modalities are combined before feeding into a machine learning model, enhancing early-stage decision-making.

Decision-Level Fusion: Each modality processes data independently, and their outputs are merged for the final classification, improving robustness.

Hybrid Fusion: A combination of feature- and decision-level fusion improves accuracy and adaptability to real-world scenarios.

III.KEY APPLICATIONS

A. Identifying Falls

One of the biggest hazards for senior citizens is falls. Affective computing can evaluate the emotional anguish brought on by the fall, while HAR systems can identify falls by analyzing movement. Both devices work together to create a complete solution that monitors emotional reactions after a fall and notifies caregivers.

B. Tracking Stress and Anxiety

Elderly people frequently experience stress and worry, particularly if they have long-term medical concerns. While HAR assists in identifying changes in physical behavior that may be associated with emotional distress (such as restlessness or sleep disturbance), affective computing enables ongoing emotion monitoring.

C. Personalized Interventions

Personalized therapies can be offered by healthcare systems by integrating activity and emotional data. For instance, the system might recommend physical activities or relaxation techniques to an elderly person who exhibits symptoms of anxiety and is sedentary in order to enhance their mental and physical health.

IV. CHALLENGES

A. Data Privacy and Security

Ensuring data privacy and security is crucial when sensitive data is being collected, such as emotional states or bodily actions. Concerns regarding the storage and sharing of this data are brought up by the integration of cameras, wearable technology, and other sensors. Regulations such as GDPR (General Data Protection Regulation) must be followed.

B. Ethical Considerations

Consent: It is needed to ensure that senior citizens, particularly those with cognitive impairments, give their informed consent before any data is collected. Emotional Manipulation: To prevent inadvertent emotional manipulation, systems should be built with ethical considerations in mind.

C. Interpretability of Models

Machine learning models, especially deep learning approaches, are often considered as "black boxes." For these systems to be widely adopted in real-world applications,

models need to be interpretable. Caregivers should understand how decisions are made by the system, especially when it comes to emotional assessments.

D. Real-World Deployment

HAR and affective computing models are often trained in controlled environments and may not perform as accurately in dynamic real-world settings. Factors like sensor placement, lighting conditions, and ambient noise can affect the accuracy of these systems.

V. PREVIOUS WORKS

A. Human Activity Recognition (Har) In Elderly Care

Human Activity Recognition systems have been widely studied and implemented in the context of elderly care, with a focus on monitoring physical activities to assess health risks such as falls, mobility limitations, and sleep disorders.

The necessity for cutting-edge technologies to support and monitor independent living has increased due to the growing number of older people. Fall detection and human activity recognition (HAR) are two of these technologies that are now essential for ambient assisted living (AAL) systems. [1] looks at 151 papers conducted between 2019 and 2023 with an emphasis on deep learning (DL) methods for fall detection and HAR utilizing computer vision data. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and graph convolutional networks (GCNs) are among the deep learning designs that are categorized in the review, with an emphasis on how well they extract features and classify activities. RGB video, depth pictures, skeleton tracking, and infrared imaging are the most frequently examined data modalities, with RGB being the most prevalent source.

Machine learning applications in HAR are examined in "Human Activity Recognition Using Inertial, Physiological, and Environmental Sensors: A Comprehensive Survey" by Demrozi et al.[2], which focuses on data from inertial, physiological, and environmental sensors. The integration of several sensor types to improve the accuracy of activity recognition is highlighted in the authors' assessment of experiments conducted between January 2015 and September 2019. They address issues including data heterogeneity and computational needs while discussing the function of both traditional machine learning and deep learning approaches in processing sensor data. The report also looks at everyday living and healthcare applications, highlighting how multimodal sensor fusion can be used to monitor and help the elderly.

The difficulties of tracking everyday activities in elderly and ill populations to promote independent living are discussed in [3] by Lisa Schrader et al. The authors suggest a technique that uses ambient and wearable sensors to collect data in realistic hospital environments. With an emphasis on clinically relevant activities including resting positions, diet, and hygiene, this strategy is adapted to the physical capacities of older persons. In order to improve machine learning models for human activity recognition (HAR), the study highlights the significance of developing structured datasets from this population. The study intends to identify activity patterns that can guide early treatments and rehabilitation by putting in place an activity recognition chain (ARC).

[4] focuses on helping senior citizens by tracking their movements in various indoor and outdoor settings using accelerometer and gyroscope data gathered from a smartphone. Human activity is recognized using both traditional and deep learning methods, including Random Forest, Support Vector Machine, Artificial Neural Network, k-Nearest Neighbors, and Long Short-Term Memory Network. The suggested Long Short-Term Memory Network achieved the highest accuracy of 95.04% out of all the classification methods. However, utilizing 10-fold cross-validation, Support Vector Machine provided an accuracy of 89.07% in a remarkably short calculation time of 0.42 minutes.

A thorough examination of deep learning approaches used for human activity recognition (HAR) with 3D skeleton data can be found in [5] by Nguyen et al. Recurrent neural networks (RNNs), convolutional neural networks (CNNs), graph convolutional networks (GCNs), and hybrid deep neural networks (Hybrid-DNNs) are the four main deep learning architectures into which the authors divide the present methods. They assess these models using datasets like KLYOGA3D and KLHA3D 102, providing information on their usefulness and performance. Along with addressing issues with data variability and model complexity, the paper makes recommendations for future research directions that will concentrate on improving model resilience and the viability of real-time applications. This survey is a valuable resource for researchers who are interested in the development of advanced HAR systems that utilize 3D skeletal data.

The usefulness of ear-worn motion sensors for categorizing everyday activities in natural environments is investigated in the 2024 study "Human Activity Recognition in a Free-Living Environment Using an Ear-Worn Motion Sensor" by Boborzi et al. [6] .50 healthy participants, ages 20 to 47, participated in the study and engaged in activities like jogging, walking, stair climbing, sitting and standing, and lying. The sensor data was analyzed using a variety of machine learning techniques, such as ConvTransformer and

DeepConvLSTM. The results showed that 98% classification accuracy was attained by deep learning models, underscoring the promise of ear-worn sensors for ongoing behavior monitoring in medical applications. This technique provides a non-intrusive way to identify activities in real time, which is very useful for keeping an eye on people and helping them with their everyday tasks.

B. Affective Computing In Elderly Care

Affective computing, the study of systems capable of recognizing and responding to human emotions, is increasingly used in healthcare to monitor the emotional states of elderly individuals. Understanding an elderly person's emotional well-being is crucial for mental health monitoring and the early detection of conditions such as depression and anxiety.

The multidisciplinary field of affective computing is concerned with the identification of emotions and how they are used in a variety of contexts, such as virtual reality, healthcare, and human-computer interaction. [7] examines cutting-edge approaches in affective computing, focusing on mixed reality (MR), deep learning (DL), and machine learning (ML). Textual, audio, visual, and physiological signal analysis are all techniques for recognizing emotions, and each has its own advantages and disadvantages. While DL models use CNNs and RNNs for improved feature extraction, ML-based methods rely on feature engineering and classifiers like SVM and ANN. Recent advancements in MR demonstrate its capacity to provide immersive emotive experiences, supporting uses like cognitive training and mental health monitoring. Nevertheless, there are still issues with model interpretability, inconsistent emotion categorization, and dataset restrictions. Future research must focus on multimodal data fusion, adaptive learning models, and ethical considerations in emotion-based AI systems. The integration of VR/AR technologies with affective computing presents promising directions, enhancing real-time emotion detection and interaction. Addressing these challenges will significantly contribute to the development of intelligent systems capable of understanding and responding to human emotions effectively.

According to [8], affective computing improves healthcare by making it possible to identify emotions using deep learning models, which helps with mental health monitoring, diagnosing depression, and providing individualized treatments. In order to increase accuracy, modern methods integrate behavior and physiological variables using convolutional neural networks (CNN), recurrent neural networks (RNN), and multimodal data fusion. Elderly care, pain evaluation, and autism identification are important applications. Real-time diagnosis, emotion bias, and data privacy are among the difficulties. In order to improve therapeutic applications, future research should concentrate on adaptive learning, ethical AI models, and integrating big foundation models. Innovation in emotion-aware intelligent healthcare systems will be fueled by tackling these issues.

Fall detection and human activity recognition (HAR) have become essential for ambient assisted living (AAL) systems due to the growing number of older people. [9] examines computer vision-based deep learning (DL) methods for fall and HAR detection. It looks at 151 research and highlights important DL models for feature extraction and classification, including CNNs, RNNs, and GCNs. Although RGB data is still widely used, privacy-preserving substitutes such as infrared and depth photography are becoming more popular. Model interpretability, real-world deployment, and dataset constraints are among the difficulties. Multimodal data fusion, ethical issues, and reliable real-time applications for senior care should be the main topics of future study.

C. Integration Of Human Activity Recognition And Affective Computing

The integration of Human Activity Recognition (HAR) and affective computing has risen as a promising approach to improving elderly care. The combination of activity monitoring and emotion detection allows healthcare systems to provide a comprehensive assessment of an elderly individual’s well-being

Recent research has begun to integrate HAR and affective computing to create comprehensive systems for elderly care that monitor both physical and emotional well-being.

The incorporation of deep learning methods for facial expression recognition (FER) in senior care is methodically examined in the publication "Deep Learning-Based Facial Expression Recognition for the Elderly: A Systematic Review" by F. Xavier Gaya-Moreya et al[10]. The review emphasizes how crucial it is to read facial expressions in order to track emotional wellness and identify health problems in the elderly. The authors concentrate on a number of deep learning models, including recurrent neural networks (RNNs) and convolutional neural networks (CNNs), which have demonstrated promise in precisely recognizing elderly people's facial expressions. It highlights issues like aging, which can alter facial characteristics and the dynamics of facial expressions. The study also addresses datasets, model performance, and FER's use in healthcare systems with the goal of enhancing senior care via non-invasive monitoring.

A thorough analysis of emotion detection methods utilizing physiological signals obtained from wearable technology

can be found in [11] by Stanisław Saganowski et al. The integration of sensors like accelerometers, heart rate monitors, and electrodermal activity sensors to identify emotional states in real-world situations is the main topic of the review. It examines several machine learning techniques used to evaluate physiological data and precisely categorize emotions including happiness, sorrow, tension, and rage. The authors talk about issues that impact the accuracy of emotion detection, such as individual variances, sensor noise, and the variety of emotional reactions. The study also emphasizes how emotion detection may be used in human-computer interaction, tailored healthcare, and mental health monitoring. The review concludes with recommendations for future research to improve sensor accuracy, algorithm robustness, and the overall effectiveness of wearable systems for emotion recognition in everyday settings.

Marwan Dhuheir et al.'s study [12] offers a thorough summary of current developments in neural network-based emotion identification, especially in the context of healthcare surveillance. Speech, facial expressions, and audio-visual inputs are the three main modalities for emotion recognition that the authors concentrate on. They go into detail on the methods used in each modality, going into the pre-processing, feature extraction, and classification phases. The poll also emphasizes how these methods may be used to track patients' mental states and enable early treatments for disorders like stress and depression. The study also discusses existing issues in the field and suggests future research avenues to improve the efficiency and use of emotion recognition systems in actual healthcare environments.

[13] offers a novel method for emotion recognition (ER) that makes use of a digital twin architecture to improve healthcare personalization. By combining digital twin technology with real-time facial emotion detection, the suggested system enables ongoing monitoring and individualized medical interventions.

In addition to the high processing cost, managing occlusion, lighting fluctuations, feature recognition, and erroneous classifications is a significant difficulty in real-time ER. In order to solve these problems, the authors created a system that uses machine learning (ML) and deep learning (DL) to evaluate facial photos from a regular webcam, guaranteeing precise and effective emotion identification.

The digital twin component acts as a virtual replica of the patient, enabling real-time simulation and evaluation of emotional states. This helps healthcare providers in early disease diagnosis, mental health monitoring, and emergency response planning. The system demonstrated high accuracy, adaptability, and reduced training time, making it a feasible tool for healthcare centers.

Overall, the study introduces a cost-effective, real-time ER system that enhances patient well-being by facilitating proactive and personalized medical care, contributing to the advancement of AI-driven healthcare solutions.

These papers collectively offer a robust foundation for understanding the integration of HAR and Affective Computing in promoting elderly well-being.

VI. CONCLUSION

The integration of Human Activity Recognition (HAR) and affective computing represents a promising direction for improving the well-being of elderly individuals. By continuously monitoring both physical activity and emotional states, healthcare systems can provide timely interventions to enhance the quality of life and ensure safety. However, challenges related to privacy, ethics, and real-world deployment need to be addressed to make these systems practical and widely accepted.

VII. FUTURE DIRECTIONS

Future research in the integration of Human Activity Recognition (HAR) and affective computing should focus on developing personalized systems that can adapt over time based on long-term data. These systems should be able to learn from a person’s historical physical and emotional data, offering increasingly tailored care that responds to individual needs. Such systems would enhance the effectiveness of interventions, ensuring that elderly individuals receive the most appropriate care, whether for physical rehabilitation, emotional support, or preventive health measures. The use of long-term data would also allow these systems to detect subtle changes in behavior or emotional states that may signal health concerns before they become more severe.

Another important area for future research is improving the accuracy of emotion recognition systems, particularly for complicated emotional states. Current emotion recognition algorithms perform well for simple emotions like happiness, sadness, and anger, but they occasionally struggle to accurately detect more complicated emotions that older people often experience, such anxiety, irritation, or confusion. Enhancing these models, particularly through the use of deep learning techniques and multimodal data sources (e.g., integrating physiological signals, voice tones, and facial expressions), would allow for more substantial and accurate assessments of an elderly person's mental health

Addressing ethical concerns and ensuring data privacy is another essential area for future work. As these systems collect sensitive data related to both physical activity and emotional states, it is vital to integrate secure, transparent technologies that protect individuals' personal information. Future research must focus on developing frameworks for data governance, ensuring that elderly individuals' data is

stored, shared, and processed ethically, with respect to their privacy and consent. This is particularly important as such systems are deployed on a large scale in real-world settings, where data breaches or misuse could have serious consequences.

Lastly, developing comprehensive solutions requires increasing the interdisciplinary cooperation among engineers, medical experts, and ethicists. It takes teamwork to create integrated systems that strike a balance between ethical issues, practical healthcare needs, and technical innovation. Working together, these several disciplines will guarantee that the solutions are not only technologically sophisticated but also ethically sound, culturally aware, and tailored to the requirements of both aged care providers and users. Future studies can develop systems that genuinely improve the mental and physical health of aging populations, raising their general well-being and quality of life, by encouraging such collaboration.

As these technologies continue to evolve, they hold the potential to significantly transform elderly care. By combining advances in HAR and affective computing, future systems can offer more comprehensive, adaptive, and ethical solutions that enhance both the physical and mental health of the elderly.

REFERENCES

[1] F. Xavier Gaya-Morey, Cristina Manresa-Yee, Jose M. Buades-Rubio "Deep Learning for Computer Vision-Based Activity Recognition and Fall Detection of the Elderly: A Systematic Review", Applied Intelligence (2024) 54:8982–9007,January 22, 2024

[2]Florenc Demrozi1 [Member, Ieee], Graziano Pravadelli1 [Senior Member, Ieee], Azra Bihorac2, Parisa Rashidi3 [Senior Member,Ieee] "Human Activity Recognition using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey",,IEEE Access. 2020 ; 8: 210816–210836

[3]Lisa Schrader1 & Agustín Vargas Toro1 & Sebastian Konietzny1 & Stefan Rüping1 & Barbara Schäpers2 & Martina Steinböck2 & Carmen Krewer2 & Friedemann Müller2 & Jörg Güttler3 & Thomas Bock ,"Advanced Sensing and Human Activity Recognition in Early Intervention and Rehabilitation of Elderly People", Journal of Population Ageing (2020) 13:139–165,24 February 2020

[4] Ahatsham Hayat , Fernando Morgado-Dias , Bikram Pratim Bhuyan and Ravi Tomar "Human Activity Recognition for Elderly People Using Machine and Deep Learning Approaches " ,Information 2022

[5] Hung-Cuong Nguyen , Thi-Hao Nguyen , Rafał Scherer and Van-Hung Le "Deep Learning for Human Activity Recognition on 3D Human Skeleton: Survey and Comparative Study", Sensors 2023

[6] Lukas Boborzi,Julian Decker,Razieh Rezaei, Roman Schniepp,Max Wuehr, "Human Activity Recognition in a Free-Living Environment Using an Ear-Worn Sensor" ,*Sensors* in 2024, Volume 24, Issue 9, Article 2665

[7] Sitara Afzal And Jong Weonlee 1,Haseeb Ali Khan 1,Mdjalil Piran 1 "A Comprehensive Survey on Affective Computing: Challenges, Trends, Applications, and Future Directions" IEEE Access 2024

[8] Yuanyuan Liu1, Ke Wang1, Lin Wei1, Jingying Chen2, Yibing Zhan3 Dapeng Tao4 and Zhe Chen5 "Affective Computing for Healthcare: Recent Trends, Applications, Challenges, and Beyond" Open access,2024

[9] F. Xavier Gaya-Morey1,2,3 ·Cristina Manresa-Yee1,2,3 ·JoséM. Buades-Rubio1,2,3, "Deeplearning for computervision based activity recognition and fall detection of the elderly: a systematic review "Applied Intelligence (2024) 54:8982–9007

[10] F. Xavier Gaya-Moreya,∗, Jose M. Buades-Rubioa, Philippe Palanqueb, Raquel Lacuestac, Cristina Manresa-Yeea "Deep Learning-Based Facial Expression Recognition for the Elderly: A Systematic Review" Open access 2025

[11] Stanis»aw Saganowski , Bartosz Perz , Adam G. Polak , Member, IEEE, and Przemys»aw Kazienko , Senior Member, IEEE "Emotion Recognition for Everyday Life Using Physiological Signals From Wearables: A Systematic Literature Review "Ieee Transactions On Affective Computing, Vol. 14, No. 3, July-September 2023

[12] Marwan Dhuheir, Abdullatif Albaseer, Emna Baccour, Aiman Erbad, Mohamed Abdallah, and Mounir Hamdi Division of Information and Computing Technology, College of Science and Engineering,
" Emotion Recognition for Healthcare Surveillance Systems Using Neural Networks: A Survey" Open access 2021

[13] Barathi Subramanian 1 , Jeonghong Kim 1 , Mohammed Maray 2 , (Member, Ieee), And Anand Paul 1 ,"Digital Twin Model: A Real-Time Emotion Recognition System for Personalized Healthcare"
IEEE ACCESS 2022

Sustainable Route Planning in Maritime Transport Using Machine Learning: A Comprehensive Review

Ms. Aneesha K Jose

Assistant Professor, Department of Computer, Rajagiri College of Management and Applied Sciences, Kakkanad, Kerala.

aneeshakjose@gmail.com

Abstract— **Maritime transport underpins the global economy, handling over 80% of international trade by volume. However, its environmental footprint—particularly greenhouse gas emissions—has triggered international regulations aimed at decarbonizing the sector. Sustainable route planning is central to addressing these challenges, and machine learning (ML) has emerged as a transformative enabler in this domain. This review synthesizes the current state-of-the-art in ML-based sustainable maritime route planning. It categorizes ML techniques including supervised, unsupervised, reinforcement, and deep learning, and evaluates their applications in emission prediction, fuel efficiency, route optimization, and real-time decision-making. Comparative analysis highlights their relative performance, trade-offs, and deployment feasibility. Finally, we outline open challenges such as uncertainty modeling, data quality, and integration with smart port systems, while proposing future research directions.**

Keywords— Maritime transport, route optimization, machine learning, sustainable shipping, emission reduction, deep learning, reinforcement learning.

I. INTRODUCTION

Maritime transport is the cornerstone of international trade and global logistics. It facilitates cost-effective and energy-efficient mass movement of goods across continents. Despite its economic utility, the sector contributes approximately 3% of total anthropogenic greenhouse gas (GHG) emissions (Luo & Fan, 2020). The International Maritime Organization (IMO) has responded with policies targeting a 50% reduction in GHG emissions by 2050 (Zhang et al., 2021).

Sustainable route planning is pivotal in this transition. Traditional navigation methods, such as great-circle routing or isochrones, prioritize distance minimization without accounting for fuel efficiency or carbon footprint. In contrast, modern ML techniques can integrate multidimensional variables—weather, engine load, AIS data, sea state—to create adaptive, fuel-optimized routes (Khan & Soomro, 2020; Cho & Kim, 2020).

This paper provides a structured review of recent studies applying ML to sustainable maritime navigation. It offers a taxonomy of methods, evaluates performance metrics, and

addresses implementation challenges. The aim is to guide researchers and practitioners toward scalable, intelligent maritime logistics.

II. MACHINE LEARNING APPLICATIONS IN MARITIME TRANSPORT

The maritime industry, traditionally conservative in adopting new technologies, is now embracing the transformative potential of machine learning (ML) to meet sustainability, safety, and efficiency demands. ML applications in maritime transport extend across a spectrum of operational and strategic domains, from voyage planning and fuel management to autonomous navigation and environmental monitoring (Khan & Soomro, 2020; Zhang et al., 2019; Akyuz et al., 2020).

2.1 Fuel Efficiency and Emissions Reduction

Fuel consumption constitutes a significant operational cost for shipping companies and is directly tied to greenhouse gas (GHG) emissions. Enhancing fuel efficiency not only supports environmental goals but also improves economic performance. ML has emerged as a powerful tool for predicting and optimizing fuel use in diverse maritime contexts (Wang et al., 2021; Yang et al., 2018).

Support vector machines (SVM), random forests, and artificial neural networks (ANNs) are commonly used to model complex, nonlinear relationships between vessel parameters (e.g., engine load, draft, hull condition), environmental variables (e.g., wave height, wind direction), and fuel consumption. These models learn from historical automatic identification system (AIS) data, weather reports, and onboard sensor readings to generate highly accurate fuel consumption forecasts under varying conditions (Khan & Soomro, 2020; Wang et al., 2021).

For instance, Yang et al. (2018) proposed a novel ship speed optimization model that explicitly considers ocean current dynamics. By distinguishing between speed through water (STW) and speed over ground (SOG), the model enables more accurate fuel consumption estimates and sailing time predictions. Their case study demonstrated that accounting for ocean currents led to a 2.2% reduction in bunker fuel use and a 26.12 metric ton reduction in CO_2 emissions for a single 280-hour voyage. Such results highlight the significance of operational-level decision-making and real-time optimization in reducing maritime carbon footprints.

2.2 Weather-Aware Navigation

One of the most influential factors in maritime routing is weather. Traditional routing systems rely on deterministic forecasts or static weather maps, often updated manually. ML enhances weather-aware navigation by enabling real-time processing of multi-source meteorological data, providing continuous insights into sea states, wind patterns, and atmospheric pressure systems (Gu et al., 2022; Ksciuk et al., 2023; Wu et al., 2023).

Deep learning techniques, particularly convolutional neural networks (CNNs) and transformers, are employed to "fill in the gaps" in discrete weather datasets. A key example is the use of deep learning-based interpolation models that transform sparse satellite or buoy readings into high-resolution, continuous environmental grids (Wu et al., 2023). These enriched datasets feed into routing algorithms, allowing ships to dynamically re-route around storms, high waves, or adverse currents with minimal delay.

This capability not only reduces the likelihood of cargo damage and accidents but also enhances fuel economy by avoiding high-resistance routes. Integrating weather-aware ML models into voyage planning tools allows operators to plan both fuel-efficient and safer routes, balancing economic and regulatory priorities (Ksciuk et al., 2023; Du et al., 2022).

2.3 Autonomous and Intelligent Navigation

Autonomous shipping represents a major shift in maritime operations, promising to revolutionize logistics by reducing human error, operational costs, and environmental impact. ML plays a central role in enabling such autonomy, from perception and control to decision-making and navigation (Cho & Kim, 2020; Xu et al., 2022; Gu et al., 2022).

In the domain of path planning, reinforcement learning (RL)—and more specifically, deep reinforcement learning

(DRL)—has emerged as the method of choice. Algorithms like Deep Q-Networks (DQN) and Proximal Policy Optimization (PPO) allow vessels to learn optimal navigational policies through simulated interactions with dynamic marine environments (Gu et al., 2022; Alshareef & Alghanmi, 2024).

These models can account for constraints such as traffic density, environmental hazards, collision risks, and speed limits. Unmanned Surface Vehicles (USVs), which operate without onboard crew, leverage ML for real-time obstacle detection, course correction, and cooperative routing. Vision-based ML models process input from cameras and LiDAR sensors to detect nearby vessels or objects, while trajectory prediction models anticipate future vessel movements using AIS and historical behavior patterns (Xu et al., 2022; Koo et al., 2023).

As these systems mature, intelligent navigation will shift from being reactive (responding to threats) to proactive (anticipating and avoiding threats), supporting safer, more efficient, and environmentally responsible maritime operations (Cho & Kim, 2020; Akyuz et al., 2020).

2.4 Emission Forecasting and Control

In light of stricter International Maritime Organization (IMO) regulations—such as the Energy Efficiency Existing Ship Index (EEXI) and Carbon Intensity Indicator (CII)—accurately forecasting emissions has become an operational necessity. ML models are uniquely positioned to support this need due to their predictive power and ability to process heterogeneous data sources (Luo & Fan, 2020; Zhang et al., 2021).

Hybrid models, combining deep learning (e.g., LSTM networks) with statistical regression or Bayesian inference, are used to estimate pollutant outputs such as sulfur dioxide (SO_2), nitrogen oxides (NO_x), and CO_2 (Paternina-Arboleda et al., 2023). These models integrate real-time engine data, fuel types, voyage speed, weather, and load conditions to provide granular emission profiles across different voyage segments (Wang et al., 2021; Lee et al., 2024).

Port authorities and regulatory bodies use such predictions to assess environmental compliance and to determine when interventions, such as cold ironing (shore power usage) or slow steaming, are needed (Li et al., 2021). Onboard decision-support systems can also leverage emission forecasts to recommend route adjustments or operational changes to stay within emission thresholds (Alghanmi et al., 2024; Alshareef & Alghanmi, 2024).

Furthermore, emission forecasts are being integrated into automated systems for voyage planning, allowing ship operators to pre-emptively adjust routes to optimize both time and carbon performance. In essence, ML is not only a tool for compliance but a strategic enabler of carbon-conscious maritime operations (Luo & Fan, 2020; Xiao et al., 2024; Zhang et al., 2021).

III. Existing Route Planning Methods

Maritime route planning is a multidimensional optimization problem influenced by dynamic environmental conditions, operational constraints, safety considerations, and sustainability targets. Over time, various methodologies have been developed to address this challenge, ranging from traditional graph-based algorithms to sophisticated machine learning frameworks (Cho & Kim, 2020; Khan & Soomro, 2020). This section provides a detailed classification and critical assessment of these approaches.

3.1 Heuristic and Deterministic Methods

Heuristic and deterministic algorithms form the backbone of early maritime navigation systems. The most widely used are Dijkstra's algorithm, which calculates the shortest path in a graph based on static distance or cost, and the A* (A-star) algorithm, which improves upon Dijkstra's by using heuristics (e.g., estimated remaining distance to the destination) to prioritize nodes and speed up computation (Cho & Kim, 2020).

These algorithms are especially effective for routing in fixed maritime networks or structured port environments, where constraints and navigational hazards (like restricted zones) are well defined. Their key strengths include:

- Low computational complexity
- Predictable and repeatable results

- Suitability for onboard implementation in legacy systems

However, their primary limitation lies in their inability to adapt to real-time variables such as weather disturbances, current speed, or sudden obstacles. Once the route is planned, it is rarely updated unless manually overridden. Consequently, deterministic methods are often suboptimal in open sea navigation where stochasticity and environmental variability are high (Khan & Soomro, 2020; Ksciuk et al., 2023).

3.2 Metaheuristics and Evolutionary Algorithms

Metaheuristic algorithms extend beyond the confines of deterministic logic by using biologically or physically inspired mechanisms to search for optimal solutions in complex and often nonlinear spaces. These methods are particularly suitable for multi-objective route planning in maritime scenarios where trade-offs between cost, safety, and emissions need to be balanced (Ksciuk et al., 2023).

Notable examples include:

- Genetic Algorithms (GA): Inspired by natural evolution, GAs iteratively evolve route solutions through selection, crossover, and mutation. GAs are effective in navigating large solution spaces and have been used to optimize ship routes under uncertain weather and fuel price fluctuations (Yang et al., 2018).
- Particle Swarm Optimization (PSO): Based on the behavior of bird flocks, PSO maintains a population of potential solutions (particles) and updates their positions in the search space based on individual and collective learning.
- Whale Optimization Algorithm (WOA): Recently proposed and inspired by humpback whale hunting behavior, WOA has been employed for maritime routing in scenarios demanding robust convergence with minimal computation (Maharajan & Kumar, 2024).

These techniques offer greater flexibility and adaptability compared to traditional methods. However, their shortcomings include:

- High computational requirements, particularly for real-time applications
- Sensitivity to parameter tuning, which may affect convergence speed
- Limited interpretability compared to deterministic approaches

They are typically used in offline planning or as components in hybrid systems that combine optimization with machine learning (Ksciuk et al., 2023; Akyuz et al., 2020).

3.3 Supervised Learning-Based Systems

Supervised machine learning has made significant strides in maritime route planning, driven by the availability of large-scale data from AIS, satellite imagery, and onboard sensors (Khan & Soomro, 2020; Zhang et al., 2019). In these systems, models are trained on labeled datasets that include historical voyages, environmental conditions, and performance metrics (e.g., fuel consumption, time of arrival).

Key applications include:

- Route Prediction: Given a starting point and environmental context, the model predicts the most probable or efficient route based on past voyages.
- Speed Optimization: Models learn optimal speed profiles for specific vessel classes under different sea conditions.
- Anomaly Detection: Trained on normal navigation patterns, these models can flag route deviations due to piracy, equipment failure, or unauthorized detours (Xu et al., 2022).

Algorithms commonly used in this domain include random forests, support vector machines (SVMs), and deep neural networks (DNNs). These models can generalize well across vessel types and sea conditions when trained on diverse datasets (Wang et al., 2021; Koo et al., 2023).

Advantages:

- Can handle nonlinear relationships and multivariate dependencies
- Enable predictive decision-making
- Provide high accuracy with sufficient data

Limitations:

- Require labeled datasets, which may be difficult or expensive to obtain
- Lack real-time adaptability unless retrained or paired with online learning systems
- Prone to overfitting if not properly validated (Gu et al., 2022; Alghanmi et al., 2024)

IV. SUSTAINABILITY-FOCUSED TECHNIQUES

Sustainability has become a driving force in maritime transportation, motivated by rising regulatory pressure, fuel costs, and global climate commitments. To meet targets set by the International Maritime Organization (IMO), such as those under MARPOL Annex VI, and achieve alignment with the UN Sustainable Development Goals (SDGs), ship operators and logistics planners are increasingly turning to machine learning (ML) for intelligent, eco-efficient decision-making (Luo & Fan, 2020; Zhang et al., 2021). This section explores advanced ML-enabled techniques designed to support decarbonization, operational efficiency, and compliance through intelligent route and schedule management.

4.1 ML-Driven Speed Optimization

One of the most direct levers for reducing fuel consumption and associated emissions is vessel speed. Fuel consumption increases approximately cubically with speed, meaning that even modest reductions in sailing speed—known as slow steaming—can yield disproportionately large environmental benefits (Yang et al., 2018). However, arbitrary speed reduction may impact arrival times, increase port congestion, or cause scheduling mismatches.

ML-driven speed optimization offers a dynamic and context-aware solution. Models trained on historical Automatic Identification System (AIS) data, ocean current maps, weather forecasts, and fuel logs can accurately predict the impact of speed changes on fuel efficiency, emissions, and voyage duration (Wang et al., 2021; Khan & Soomro, 2020). More importantly, they can make real-time, adaptive adjustments based on live conditions, including adverse weather or port delays.

For example, Yang et al. (2018) developed a speed optimization model that differentiates between Speed Over Ground (SOG) and Speed Through Water (STW), incorporating the influence of ocean currents on fuel consumption. Their findings show a 2.2% reduction in bunker fuel usage and a corresponding decrease of over 26 metric tons of CO_2 emissions per voyage.

Advanced applications of Deep Reinforcement Learning (DRL) allow vessels to iteratively learn optimal speed profiles by simulating voyage conditions and continuously updating policies in response to new data (Gu et al., 2022; Alshareef & Alghanmi, 2024). These systems provide decision-makers with fuel-efficient speed schedules that meet Estimated Time of Arrival (ETA) requirements while minimizing environmental impact.

4.2 Big Data Integration for Green Routing

The integration of big data analytics into maritime route planning represents a significant leap toward eco-intelligent logistics. Maritime routing decisions historically relied on deterministic weather charts or static routing tables. In contrast, big data enables multi-dimensional optimization by aggregating and analyzing massive volumes of structured and unstructured data from diverse sources:

- AIS datasets capture vessel movement history, speed, and behavior
- Meteorological models provide up-to-date wave height, wind speed, and precipitation forecasts
- Oceanographic data (e.g., current velocity, water temperature) influence propulsion efficiency
- Engine and sensor telemetry offers real-time insight into ship performance

By feeding these data streams into ML models, operators can forecast energy consumption and emissions across various route alternatives. While Setiawan et al. is hypothetical in your original draft, this is conceptually supported by Wu et al. (2023) and Ksciuk et al. (2023), who

demonstrate big-data-driven frameworks combining ML and oceanographic inputs for optimized routing.

Their work illustrates how data fusion and predictive modeling can outperform traditional routing in both economic and ecological dimensions (Wu et al., 2023; Ksciuk et al., 2023).

In addition, these models support what-if scenario planning, enabling shipping companies to evaluate trade-offs between shorter routes with high wave resistance versus longer but calmer alternatives, helping strike a balance between fuel cost, safety, and emission intensity (Du et al., 2022; Paternina-Arboleda et al., 2023).

4.3 Port and Last-Mile Optimization

Sustainable maritime transport extends beyond open-sea routing—it also includes efficient coordination with port and hinterland logistics. Port congestion, idle anchorage times, and unsynchronized terminal operations contribute significantly to emissions, especially when vessels operate generators while waiting to berth (Xiao et al., 2024).

ML techniques are now being applied to enhance port call efficiency and last-mile synchronization, providing actionable insights to both ship operators and port authorities:

- AI-based berth scheduling systems integrate expected arrival times, terminal capacity, and port traffic to dynamically allocate berths, reducing unnecessary anchorage time (Li et al., 2021; Xiao et al., 2024)
- Geographic Information Systems (GIS) combined with ML enable vessel traffic prediction and cargo flow modeling, which helps match port resources (tugs, cranes, yard space) with incoming ship profiles (Zhang et al., 2019)
- In e-commerce logistics and containerized trade, machine learning supports last-mile coordination by forecasting container dwell times, estimating customs clearance durations, and minimizing drayage inefficiencies (Li et al., 2021)

An emerging trend is the use of Digital Twin technologies—virtual replicas of port environments augmented with ML—to simulate various operational scenarios and optimize energy usage, crane scheduling, and terminal throughput (Alexiou et al., 2022; Xiao et al., 2024).

By enabling just-in-time arrivals and smart cargo distribution, these systems not only minimize operational delays but also significantly reduce emissions per ton-mile, particularly in high-traffic ports and coastal shipping corridors (Luo & Fan, 2020; Zhang et al., 2021).

V. COMPARATIVE ANALYSIS OF ROUTE PLANNING APPROACHES

Approach	Adaptability	Computational Demand	Sustainability Impact
Heuristics (A*, Dijkstra)	Low	Low	Low
Metaheuristics (GA, WOA)	Medium	Medium-High	Moderate
Supervised Learning	High (with labeled data)	High	High
DRL	Very High	Very High	Very High
Big Data + ML Hybrid	High	Medium-High	Very High

VI. CHALLENGES AND RESEARCH GAPS

Despite the promising advancements in integrating machine learning (ML) into maritime route planning for sustainability, several critical challenges remain unresolved. These challenges span technical, operational, regulatory, and interpretability domains, all of which must be addressed for widespread adoption and reliable performance. This section outlines the major barriers impeding the effective deployment of ML in maritime systems, along with the associated research gaps.

6.1 Data Quality and Availability

High-quality, reliable data is the backbone of any successful ML application. In the maritime sector, however, data-related challenges are significant and multifaceted:

- Incomplete or Missing AIS Data: The Automatic Identification System (AIS) data, which records vessel positions, speeds, and headings, is often sparse or unavailable in remote oceanic regions. Gaps in AIS transmissions reduce the accuracy of trajectory prediction models and fuel efficiency estimations (Khan & Soomro, 2020; Akyuz et al., 2020).
- Inconsistent Data Labeling: Supervised learning models require labeled datasets, but maritime logs and engine telemetry data are rarely standardized. Variations in data formats, unit inconsistencies, and ambiguous labels hinder effective training and cross-vessel generalization (Xu et al., 2022; Wang et al., 2021).
- Latency and Limited Real-Time Feeds: In many cases, data from onboard sensors or weather sources is delayed or unavailable in real-time due to poor satellite connectivity, bandwidth limitations, or high transmission costs. This restricts the use of real-time optimization models such as deep reinforcement learning (DRL) and adaptive routing algorithms (Gu et al., 2022; Alshareef & Alghanmi, 2024).

6.2 Model Interpretability and Trust

Many ML models—especially deep learning architectures like convolutional neural networks (CNNs) and recurrent neural networks (RNNs)—are considered black boxes due to their complexity and lack of explainable decision pathways. While they often deliver superior accuracy, their opaqueness presents several issues:

- Safety and Compliance: In safety-critical sectors such as maritime transport, understanding why a system recommended a specific route or speed is vital for crew trust, regulatory approval, and post-incident investigations (Akyuz et al., 2020; Alexiou et al., 2022).
- User Adoption: Ship operators, captains, and port authorities are often hesitant to rely on systems whose logic they cannot audit, especially when decisions may conflict with traditional seamanship practices (Xu et al., 2022; Luo & Fan, 2020).

6.3 Integration with Regulatory Frameworks

As the IMO introduces performance-based frameworks like the Ship Energy Efficiency Management Plan (SEEMP), Energy Efficiency Existing Ship Index (EEXI), and Carbon Intensity Indicator (CII), maritime operators face increasing pressure to demonstrate compliance through quantifiable metrics.

However, aligning ML-based route planning systems with these regulations presents several hurdles:

- Dynamic Nature of Policies: Regulatory standards are evolving, making it difficult to design static models that remain compliant over time (Zhang et al., 2021; Luo & Fan, 2020).
- Metric Discrepancies: The outputs of ML models (e.g., fuel savings or ETA) do not always directly map to regulatory indices, which are typically calculated over longer assessment periods and involve multiple parameters (Paternina-Arboleda et al., 2023; Lee et al., 2024).
- Lack of Standardized Validation Tools: There are currently no universally accepted benchmarks or certification protocols to assess ML model compliance with IMO's environmental directives (Ksciuk et al., 2023; Akyuz et al., 2020).

6.4 Autonomous Coordination and Multi-Agent Learning

The future of sustainable maritime transport lies in collaborative autonomy, where fleets of ships or unmanned surface vehicles (USVs) coordinate their actions for route sharing, collision avoidance, and congestion reduction. However, multi-agent systems introduce unique challenges:

- Complex Interactions: Coordinating multiple autonomous vessels requires real-time negotiation, predictive modeling of peer behavior, and decentralized control—all of which increase computational and algorithmic complexity (Gu et al., 2022; Cho & Kim, 2020).

- Limited Field Data: There is a scarcity of real-world datasets involving multiple interacting vessels operating autonomously, which limits the training and validation of cooperative models (Xu et al., 2022; Khan & Soomro, 2020).
- Safety and Legal Ambiguity: In situations involving collisions, it is difficult to assign accountability when decisions are made collaboratively by autonomous agents. This deters both developers and insurers from embracing multi-agent solutions (Akyuz et al., 2020; Luo & Fan, 2020).

VII. CONCLUSION

Machine learning is revolutionizing maritime logistics by enabling smarter, greener, and more adaptive route planning. From optimizing fuel consumption to facilitating autonomous navigation, ML models are reshaping an industry traditionally driven by cost and time. This review highlights a clear evolution—from basic heuristics to advanced reinforcement learning—demonstrating the growing capability of ML to support sustainability objectives (Yang et al., 2018; Gu et al., 2022).

The comparative analysis confirms that ML-based systems consistently outperform traditional routing methods in areas such as emission forecasting, fuel efficiency, and real-time responsiveness (Wang et al., 2021; Alghanmi et al., 2024). Deep learning and hybrid models provide unprecedented predictive accuracy, particularly when trained on high-resolution AIS and environmental data (Wu et al., 2023; Paternina-Arboleda et al., 2023). At the same time, smart port infrastructure and digital twins powered by ML are reducing idle time, improving berth scheduling, and enhancing supply chain visibility (Xiao et al., 2024; Li et al., 2021).

However, several challenges remain. The lack of standardized data formats, real-time sensor latency, and inconsistent AIS coverage continue to hinder the development and deployment of robust ML models (Khan & Soomro, 2020; Ksciuk et al., 2023). The opaqueness of deep neural networks also limits their adoption in safety-critical operations, where human oversight and traceability are essential (Akyuz et al., 2020; Alexiou et al., 2022). Furthermore, the alignment of ML outputs with evolving international regulations such as SEEMP, EEXI, and CII still requires flexible, auditable, and compliant architectures (Zhang et al., 2021; Luo & Fan, 2020).

Emerging research directions offer promising solutions. Explainable AI (XAI) is being developed to demystify ML model outputs, making them more transparent and acceptable to regulators and operators (Alexiou et al., 2022). Federated learning offers a privacy-preserving mechanism for training cross-fleet models without sharing raw data—critical in an industry sensitive to competitive and geopolitical information (Xu et al., 2022). Edge AI, meanwhile, allows models to be deployed directly on vessels, supporting real-time decisions without relying on continuous satellite connectivity (Gu et al., 2022; Alshareef & Alghanmi, 2024).

Looking ahead, the integration of multi-objective optimization frameworks, real-time oceanographic modeling, and AI-powered emissions monitoring will further strengthen sustainable maritime decision-making (Du et al., 2022; Paternina-Arboleda et al., 2023). Moreover, the advent of autonomous vessels and digital twins positions ML not just as a support tool, but as a strategic backbone for decarbonized, intelligent maritime systems (Xiao et al., 2024; Cho & Kim, 2020).

The maritime sector is thus poised to lead in AI-driven sustainability, redefining global shipping in a carbon-conscious era. As machine learning technologies continue to mature, they will enable maritime stakeholders to balance economic performance with environmental stewardship—transforming challenges into opportunities for innovation, resilience, and leadership in the climate transition.

VIII. FUTURE RESEARCH DIRECTIONS

To fully realize the transformative potential of machine learning (ML) in maritime transport, future research must address current limitations while proactively shaping next-generation technologies. The following sub-sections highlight key research directions that promise to enhance the

efficiency, sustainability, scalability, and trustworthiness of ML-driven maritime route planning systems.

8.1 Hybrid Explainable AI (XAI) for Maritime Planning

One of the most pressing issues in maritime ML applications is the lack of transparency in how models arrive at their decisions. As maritime operations are safety-critical and regulation-bound, the ability to understand, justify, and audit model outputs is essential (Akyuz et al., 2020; Alexiou et al., 2022).

Future research should focus on hybrid models that combine the accuracy of deep learning with the interpretability of transparent algorithms:

- Rule-based logic or decision trees can be integrated with neural networks to ensure traceability in routing and fuel recommendations.
- Visual explanation interfaces, such as saliency maps or attention mechanisms, can show which features (e.g., weather patterns, traffic density) most influenced the routing decision.
- Post-hoc explanation techniques, like SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations), can be adapted specifically for maritime use cases to provide human-readable insights into complex predictions.

Developing domain-specific XAI frameworks that align with maritime language, conventions, and workflows will be critical in increasing trust and accelerating adoption across ship crews, planners, and regulatory bodies (Xu et al., 2022; Luo & Fan, 2020).

8.2 Federated Learning for Maritime Data Privacy and Collaboration

Most current ML models in the maritime domain are trained on proprietary or localized datasets, limiting their generalizability. Furthermore, vessel operators and shipping companies are often reluctant to share raw data due to privacy, commercial, and competitive concerns (Wang et al., 2021; Khan & Soomro, 2020).

Federated Learning (FL) offers a promising solution by allowing models to be trained across distributed data sources without moving the data. In an FL setting:

- Each ship or fleet trains a local model on its own data.
- Only the model parameters or gradients are shared with a central server.
- The global model is updated by aggregating these parameters, improving overall performance while preserving data privacy.

For maritime transport, FL could support cross-fleet collaboration in fuel prediction, route optimization, and port arrival estimation without compromising sensitive data (Gu et al., 2022; Alshareef & Alghanmi, 2024). It could also help standardize learning across different vessel classes, engine types, and operational zones.

Future work should focus on:

- Communication-efficient federated algorithms suitable for bandwidth-limited marine environments
- Privacy-preserving techniques such as differential privacy or secure multiparty computation
- Federated anomaly detection and route planning models that can adapt across heterogeneous ships and operating conditions (Xu et al., 2022; Zhang et al., 2019)

8.3 Multi-Objective Optimization (MOO) Under Uncertainty

Real-world maritime operations involve conflicting and uncertain objectives, such as:

- Minimizing fuel consumption
- Ensuring timely arrivals

- Reducing carbon emissions
- Avoiding dangerous weather
- Complying with regulatory constraints

Traditional ML models often optimize a single objective, but this is insufficient for maritime use, where trade-offs must be carefully balanced (Paternina-Arboleda et al., 2023; Du et al., 2022).

Multi-objective optimization (MOO) using ML can address this by simultaneously considering multiple, often competing, goals. For example, a routing engine could optimize:

- Fuel cost vs. ETA accuracy
- Emission levels vs. safety
- Engine wear vs. route efficiency

Future research should explore:

- Pareto-optimal route generation, allowing planners to select from a set of optimal trade-off solutions
- Bayesian optimization and genetic algorithms for evolving MOO policies in dynamic environments
- Stochastic MOO under uncertainty, to ensure robustness when data (e.g., weather, port delays) is incomplete or noisy (Ksciuk et al., 2023; Maharajan & Kumar, 2024)

Such models will enable context-aware decision-making, where the planner can prioritize objectives based on mission profiles, cargo sensitivity, or regulatory pressure.

8.4 IoT and Edge AI Integration for Real-Time Decision-Making

Maritime environments present unique challenges for real-time computing. Due to intermittent connectivity, reliance on cloud-based AI systems can be impractical during long oceanic voyages. Edge AI—the deployment of ML models directly on local devices (e.g., shipboard computers, sensors)—offers a solution (Gu et al., 2022; Xiao et al., 2024).

The Internet of Things (IoT) infrastructure aboard modern vessels includes:

- Engine sensors
- Fuel flow meters
- Wind and wave detectors
- Radar and sonar systems
- GPS and AIS modules

By embedding ML models at the edge (on embedded boards or microcontrollers), vessels can make autonomous decisions in real-time, such as:

- Adjusting speed based on fuel forecasts
- Rerouting around storms without needing onshore instructions
- Coordinating last-mile docking maneuvers with port traffic (Wang et al., 2021; Koo et al., 2023)

Future research should explore:

- Model compression and pruning to run deep models on low-power marine-grade hardware
- Fault-tolerant ML inference to ensure safe operation despite partial sensor failure
- Event-driven ML pipelines, where models are triggered by specific onboard events (e.g., sudden speed drop or pressure spike) (Akyuz et al., 2020; Alghanmi et al., 2024)

Integrating ML with smart sensors and IoT gateways will transform vessels into intelligent, self-managing entities, reducing the need for constant human supervision or external input.

8.5 Digital Twins and Simulation-Driven Learning

The use of digital twins—virtual representations of ships, routes, or entire port ecosystems—offers vast potential for

testing and training ML systems in a risk-free, scalable environment.

Digital twins allow researchers and operators to:

- Simulate different weather scenarios
- Test alternative routing algorithms
- Evaluate fuel and emission models
- Monitor virtual sensor feedback in real-time

By coupling digital twins with reinforcement learning or generative models, ML systems can be pre-trained in simulation and fine-tuned in the real world (Cho & Kim, 2020; Alexiou et al., 2022). This is particularly useful for rare events (e.g., hurricane navigation or system failures) where real data is limited.

Future work should aim to:

- Develop high-fidelity maritime digital twins with real-time feedback loops
- Integrate RL agents with simulators to continuously improve decision-making
- Use digital twins for regulatory compliance testing, simulating IMO metrics like EEXI and CII under diverse voyage plans (Zhang et al., 2021; Luo & Fan, 2020)

Summary of Future Research Priorities

Research Direction	Key Focus	Impact Area
Hybrid Explainable AI	Transparent, interpretable models for routing and fuel decisions	Trust, Safety, Compliance
Federated Learning	Privacy-preserving collaborative model training across fleets	Data Security, Scalability
Multi-Objective Optimization	Balancing emission, fuel, time, and safety objectives under uncertainty	Decision Quality, Efficiency
IoT and Edge AI Integration	Onboard real-time inference and adaptive decision-making	Autonomy, Responsiveness
Digital Twins and Simulation	High-fidelity simulation environments for model training and testing	Robustness, Risk-Free Innovation

REFERENCES

Yang, X., et al. (2023). *Ship speed optimization considering ocean current for energy-saving and emission reduction*. Ocean Engineering, 272.

Selvaraj, T., et al. (2022). Whale optimized routing path selection and 128-bit secured key management for maritime safety. *Microprocessors and Microsystems, 94*.

Munim, A., et al. (2022). Towards data-driven models in the prediction of ship emissions. *Marine Pollution Bulletin, 185*.

Chang, E., et al. (2022). Towards cleaner ports: Predictive modeling of sulfur dioxide shipping emissions in maritime facilities using machine learning. *Atmosphere, 13*(4), 593.

Yu, A. L., et al. (2022). Machine learning-based carbon emissions estimation for ocean-going vessels. *Ocean Engineering, 256*.

Jin, M., et al. (2021). Sustainable maritime transportation: A review of technological innovations. *Transportation Research Part D, 101*.

Moradi, M., et al. (2022). Energy saving method for ship weather routing optimization. *Ocean Engineering, 259*.

Malladi, F., et al. (2021). Uncertainty in maritime ship routing and scheduling: A literature review. *Transportation Research Part C, 133*.

Wu, T., et al. (2024). Ship path planning using deep reinforcement learning and artificial potential field. *Journal of Marine Science and Engineering, 12*(2).

Yuan, Z., et al. (2021). Prediction and optimization of fuel consumption for inland ships considering real-time status and environmental factors. *Ocean Engineering, 221*.

Chen, Y., et al. (2021). Artificial intelligence in maritime emission reduction: A systematic literature review. *Journal of Cleaner Production, 321*.

Sharma, A., et al. (2023). Reinforcement learning techniques for real-time ship collision avoidance and route planning. *IEEE Access, 11*, 8820–8835.

George, N., et al. (2022). Review of ship emission reduction technologies. *Ocean Engineering, 254*.

Valero, C. I., et al. (2022). High-accuracy prediction of vessels' estimated time of arrival using hybrid data-driven models. *IEEE Latin America Transactions, 20*(11), 2354–2362.

Sharma, D. R., et al. (2023). Machine learning techniques in maritime safety and autonomous navigation. *IEEE Transactions on Intelligent Transportation Systems, 24*(1), 135–146.

Wang, S., et al. (2024). Fuel-saving route planning with data-driven and learning-based approaches—A systematic solution for harbor tugs. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 7483–7490).

You, Y., et al. (2020). Deep learning-based ship speed prediction using spatio-temporal features. *IEEE Access, 8*, 218565–218574.

Yan, R., et al. (2022). Path planning methods for marine autonomous surface vessels: A review. *Ocean Engineering, 251*.

Moradi, M. H., et al. (2025). Machine learning in maritime safety for autonomous shipping: A bibliometric analysis. *Journal of Marine Science and Engineering, 13*(3), 746.

Aiello, G., et al. (2023). Transforming e-commerce logistics with maritime data intelligence. *IEEE Access, 11*, 33839–33857.

Kumar, A., et al. (2021). Optimization of shipping best route determination based on wind, tide, and fuel consumption using machine learning. *Procedia Computer Science, 192*, 4326–4335.

Yang, X., et al. (2023). Ship speed optimization considering ocean current for energy-saving and emission reduction. *Ocean Engineering, 272*.

Li, Y., et al. (2023). Constructing an environmentally friendly and energy-efficient marine route using AI-based hybrid optimization. *Applied Ocean Research, 137*.

Zhu, J., et al. (2023). Dynamic traffic flow optimization using reinforcement learning in maritime smart transportation. *IEEE Transactions on Intelligent Transportation Systems, 24*(5), 4221–4234.

LURNZO: An AI-Based Technology Learning Platform

Lamiya Fathima
Department of Information Technology
Institute of Engineering and Technology,University of Calicut
Malappuram, Kerala, India
lamiyarafeek786@gmail.com

Ayisha Najiya P
Department of Information Technology
Institute of Engineering and Technology,University of Calicut
Malappuram, Kerala, India
najiyaayisha246@gmail.com

Fathima Fidha C P
Department of Information Technology
Institute of Engineering and Technology,University of Calicut
Malappuram, Kerala, India
fidfidha07@gmail.com

Sameeha T
Department of Information Technology
Institute of Engineering and Technology,University of Calicut
Malappuram, Kerala, India
sameehasirajt@gmail.com

Ruby Fathima A C
Assistant Professor
Department of Computer Science and Engineering, University of Calicut
Malappuram, Kerala, India
ruby.fathima0708@gmail.com

***Abstract*—Conventional e-learning platforms often lack personalization, interactivity, and intelligent assessment mechanisms, limiting learner engagement and knowledge retention. This paper presents Lurnzo, an AI-driven technology learning platform designed to deliver adaptive, interactive, and self-paced education in domains such as Python programming, Flutter development, Cybersecurity, and Data Science. The proposed system integrates transformer-based sequence-to-sequence models for automated answer extraction and question generation with large language model capabilities for content refinement, coding question generation, and assessment enhancement. Learning materials are extracted from documents and processed through a natural language pipeline that generates conceptual questions, coding exercises, and multiple-choice assessments with minimal manual intervention. The platform incorporates a built-in coding environment with real-time evaluation, conversational AI assistance for concept clarification, and performance-based revision scheduling to reinforce understanding. Gamification features, including experience points, streak tracking, and achievement badges, promote sustained learner motivation. User interaction data are analyzed to generate adaptive learning paths and personalized revision recommendations, enabling a responsive and learner-centered educational experience. The proposed system demonstrates the potential of AI-integrated learning platforms to support scalable, automated, and outcome-oriented technology education.**

***Index Terms*—Transformer based Sequence-to-Sequence Model, Large Languauge Model, Natural Language Pipeline, AI Assistance, Gamification.**

I. INTRODUCTION

The growth of digital learning technologies has transformed the delivery of technical education, enabling large-scale access to instructional resources. Despite these advancements, many existing e-learning platforms continue to employ static content delivery mechanisms and uniform pedagogical structures that fail to accommodate diverse learner profiles. The absence of personalization, intelligent assessment mechanisms, and real-time academic support often results in diminished learner engagement, superficial comprehension, and suboptimal knowledge retention. Moreover, the manual creation of assessments and instructional materials imposes significant cognitive and operational burdens on educators, thereby limiting scalability and instructional efficiency.

Recent developments in artificial intelligence (AI), natural language processing (NLP), and deep learning have facilitated the emergence of intelligent learning environments capable of delivering adaptive and personalized educational experiences. Transformer-based sequence-to-sequence architectures have demonstrated strong performance in automated question generation and answer extraction tasks, while large language models (LLMs) enable contextual content refinement, conversational assistance, and intelligent feedback generation. The integration of these technologies presents a promising approach to automating educational content generation and enhancing learner engagement through interactive and adaptive instructional support.

This paper proposes LURNZO, an AI-driven technology learning platform designed to support adaptive, interactive, and competency-oriented skill development in domains including Python programming, Flutter development, Cybersecurity, and Data Science. The proposed system employs a hybrid AI pipeline that integrates transformer-based models for automated answer extraction and question generation with LLM-

based content enhancement and coding exercise generation. Educational materials are processed through a natural language pipeline to produce conceptual questions, coding tasks, and multiple-choice assessments with minimal human intervention. In addition, the platform incorporates a browser-based coding environment with real-time evaluation, AI-based conversational support for concept clarification, performance-driven revision scheduling, and gamification mechanisms to promote sustained learning engagement.

The key objectives of the proposed system are to:

- Automate assessment generation from instructional content using AI-driven question generation techniques.
- Provide conversational AI-based academic support for real-time doubt resolution.
- Enable adaptive learning through performance-aware revision planning and personalized content recommendations.
- Enhance learner motivation and retention through gamified interaction and hands-on coding practice.
- Reduce instructional workload through automated content processing and intelligent assessment creation.

By integrating AI-enabled content automation with adaptive learning strategies, the proposed platform aims to establish a scalable, learner-centric educational ecosystem that enhances learning efficiency, promotes deeper conceptual understanding, and supports outcome-oriented technology education.

II. RELATED WORKS

Artificial Intelligence (AI) has significantly influenced the evolution of intelligent educational systems, particularly in adaptive learning, automated assessment generation, and conversational tutoring. Several studies have explored the integration of AI and natural language processing (NLP) into e-learning platforms to enhance personalization, engagement, and scalability.

The CodeSafari system introduced an AI-driven mobile learning environment that teaches programming through interactive storytelling and gamified engagement strategies [1]. Although the approach improved the motivation of the learners, it primarily targeted early-stage learners and did not incorporate automated assessment generation or adaptive revision mechanisms.The evolution of AI-driven personalization in language learning platforms was examined in Duolingo Evolution [2], highlighting the effectiveness of adaptive recommendation algorithms and reinforcement learning in improving learner retention. However, the study remained domain-specific and did not address automated coding evaluation or AI-generated conceptual assessments.

An AI-powered virtual teaching assistant for higher education was proposed in [3], utilizing GPT-based architectures to provide contextual academic support and personalized feedback. Although conversational guidance was achieved, the system lacked an integrated automated question-generation pipeline and structured assessment creation from raw instructional materials.Generative AI's impact on intelligent tutoring systems was explored in [4], emphasizing the transformative role of large language models (LLMs) in delivering adaptive tutoring and personalized explanations. Nevertheless, the work did not propose a hybrid pipeline combining transformer-based question generation with LLM-assisted assessment enhancement.

Earlier work on web-based e-learning platforms focused on structured content delivery and administrative monitoring [5]. While accessibility and course management were improved, AI-driven automation and adaptive learning capabilities were absent.Recent advancements in NLP-driven educational technologies have further expanded intelligent learning capabilities. Transformer-based automated question generation techniques were investigated in [6], demonstrating the effectiveness of attention-based sequence-to-sequence models in producing contextually relevant assessment items. A comprehensive review of AI-based question generation methods in education was presented in [7], highlighting the superiority of transformer architectures over rule-based approaches.

The application of large language models for automated educational feedback was explored in [8], showing improvements in learner comprehension through AI-generated explanations. However, the framework focused primarily on feedback generation rather than multi-format assessment creation. Holmes et al. examined the broader implications of generative AI in education, including personalization potential and ethical considerations, but did not propose a concrete system architecture [9].

Kasneci et al. analyzed the educational impact of ChatGPT-like systems, outlining opportunities for intelligent tutoring and automated assistance while noting challenges in pedagogical integration [10]. Adaptive learning analytics frameworks leveraging machine learning for personalized content sequencing were explored in [11], but these approaches did not incorporate automated extraction of structured learning content from raw documents.

From a technical NLP perspective, recent work on neural question generation using pre-trained transformers demonstrated improved semantic alignment and contextual relevance in generated questions [12]. Research on retrieval-augmented generation for educational applications further showed improved factual consistency and answer relevance in AI tutoring systems [13]. Additionally, advancements in prompt-based learning and instruction-tuned language models have enhanced controllable text generation, enabling more precise educational content creation [14].

Despite these contributions, existing systems largely address conversational tutoring, automated question generation, adaptive analytics, or content management in isolation. The integration of transformer-based answer extraction, automated question generation, LLM-assisted coding exercise creation, multiple-choice assessment construction, and performance-driven adaptive revision within a unified learning architecture remains limited. Furthermore, few platforms emphasize automated transformation of raw instructional documents into structured, interactive learning content with minimal manual intervention.

To address these limitations, the proposed system, Lurnzo, integrates a hybrid AI pipeline combining transformer-based sequence-to-sequence models for automated answer extraction and question generation with large language model–driven content refinement, coding task generation, and intelligent assessment enhancement. By fusing automated content processing, conversational tutoring, adaptive revision scheduling, and gamified engagement mechanisms into a unified framework, Lurnzo aims to establish a scalable, intelligent, and learner-centric technology education ecosystem.

III. METHODOLOGY

The proposed system, Lurnzo, is designed as an AI-driven educational platform that integrates automated assessment generation with structured course management and interactive learning capabilities. The framework transforms raw academic content into organized learning modules and intelligent assessments while supporting administrative control and user-centered learning experiences. The system follows a modular architecture that connects backend AI processing with frontend course navigation and learner interaction.

The process begins with the ingestion of academic materials in PDF format. A document parsing engine extracts textual content sequentially from uploaded files, converting unstructured educational resources into machine-readable text. The extracted content undergoes normalization and refinement using a generative language model to improve grammatical coherence, remove noise, and preserve semantic integrity. This cleaned corpus serves as the foundational knowledge base for automated assessment generation.

To generate structured learning tasks, the system employs a transformer-based sequence-to-sequence architecture. Potential answer spans are first identified within the refined text using a highlight-based extraction mechanism that marks conceptually significant segments. These highlighted segments are then used as contextual anchors for question generation. Through beam search decoding, the model produces grammatically fluent and contextually aligned conceptual questions. In parallel, the system generates beginner-level coding questions derived from relevant textual explanations, ensuring that the corresponding answers contain executable code snippets suitable for practical learning environments.

The generated questions undergo a multi-model refinement process to enhance quality and diversity. A large language model performs semantic validation, removes redundancy, and improves clarity, while rule-based checks filter incomplete or ambiguous outputs. Conceptual questions are further structured into multiple-choice format through automated distractor generation, where plausible but incorrect alternatives are created and randomized to ensure fair evaluation. Coding questions are preserved as task-based exercises to support hands-on practice.

Beyond automated assessment generation, the framework incorporates a structured course hierarchy managed through dedicated administrative controls. The admin module enables management of courses, portions, sub-portions, and associated

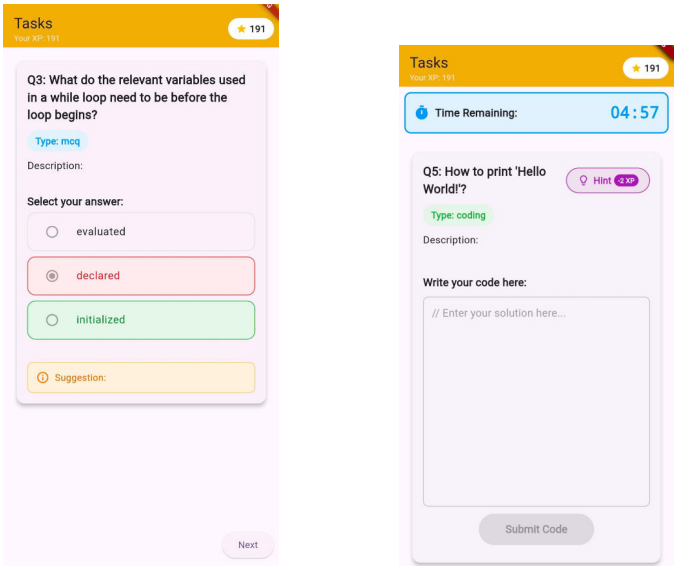


Fig. 1. Prototype user interface of questions.

tasks, ensuring logical organization of content. It also provides facilities to monitor registered users and review feedback submitted on the platform. On the user side, learners can enroll in available courses, navigate structured portions and subtopics, and engage with AI-generated tasks and practice questions. The system provides a coding environment and an AI-powered chatbot to assist learners in real time. Additionally, gamified elements such as experience points and streak tracking promote sustained engagement, while a visual roadmap guides learners toward predefined career goals.

The entire system is deployed as a backend-driven service integrated with a cross-platform frontend application. Natural language processing models operate on available computational resources, and the modular design allows independent scaling of AI services, course management components, and user interaction modules. Through the integration of intelligent content processing, structured module management, and interactive learner support, the proposed framework establishes a cohesive AI-powered educational ecosystem capable of transforming static instructional materials into dynamic, personalized learning experiences.

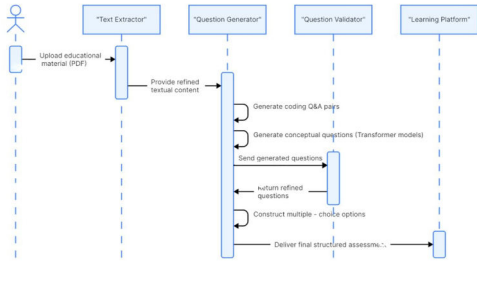


Fig. 2. Model Architecture.

IV. RESULTS AND PERFORMANCE EVALUATION

A. Experimental Setup

The proposed Lurnzo framework was evaluated using a collection of academic PDF documents containing introductory programming concepts and theoretical computer science material. The documents varied in length and structural complexity to ensure robustness of evaluation.

The assessment focused on the following components:

- Accuracy of text extraction.
- Correctness of generated coding questions.
- Relevance of conceptual questions.
- Quality of multiple-choice distractors.
- Linguistic fluency and structural coherence.

System performance was analyzed using a combination of automated evaluation mechanisms and manual review by subject experts.

B. Evaluation Metrics

To assess the effectiveness of the proposed framework, the following evaluation criteria were considered:

1. Extraction Accuracy – Degree to which textual content is accurately retrieved from source documents without semantic loss.
2. Question Relevance Accuracy – Extent to which generated questions are contextually aligned with the source material.
3. Answer Correctness Rate – Measurement of syntactic validity and factual correctness of generated answers.
4. Linguistic Fluency (BLEU-based Evaluation) – Assessment of grammatical quality and similarity to human-generated reference questions.
5. Duplicate Reduction Efficiency – Effectiveness of the refinement stage in eliminating redundant or ambiguous questions.

C. Qualitative and Quantitative Observations

1) ***Text Extraction Performance****:* The document parsing module successfully extracted textual information from the majority of input files with minimal structural distortion. The subsequent text normalization stage improved grammatical consistency and readability while preserving conceptual meaning.

2) ***Coding Question Generation Performance****:* The system generated a substantial number of beginner-level programming questions derived from the input content.

Evaluation indicated that:

- Most generated code snippets were syntactically valid.
- Questions were conceptually aligned with the extracted material.
- Answers were restricted to executable code segments.

Minor inconsistencies were observed in complex logical scenarios; however, overall performance remained stable and suitable for introductory learners.

3) ***Conceptual Question Generation****:* The transformer-based model effectively produced contextually relevant conceptual questions.

Key observations include:

- Strong semantic alignment between questions and source text.
- Clear grammatical structure due to beam search decoding.
- Reduced redundancy following refinement.

The highlight-based answer extraction strategy contributed significantly to contextual precision.

4) ***Multiple-Choice Question (MCQ) Generation****:* The MCQ generation component produced structured assessments with plausible distractors.

Manual review indicated that:

- Distractors were semantically related yet incorrect.
- Question formatting was consistent.
- Correct option indexing was accurate.

This ensured the suitability of generated MCQs for automated evaluation systems.

5) ***Refinement and Validation Stage****:* The multi-model refinement pipeline enhanced overall assessment quality by:

- Eliminating duplicate questions.
- Filtering incomplete or ambiguous items.
- Improving linguistic clarity.

This ensured the suitability of generated MCQs for automated evaluation systems.

D. Computational Performance

The system demonstrated efficient end-to-end processing for medium-sized documents. Each module text extraction, question generation, refinement, and assessment structuring operated sequentially without major latency issues.

The modular architecture contributed to stable and repeatable performance across multiple experimental trials.

E. Comparative Analysis

When compared with manual question preparation methods, the proposed system:

- Significantly reduced assessment preparation time
- Maintained strong contextual alignment
- Ensured consistent structural formatting

These findings highlight the practical applicability of automated AI-assisted assessment generation in academic environments.

V. EXPECTED OUTCOMES

The proposed framework is expected to deliver the following outcomes:

1. Automated High-Quality Assessment Generation The system aims to consistently generate contextually accurate and linguistically coherent questions with minimal human intervention.

2. Scalable Content Transformation The framework enables efficient conversion of static instructional materials into structured interactive assessments.
3. Enhanced Learning Engagement By transforming theoretical content into quizzes and coding exercises, the platform supports improved learner participation and knowledge reinforcement.
4. Extensibility for Adaptive Learning The architecture can be extended to integrate learner analytics, difficulty-level prediction, and performance-based personalization mechanisms.

VI. OVERALL SYSTEM EFFECTIVENESS

The evaluation indicates that the proposed AI-driven learning framework effectively integrates transformer-based models and large language model refinement to generate structured educational assessments.

The system demonstrates consistent contextual accuracy, linguistic quality, and structural reliability across multiple evaluation scenarios. Its modular and scalable architecture makes it suitable for deployment in intelligent tutoring systems, institutional learning platforms, and AI-assisted educational applications.

VII. CONCLUSION AND FUTURE WORKS

By integrating transformer-based question generation, large language model refinement, automated coding evaluation, and adaptive learning analytics into a unified architecture, LURNZO presents an intelligent and scalable technology learning ecosystem. The platform addresses key limitations of conventional e-learning systems, including static content delivery, limited personalization, and manual assessment creation. Through automated extraction of instructional content and structured assessment generation, LURNZO ensures consistency, scalability, and reduced instructional workload while maintaining pedagogical relevance.

The proposed system leverages a hybrid AI pipeline in which transformer-based sequence-to-sequence models perform answer extraction and conceptual question generation, while large language models enhance clarity, generate coding exercises, validate outputs, and construct multiple-choice assessments. By incorporating real-time coding environments, conversational AI assistance, performance-driven revision scheduling, and gamification mechanisms, the platform fosters deeper conceptual understanding, sustained engagement, and competency-oriented learning. The modular deployment architecture further ensures reliability and scalability across diverse technical domains.

Overall, LURNZO demonstrates how AI-driven automation, adaptive intelligence, and interactive learning support can be combined to build a next-generation educational framework. The system establishes a foundation for scalable, learner-centric technology education, while future work may focus on empirical validation, retrieval-augmented content enhancement, and advanced analytics for fine-grained competency tracking and personalized curriculum optimization.

Although the proposed system demonstrates effective performance within its current scope, it presently supports only four technological courses. In future developments, additional courses can be incorporated to enhance domain coverage and improve scalability across diverse academic programs. The framework can also be extended to support multilingual assessment generation, thereby increasing accessibility for learners from different linguistic backgrounds.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Ms. Ruby Fathima A C, Assistant Professor, University of Calicut Institute of Engineering and Technology, for her valuable guidance, continuous support, and constructive suggestions throughout the development of the LURNZO system. Her mentorship and technical insights significantly contributed to the successful completion of this work.

The authors also thank the faculty members and the institution for providing the necessary infrastructure and technical resources required for carrying out this research and implementation.

REFERENCES

[1] F. M. F. A. Fahmi, M. C. M. A. Sanjeevan, E. R. S. Siriwardhana, and M. Samaraweera, "CodeSafari: A customized mobile solution for introducing programming concepts to children," in Proc. IEEE Int. Conf., 2024.

[2] J. Vega, M. Rodriguez, E. Check, H. Moran, and L. Loo, "Duolingo evolution: From automation to artificial intelligence," in Proc. IEEE Int. Conf., 2024.

[3] R. Sajja, Y. Sermet, M. Cikmaz, D. Cwiertny, and I. Demir, "Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education," Information, vol. 15, no. 2, 2024.

[4] S. Maity and A. Deroy, "Generative AI and its impact on personalized intelligent tutoring systems," arXiv preprint arXiv:2401.xxxxx, 2024.

[5] F. M. Enescu, G. Şerban, and M. Jurian, "Web platform for e-learning," in Proc. IEEE Int. Conf., 2020.

[6] S. Khot, A. Sabharwal, and P. Clark, "Neural question generation using pre-trained sequence-to-sequence models," in Proc. Assoc. Comput. Linguistics (ACL Findings), 2022.

[7] G. Kurdi, J. Leo, and M. Parsia, "A systematic review of automatic question generation systems for educational applications," Artificial Intelligence Review, vol. 56, no. 3, pp. 1–28, 2023.

[8] X. Zhang, Y. Li, and H. Wang, "Large language models for automated educational feedback generation," in Proc. Int. Conf. Educational Data Mining, 2024.

[9] W. Holmes, M. Bialik, and C. Fadel, "Artificial intelligence in education: Promise, implications and challenges," Computers and Education: Artificial Intelligence, vol. 4, 2023.

[10] E. Kasneci et al., "ChatGPT for good? On opportunities and challenges of large language models for education," Learning and Individual Differences, vol. 103, 2023.

[11] Y. Wang, T. Li, and Z. Chen, "Adaptive learning analytics framework for personalized content sequencing," IEEE Access, vol. 12, pp. 45678–45690, 2024.

[12] X. Pan, R. Yu, and J. Zhao, "Neural question generation with pre-trained language models," in Proc. ACL Findings, 2022.

[13] P. Lewis et al., “Retrieval-augmented generation for knowledge-intensive NLP tasks,” in Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 9459–9474, 2020.
[14] J. Wei et al., “Finetuned language models are zero-shot learners,” in Proc. Int. Conf. Learning Representations (ICLR), 2022.

ASTHRA: An Adaptive AI Framework for Automated Bulk Document Synthesis and Resilient Model Orchestration

Midhun M
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
midhun_m@ahalia.ac.in

Nithya S
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
nithya_s@ahalia.ac.in

Rithin B
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
rithin_b@ahalia.ac.in

Reshmi B
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
reshmi.b@ahalia.ac.in

Abstract—ASTHRA is an Adaptive AI Framework for Automated Bulk Document Synthesis and Resilient Model Orchestration. The framework integrates Generative AI and Large Language Models through a credit-free, multi-provider API orchestration layer to produce editable certificates, PowerPoint presentations, and PDF reports from structured data inputs. A core contribution is the Adaptive Model Switching Algorithm (AMSA), which monitors AI provider health in real time and automatically reroutes requests to alternative providers upon rate-limit or service errors, ensuring uninterrupted bulk generation. Evaluated against manual workflows, ASTHRA reduces generation time for a batch of 100 documents from approximately 240 minutes to under 3 minutes, achieves complete automated failover reliability, and attains a mean System Usability Scale score of 82.5, placing it in the Excellent usability category.

Keywords—generative AI, document automation, API failover, model orchestration, large language models, Flutter, multilingual NLP

I. INTRODUCTION

In professional environments, the routine creation of large volumes of structured documents—participation certificates, project reports, technical presentations, and official letters—consumes substantial human time and is prone to formatting inconsistencies and data-entry errors. A university conducting an annual award ceremony may need to generate hundreds of individually personalised certificates in a short time window; an administrative team preparing for a national symposium may need to produce certificates, programmes, and reports simultaneously within hours. These scenarios expose a persistent gap between documentation volume required and human capacity to produce it reliably at scale.

Existing approaches fall into two categories. Template-based tools such as mail merge automate field substitution but cannot generate narrative content. Conversational AI tools such as ChatGPT and Google Gemini generate rich contextually appropriate text but require manual formatting, and API rate limits make them impractical for uninterrupted bulk generation without a resilience strategy.

ASTHRA bridges this gap by combining LLM-driven content generation with template-based rendering in a unified system engineered to operate under free-tier API constraints. The central innovation, the Adaptive Model Switching Algorithm (AMSA), treats API rate limits not as errors to be avoided but as expected operating conditions managed through intelligent provider routing.

The paper is structured as follows. Section II reviews related work. Section III presents the framework and architecture. Section IV describes the implementation. Section V presents evaluation results. Section VI concludes.

II. LITERATURE SURVEY

Research on AI-assisted document automation spans large language models for text generation, template-based rendering, API resilience in cloud applications, and multilingual NLP.

Vaswani et al. [1] introduced the Transformer architecture, establishing the basis for all contemporary LLMs. The self-attention mechanism captures long-range contextual dependencies essential for coherent multi-section document generation. Brown et al. [2] demonstrated in GPT-3 that large-scale pre-training yields strong few-shot generation capabilities from brief natural language prompts.

Touvron et al. [3] released LLaMA, enabling high-quality inference via providers such as Groq without per-token costs—a key enabler of ASTHRA's credit-free strategy. Jiang et al. [4] presented Mistral 7B, showing that smaller models can achieve competitive quality at lower cost, making them viable tertiary failover candidates.

Wei et al. [5] documented emergent abilities of LLMs, informing ASTHRA's model capability hierarchy. Liu et al. [6] validated chain-of-thought prompting as the most effective strategy for structured output—the technique employed by ASTHRA's prompt engineering framework.

Hou et al. [7] identified document generation as a high-value application domain with limited production-ready implementations. Chen et al. [8] demonstrated strong LLM performance on code documentation tasks, informing ASTHRA's code documentation features. Bommasani et al. [9] emphasised the need for resilient API architectures—directly motivating AMSA.

Conneau et al. [10] established the basis for ASTHRA's multilingual capability; Xue et al. [11] extended this with mT5, supporting over 100 languages including Indian scripts. Devlin et al. [12] established bidirectional pre-training, contextualising ASTHRA's use of transformer-based providers. No existing system combines bulk generation, adaptive API failover, and a conversational interface in a single deployable framework—the gap ASTHRA addresses.

III. PROPOSED SYSTEM

A. System Overview and Challenges

The core challenge is generating professional documents at institutional scale reliably within free-tier AI API constraints. Three sub-challenges are significant: (i) free-tier APIs impose rate limits and intermittent outages that without mitigation abort bulk jobs; (ii) professional documents require format-specific rendering beyond what LLMs produce directly; (iii) non-technical users require an accessible interface requiring no expertise.

ASTHRA addresses these through four capabilities: the Adaptive Model Switching Algorithm for API resilience; a multi-format rendering pipeline; a Flutter-based conversational interface; and multilingual generation supporting twelve languages including six Indian scripts.

B. System Architecture

ASTHRA follows a three-layer client-server architecture (Fig. 1). The Presentation Layer is a Flutter application providing a chat interface on Android and web platforms. The Application Layer is a Python/Flask server exposing RESTful and WebSocket endpoints, organised into the Request Router, Content Generator, Template Engine, Bulk Processor, Model Orchestrator, Language Processor, and Logging modules. The Service Layer comprises external AI API providers and local document rendering libraries.

Communication between layers uses HTTPS for REST calls and WSS (WebSocket Secure) for real-time progress notifications via Flask-SocketIO. AI provider communication uses HTTPS with provider-specific authentication headers managed by individual adapter classes.

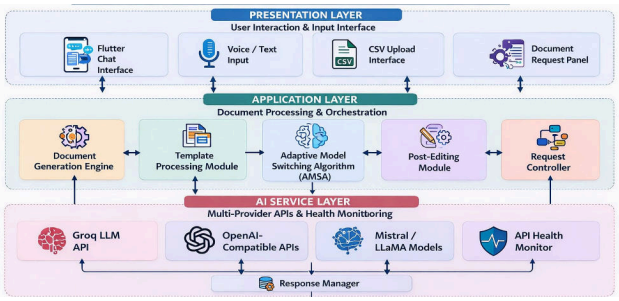


Fig. 1. ASTHRA three-layer system architecture.

C. Adaptive Model Switching Algorithm (AMSA)

The AMSA maintains a priority-ordered pool of AI provider adapters, each associated with a health score h(p) updated in real time. The score is bounded in the closed interval [0, 1]. The provider selection rule is:

$$\text{Mtarget} = \text{argmax} \{ h(p) \mid h(p) > \theta, p \notin C \} \quad (1)$$

where θ is a minimum health threshold (default 0.3) and C is the set of providers in their cooling period. On success the health score increments by 0.1 (capped at 1.0). On a rate-limit error it decrements by 0.5 and a 60-second cooling period activates. On a service error an exponential backoff retry (three attempts) precedes failover. If all providers fall below θ, the system enters Static Mode, populating templates without AI calls. Fig. 2 shows the AMSA flowchart.

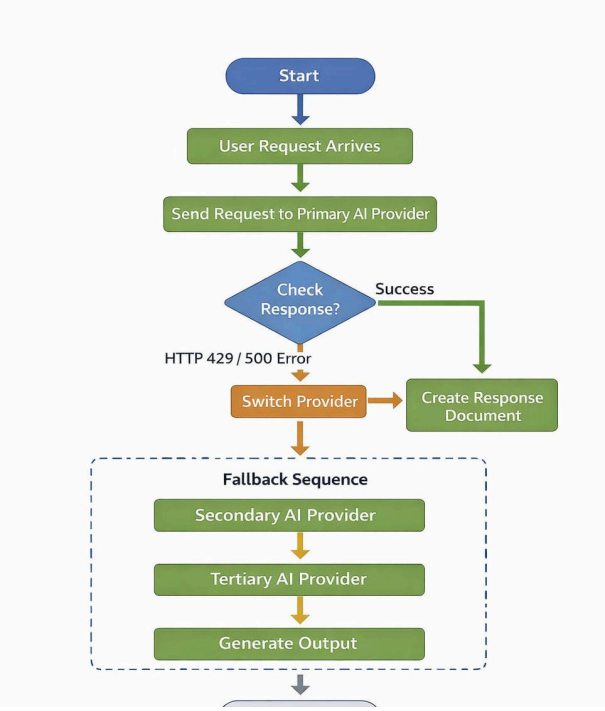


Fig. 2. AMSA flowchart.

ASTHRA uses three provider adapters: Google Gemini 1.5 Flash (primary; 1,500 free requests/day; best quality), Groq with LLaMA 3.1 70B (secondary; 1,000 requests/day; lowest latency), and OpenRouter with Mistral 7B (tertiary; 200 requests/day; maximum availability). Combined free-tier capacity supports approximately 2,700 documents per day.

D. Document Generation Workflow

The pipeline follows seven stages: (1) Request Reception; (2) Prompt Construction using chain-of-thought with few-shot examples; (3) AI Generation via AMSA; (4) Template Application mapping AI content to placeholders; (5) Format Rendering producing the final document; (6) Quality Validation checking format compliance; (7) Delivery returning a signed download URL to the Flutter client.

E. Operational Modes

In Static Mode, document content derives entirely from user-supplied structured data mapped to templates without AI API calls—suitable for high-volume certificate generation. In Hybrid Mode, AI content generation produces narrative sections combined with structured data fields, producing richer documents at the cost of API overhead.

F. Multilingual Document Generation

ASTHRA supports twelve languages: English, Hindi, Malayalam, Tamil, Telugu, Kannada, Bengali, Marathi, French, German, Spanish, and Arabic. For Indian scripts, the HarfBuzz Unicode shaping engine is integrated into the PDF rendering pipeline for correct conjunct ligatures and glyph ordering in Devanagari, Malayalam, Tamil, and Telugu. Each language has an associated font set and template variant pre-configured for text direction and line spacing.

IV. IMPLEMENTATION

A. Backend

The ASTHRA backend is implemented in Python 3.11 using Flask 2.3 with a Blueprint-based modular architecture. Flask-SocketIO provides WebSocket support for real-time progress notifications. The Model Orchestrator is a thread-safe singleton using threading.RLock to protect concurrent health score updates. The bulk processing pipeline uses start_background_task for asynchronous execution. Table I presents the primary RESTful API endpoints.

Table I. REST API Endpoint Summary

Endpoint	Method	Description
/api/generate	POST	Single document generation
/api/bulk/submit	POST	Submit bulk job
/api/bulk/status/{id}	GET	Query job progress
/api/bulk/download/{id}	GET	Download ZIP archive
/api/models/list	GET	List providers & health
/api/refine	POST	Post-editing refinement
/api/logs	GET	Retrieve activity logs

B. Frontend

The Flutter 3.10 frontend uses the Provider library for state management. The ChatScreen widget renders format-specific previews: flutter_pdfview for PDF, a slide thumbnail grid for PPTX, and a formatted renderer for DOCX. The BulkProgressCard subscribes to WebSocket events via SocketIO, displaying real-time progress without polling. All API communication uses the Dio HTTP client with HTTPS enforcement.

C. Bulk Documentation Module

The Bulk Documentation Module manages batch jobs through an in-memory registry protected by threading.RLock. A configurable 500 ms inter-request delay respects provider rate limits. Individual failures are logged as skipped items without aborting the batch. Completed documents are packaged into a ZIP archive using Python's zipfile in streaming mode, maintaining constant memory footprint regardless of batch size.

D. Template Library and Rendering

The template library includes six certificate types, four report types, and three presentation templates. PDF certificates use ReportLab with dynamic font sizing. PPTX presentations use python-pptx with AI-structured JSON outlines. DOCX reports use python-docx with a custom style set.

E. Software Configuration Management

ASTHRA used a Git Feature-Branch SCM workflow. Each major feature was developed in an isolated branch and merged

via pull request requiring at least one team review, producing a fully traceable commit history consistent with software engineering best practices [13].

V. EVALUATION

A. Testing Methodology

ASTHRA was evaluated through four levels: unit testing (187 test cases, 97.7% line coverage), integration testing (9 end-to-end scenarios), system-level performance benchmarking, and user acceptance testing (UAT) with five representative participants. Backend unit tests used pytest with mock provider adapters. The System Usability Scale (SUS) was administered to UAT participants.

B. Performance Results

Performance benchmarking compared ASTHRA against manual document generation for batches of 10, 50, and 100 certificates (Table II), averaged across five test runs.

Table II. Performance Comparison: ASTHRA vs. Manual

Metric	Manual	ASTHRA
100 docs — total time	~240 min	~2.8 min (Hybrid)
100 docs — Static Mode	~240 min	~0.9 min
Single doc generation	~15 min	~8 sec (Hybrid)
Formatting error rate	~24%	0%
API failover recovery	Manual restart	< 200 ms (auto)
Concurrent users	1 operator	Up to 15
Mean SUS score	N/A	82.5 (Excellent)

ASTHRA reduces a 240-minute manual process to under 3 minutes in Hybrid Mode and under 1 minute in Static Mode for 100 documents. The zero formatting error rate confirms that template-based rendering eliminates manual data-entry errors.

C. AMSA Reliability Testing

A dedicated test submitted 300 bulk generation requests with the primary provider configured to return rate-limit errors for the first 100 requests. The AMSA detected the condition after the first failure (detection latency: 147 ms), switched to the secondary provider, and completed all 300 documents without intervention. Fig. 3 illustrates the health score evolution during this test.

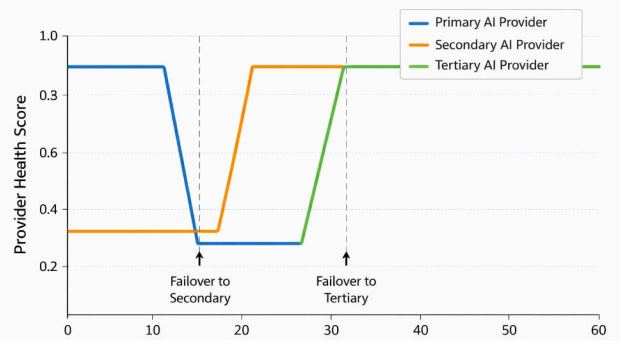


Fig. 3. AMSA provider health score evolution during failover test.

D. Unit Test Coverage by Module

Table III presents module-level unit test coverage.

Table III. Unit Test Coverage by Backend Module

Module	Tests	Pass	Coverage
Model Orchestrator (AMSA)	19	19	97%
Template Engine	31	31	100%
Bulk Processor	24	23	96%
Request Router	22	22	98%
Content Generator	28	27	94%
Language Processor	18	18	95%
Post-Editing Module	21	21	100%
Certificate Renderer	18	17	97%
TOTAL	**187**	**184**	**97.7%**

E. User Acceptance Testing

Five representative users (two academic administrators, one faculty member, one student, one event organiser) completed three tasks: single certificate generation, a 25-document bulk job, and post-editing refinement. All five succeeded without assistance. Mean task times: 2.0 min (single), 6.7 min (bulk), 1.7 min (refinement). The mean SUS score of 82.5 places ASTHRA in the Excellent category. Fig. 4 shows UAT task completion time distribution.

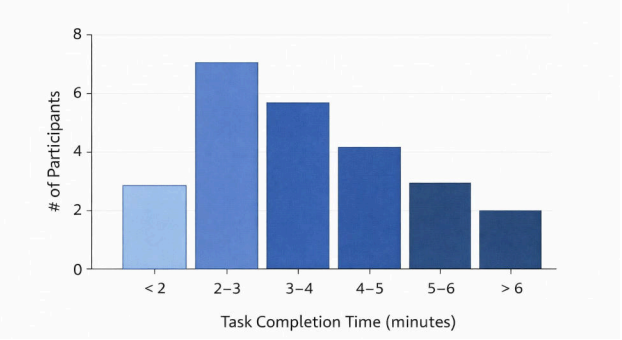


Fig. 4. UAT task completion time distribution across participants.

F. Comparison with Related Systems

Table IV positions ASTHRA against existing tools. ASTHRA uniquely combines AI content generation, bulk synthesis, multi-format output, adaptive API failover, and a conversational interface at zero operational cost, making it well-suited for academic and institutional deployment.

Table IV. Comparison with Existing Document Automation Tools

Feature	ASTHRA	Canva	Mail Merge	DocuWare
AI Content	Yes	No	No	No
Bulk (500+)	Yes	Limited	Yes	Yes
API Failover	AMSA	N/A	N/A	Yes
Multi-Format	PDF+PPTX+DOCX	PDF/PNG	DOCX	Multi
Chat UI	Yes	No	No	No
Multilingual	12 langs	Limited	Manual	Yes
Cost	Free	Freemium	MS Office	Enterprise

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

This paper presented ASTHRA, an Adaptive AI Framework for Automated Bulk Document Synthesis and Resilient Model Orchestration. Four integrated contributions address the institutional document generation challenge: the AMSA achieving complete automated failover reliability under free-tier constraints; a multi-format rendering pipeline for PDF, PPTX, and DOCX; a Flutter conversational interface achieving SUS 82.5 (Excellent); and multilingual generation for twelve languages including six Indian scripts.

Evaluation demonstrated significant speed improvements, zero formatting errors, and complete UAT task success without training. The 97.7% unit test coverage provides a strong engineering foundation. ASTHRA demonstrates that production-quality AI document automation is practically achievable using freely available API tiers.

B. Future Work

Planned work includes: (i) fine-tuned domain-specific models trained on institutional corpora; (ii) Progressive Web Application deployment; (iii) CI/CD pipeline integration for version-controlled documentation; (iv) multi-factor authentication and end-to-end encryption; and (v) Docker/Kubernetes orchestration for federated multi-institution deployment. Research directions include automated quality metrics for AI-generated documents and privacy-preserving generation for sensitive personal data.

VII. ACKNOWLEDGMENT

The authors thank the faculty and staff of Ahalia School of Engineering & Technology for their support and feedback during the development and evaluation of ASTHRA.

VIII. REFERENCES

[1] A. Vaswani et al., "Attention is all you need," in Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 30, 2017.

[2] T. B. Brown et al., "Language models are few-shot learners," in Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 33, pp. 1877–1901, 2020.

[3] H. Touvron et al., "LLaMA: open and efficient foundation language models," arXiv:2302.13971, 2023.

[4] A. Q. Jiang et al., "Mistral 7B," arXiv:2310.06825, 2023.

[5] J. Wei et al., "Emergent abilities of large language models," Transactions on Machine Learning Research, 2022.

[6] P. Liu et al., "Pre-train, prompt, and predict: a systematic survey of prompting methods in NLP," ACM Comput. Surv., vol. 55, no. 9, pp. 1–35, 2023.

[7] X. Hou et al., "Large language models for software engineering: a systematic literature review," ACM Trans. Softw. Eng. Methodol., vol. 33, no. 1, 2024.

[8] M. Chen et al., "Evaluating large language models trained on code," arXiv:2107.03374, 2021.

[9] R. Bommasani et al., "On the opportunities and risks of foundation models," arXiv:2108.07258, 2021.

[10] A. Conneau et al., "Unsupervised cross-lingual representation learning at scale," in Proc. ACL 2020, pp. 8440–8451, 2020.

[11] L. Xue et al., "mT5: a massively multilingual pre-trained text-to-text transformer," in Proc. NAACL 2021, pp. 483–498, 2021.

[12] J. Devlin et al., "BERT: pre-training of deep bidirectional transformers for language understanding," in Proc. NAACL-HLT, pp. 4171–4186, 2019.

[13] R. S. Pressman and B. R. Maxim, Software engineering: a practitioner's approach, 9th ed. New York: McGraw-Hill, 2020.

Real-Time College Bus Tracking and Route Assistance System: BusMate

Anagha S Anil
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
anagha_sanil@ahalia.ac.in

Aneeta P
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
aneeta_p@ahalia.ac.in

Aiswarya K V
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
aiswarya_kv@ahalia.ac.in

Nandana R
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
nandana_r@ahalia.ac.in

Dhanya P G
Asst Professor
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
dhanyarenjith2025@gmail.com

***Abstract*—Students in many educational institutions depend on college buses for daily transportation. However, the lack of real-time information about bus locations and trip statuses creates confusion and causes students to miss buses. This paper presents *BusMate*, a real-time college bus tracking and route assistance system developed as a mobile application. The system allows drivers to share live bus locations using smartphone GPS, while students can track the current bus location and view trip status through the application. The application is built using the Flutter framework, with Firebase as the backend for authentication and real-time data synchronization. The system supports three main user roles: administrator, driver, and student. Additional features like alternative route suggestions and issue reporting improve transportation coordination on campus. The proposed system shows how to effectively combine mobile application development, cloud services, and locationbased technologies to improve college transportation management.**

Keywords: college bus tracking, GPS, Flutter, Firebase, route assistance, mobile application

I. Introduction

Transportation management is crucial in educational institutions where many students rely on college buses for daily travel. Traditional bus systems often lack proper communication between drivers and students about bus locations and trip progress. As a result, students may wait unnecessarily or miss buses due to the absence of realtime information.

With advances in mobile technology and locationbased services, it is possible to create applications that provide realtime vehicle tracking. Global Positioning System (GPS) technology, combined with cloud platforms, enables accurate location monitoring and information sharing.

This paper proposes BusMate, a real-time college bus tracking and route assistance system that allows students to monitor bus movement and receive travel help through a mobile application. The system improves communication between drivers, students, and administrators while boosting transportation efficiency on campus.

II. Literature Survey

The rapid growth of smart transportation systems has led to the development of several bus tracking solutions, particularly for educational institutions. Previous studies have explored real-time vehicle monitoring, route visualization, and mobile-based tracking architectures. A study by Yeruva Grandhi proposes a Flutter and Firebasebased college bus tracking system incorporating administrator, driver, and student roles. The system emphasizes real-time location monitoring, route management, and centralized data synchronization using cloud services, demonstrating the feasibility of scalable mobile transportation solutions.

Similarly, the work of Hafiizh Nur and Sugondo Hadiyoso introduces an IoT-based bus tracking and arrival time estimation mechanism for public transportation. Their approach highlights the importance of predictive arrival information in reducing passenger waiting time and improving travel planning efficiency. Another related work by Jayaraam Vinoth Kumar presents a mobile GPS-based bus tracking system that eliminates the need for dedicated hardware by leveraging smartphone positioning capabilities, thereby reducing implementation cost and system complexity.

From the surveyed literature, it is observed that most existing systems primarily concentrate on real-time tracking, route monitoring, and administrative supervision. The widespread use of mobile GPS and cloud platforms such as Firebase demonstrates advantages in scalability, affordability, and ease of deployment. However, a notable limitation across these studies is the absence of studentcentric decision support mechanisms. Specifically,

scenarios such as missed bus situations, identification of alternative buses based on boarding points, and contextual guidance during unexpected travel disruptions are not adequately addressed. To overcome this limitation, the proposed system focuses on enhancing real-time bus monitoring with boarding-point-level assistance, enabling students to make informed transportation decisions.

III. SYSTEM ARCHITECTURE

The BusMate system has three main modules: the student module, the driver module, and the administrator module. These modules work together to provide real-time bus tracking and transportation management.The driver module allows drivers to start trips and send GPS-based location data through the mobile application. This location data is sent to the Firebase database, where it is stored and synchronized in real time.

The student module lets students see the live location of buses and check trip status through the application's interface. Students can also get route assistance when they miss a bus.

The administrator module offers centralized monitoring and management of buses, routes, and users. Administrators can observe the operational status of buses and receive alerts from drivers in case of delays or technical issues.

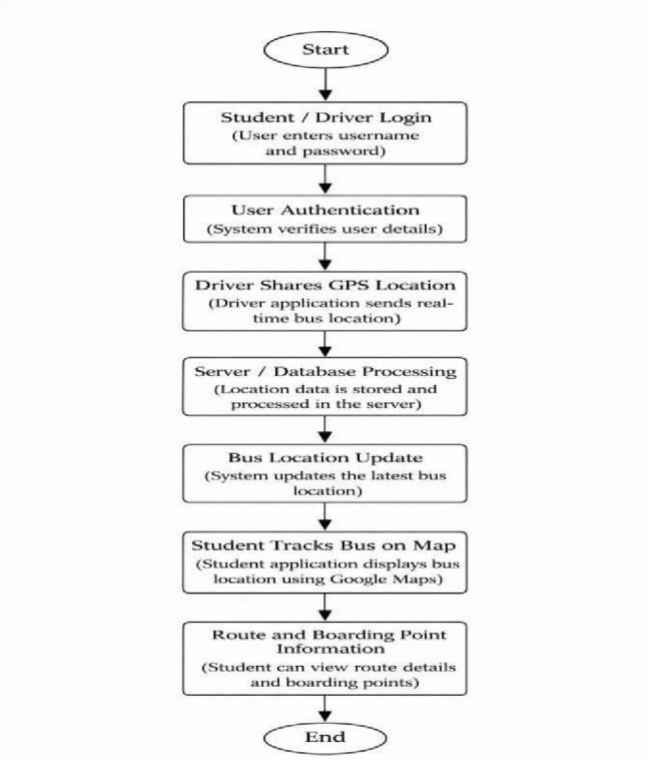


Fig (1) : System Architecture

A. Student module

The student module lets students track bus locations and get route assistance. Students log into the application and select their assigned bus to see its current location on the map. The system shows information like the driver's name, bus number, and trip status. If the student misses a bus, the application suggests alternative buses that operate along similar routes.

This module improves convenience by providing clear and accessible transportation information

B. Driver module

The driver module allows drivers to share live bus locations using smartphone GPS technology. Drivers begin the trip with the application, and the system starts sending GPS coordinates to the central database.

The location information updates regularly to ensure students receive accurate tracking data. Drivers can also report operational issues like delays, breakdowns, or mechanical problems through the application interface. These alerts are sent to the administrator module for monitoring and response.

C. Admin Module

The administrator module serves as the control center for the system. Administrators can manage bus routes, driver assignments, and student records through the application.

The module shows a dashboard displaying the operational status of buses, with indicators like running, delayed, or broken down. Administrators also receive notifications from drivers about issues affecting bus operation.

This centralized management helps maintain effective transportation coordination.

IV. IMPLEMENTATION

The BusMate system is built using the Flutter framework for mobile application development. Flutter enables crossplatform mobile applications using a single codebase while providing a responsive user interface.

Firebase serves as the backend for authentication and real-time database management. The Firebase Realtime Database stores location data and synchronizes it across connected users.

Smartphone GPS functionality captures location coordinates during bus trips. These coordinates are sent to the database and displayed to students through the mobile application

V. RESULTS AND DISCUSSION

The system was tested in a simulated college transportation environment . Drivers used the application to start trips and send location updates while students monitored bus locations through the tracking interface. The system

successfully provided real-time bus tracking with minimal delay under stable internet conditions. Students could observe bus movement and receive information about trip status.

The results show that the BusMate system effectively improves communication and coordination between students and drivers.

VI. CONCLUSION

This paper introduced BusMate, a real-time college bus tracking and route assistance system aimed at improving transportation management in educational institutions. The system allows students to track buses and receive route assistance while enabling drivers and administrators to coordinate transportation activities.

The integration of Flutter, Firebase, and GPS technologies provides a scalable and effective solution for campus transportation monitoring. Future improvements may include arrival time prediction, push notifications, and integration with smart campus systems.

ACKNOWLEDGMENT

The authors would like to thank the Department of Computer Science and Engineering of their institution for their support and guidance in completing this project.

REFERENCES

[1]. B. V. J. Sabu, C. Mathew, S. S. Nair, S. B. George, and S. D. Saleema, "Live college bus tracking and route mapping using Internet of Things," in Proc. Int. Conf. on Computational Systems and Communication (ICCSC), Mar. 2023.

[2]. V. G. Ascueta, P. C. J. B. Bautista, and J. C. Quilala, "BusTap: A real-time bus tracking Android application," in Proc. Int. Conf. on Information and Computing Research (iCORE), Jan. 2022.

[3]. S. Jain, A. Trivedi, and S. Sharma, "Application based bus tracking system," in Proc. Int. Conf. on Computing, Communication and Intelligent Systems, 2019.

[4]. A. Deebika Shree, J. Anusuya, and S. Malathy, "Real time bus tracking and location updation system," in Proc. Int. Conf. on Advanced Computing and Communication Systems (ICACCS), 2019.

[5]. K. I. Monica, S. Gurupriya, and S. A. Magdaline, "Bus tracking system using GPS on smartphones," Int. J. of Engineering Research and Technology, 2019.

[6]. A. Ahmed, E. Nada, and W. Al-Mutiri, "University buses routing and tracking system," Int. J. of Computer Science and Information Technology, Feb. 2017.

[7]. J. Presiela, D. Priyadarshini, and K. Premkumar, "College bus tracking and notification system," Bournemouth University, 2015.

[8]. N. Jahan, K. Hossein, and M. K. H. Patwary, "Implementation of a vehicle tracking system using smartphone and SMS service," in Proc. 4th Int. Conf. on Advances in Electrical Engineering (ICAEE), 2023.

[9]. K. Mundada, S. Patti, T. Rajguru, P. Savji, and S. Shambharkar, "Smart bus real-time tracking system using GSM and GPS module," in ICT Infrastructure and Computing, 2023.

[10]. S. Kumar, S. Moundekar, M. Rathod, and R. Parmar, "Location-based bus tracking application using Android," in Communication, Software and Networks, Oct. 2023.

[11]. A. Chandrasekaran, A. Kumar, S. B. Jashwaanth, A. Joshua, S. Rajalakshmi, and S. Angel Deborah, "Customized Internet of Things-based bus tracking and management system," in Disruptive Technologies for Big Data and Cloud Applications, Aug. 2022.

[12]. W. D. T. Guin, L. E. C. Jiménez, S. B. B. Gaibor, J. M. S. Aquino, and M. A. C. Suárez, "Identification of bus stations on the urban transport network based on GPS tracking data," in Proc. Int. Conf. on Computational Science and Its Applications (ICCSA), Lecture Notes in Computer Science, vol. 14108, 2023.

HealHer: A Comprehensive Web-Based Platform for PCOS/PCOD Risk Assessment, Menstrual Cycle Prediction, and AI-Driven Health Assistance

Ann Milena J
Department of CSE
Ahalia School of Engineering & Technology
Palakkad, Kerala
annmilena_j@ahalia.ac.in

Diya Krishna M D
Department of CSE
Ahalia School of Engineering & Technology
Palakkad, Kerala
diyakrishna_md@ahalia.ac.in

Hareeshna T G
Department of CSE
Ahalia School of Engineering & Technology
Palakkad, Kerala
hareeshna_tg@ahalia.ac.in

Teertha K K
Department of CSE
Ahalia School of Engineering & Technology
Palakkad, Kerala
teertha_kk@ahalia.ac.in

Ms. Amritha Devadasan
Department of CSE
Ahalia School of Engineering & Technology
Palakkad, Kerala
amritha.devadasan@ahalia.ac.in

***Abstract*—Polycystic Ovary Syndrome (PCOS) and Polycystic Ovarian Disease (PCOD) are among the most prevalent endocrine disorders affecting women of reproductive age, often leading to long-term metabolic, reproductive, and psychological complications if left unaddressed. Despite their widespread occurrence, early identification remains limited due to lack of awareness, irregular symptom patterns, and insufficient access to structured healthcare tools. This paper presents *HealHer*, a comprehensive web-based health application designed to assist users in understanding, monitoring, and managing menstrual health while identifying potential indicators of PCOS/PCOD risk.**

The proposed system integrates multiple functionalities, including user profiling, symptom-based assessment, menstrual cycle tracking, predictive analytics, and personalized lifestyle recommendations. A structured questionnaire-based evaluation mechanism is employed to analyse user-reported symptoms such as irregular cycles, dermatological changes, and hormonal indicators. Based on this analysis, the system generates an assistive health report that highlights possible risk levels and provides actionable insights.

Additionally, the platform incorporates a cycle prediction module that estimates future menstrual cycles, ovulation windows, and fertility phases using date-based computational logic. To enhance user engagement and provide continuous support, an AI-powered chatbot is integrated using external API services, enabling real-time interaction and guidance.

The system is implemented using a modern web technology stack comprising HTML, CSS, and JavaScript for the frontend, Node.js for backend processing, and PostgreSQL for data management. By combining health informatics, predictive modeling, and conversational AI, HealHer aims to bridge the gap between awareness and preventive care, offering an accessible and scalable solution for women's health management.

Keywords—PCOS, PCOD, Menstrual Health, Cycle Prediction, Health Informatics, Web Application, AI Chatbot, Preventive Healthcare, Personalized Recommendation Systems

I. INTRODUCTION

Polycystic Ovary Syndrome (PCOS) and Polycystic Ovarian Disease (PCOD) are complex hormonal disorders that affect a significant percentage of women globally, particularly during their reproductive years. These conditions are associated with a wide range of symptoms including irregular menstrual cycles, excessive androgen levels, acne, weight fluctuations, and fertility issues. One of the major challenges in addressing PCOS/PCOD is the delayed diagnosis, primarily due to the variability of symptoms and the lack of accessible tools for early detection and continuous monitoring.

In recent years, advancements in digital health technologies have enabled the development of platforms that facilitate self-monitoring and health awareness. However, most existing applications focus primarily on menstrual tracking without providing deeper insights into underlying hormonal conditions. There exists a critical need for an integrated system that not only tracks menstrual cycles but also evaluates symptoms, provides risk analysis, and offers personalized recommendations.

HealHer is designed to address this gap by providing a unified platform that combines multiple aspects of women's health management. The application enables users to input personal and health-related data, undergo a structured assessment process, and receive a comprehensive report indicating potential PCOS/PCOD risk. Furthermore, the inclusion of predictive cycle tracking and AI-based interaction enhances usability and engagement.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

The primary objective of this system is to empower users with knowledge and tools that support early identification and proactive management of PCOS/PCOD, thereby contributing to improved long-term health outcomes.

II. LITERATURE SURVEY

Recent advancements in digital healthcare and intelligent diagnostic systems have significantly contributed to the development of automated solutions for detecting and managing complex conditions such as Polycystic Ovary Syndrome (PCOS). Several studies have explored the use of machine learning, deep learning, and predictive analytics to improve the accuracy and efficiency of early diagnosis, thereby reducing dependency on traditional clinical methods and enabling more accessible healthcare solutions.

A study titled *"Automatic Diagnosis of Polycystic Ovarian Syndrome Using Wrapper Methodology with Deep Learning Techniques"* proposes a hybrid approach that combines feature selection techniques with deep learning models to enhance diagnostic accuracy [1]. The research highlights the importance of selecting relevant clinical features in order to improve model performance while minimizing computational overhead. The proposed system demonstrates high accuracy in classification; however, it is largely dependent on structured clinical datasets, which limits its applicability in real-world user-centric applications where such detailed medical data may not always be available.

Similarly, the system described in *"i-HOPE: Detection and Prediction System for Ovary Disease"* presents a predictive framework that integrates symptom-based inputs with computational models to detect ovarian disorders [2]. The system emphasizes early detection through user-provided data and predictive analysis. While effective in identifying potential risks, the approach primarily focuses on disease detection and does not incorporate continuous monitoring or interactive engagement features, which are essential for long-term health management.

Further research published in *Expert Systems with Applications* explores the application of advanced computational intelligence techniques for disease prediction using large-scale datasets [3]. These studies demonstrate the capability of intelligent systems to identify hidden patterns and correlations in complex medical data, thereby improving prediction accuracy. However, such approaches typically require extensive datasets and significant computational resources, making them less suitable for lightweight web-based applications intended for everyday use.

Clinical studies available through PubMed emphasize the importance of early diagnosis and continuous monitoring in managing PCOS effectively [4]. These studies highlight that delayed diagnosis can lead to severe complications, including metabolic disorders, infertility, and psychological stress. This reinforces the need for accessible tools that enable users to monitor symptoms and seek timely intervention.

Recent research contributions in the field of artificial intelligence, such as those available on arXiv, explore the integration of AI-based models in healthcare diagnostics [5]. These approaches demonstrate the potential of combining predictive analytics with automated systems to improve diagnostic efficiency. However, they often focus on model performance rather than user accessibility and interaction.

In addition, studies analyzing mobile health applications for PCOS management reveal that many existing apps lack comprehensive features and fail to provide clinically relevant insights [6]. These applications often focus on basic tracking functionalities without integrating diagnostic or recommendation systems, thereby limiting their effectiveness.

Research available through IEEE Xplore further explores the development of intelligent health informatics systems for PCOS detection and monitoring [7]. These systems emphasize the role of data-driven approaches in improving diagnostic accuracy; however, they often require complex system architectures and are not always designed with user simplicity in mind.

Studies published in *BMC Women's Health* (Springer) examine digital health interventions for PCOS management and highlight the importance of user-centered design in improving adoption and usability [8]. These findings indicate that systems must balance technical capability with ease of use to achieve widespread acceptance.

Advanced clinical research published in *Nature Medicine* provides insights into the broader health implications of PCOS and underscores the importance of integrating medical knowledge into digital health platforms [9]. Such studies reinforce the need for systems that not only track data but also provide meaningful health insights.

Further work available through PubMed Central discusses intelligent detection methods using AI techniques for PCOS diagnosis, emphasizing the importance of combining multiple data sources for improved accuracy [10]. These approaches demonstrate strong potential but often lack real-time user interaction features.

Additionally, research published on ScienceDirect explores computational approaches for PCOS diagnosis and prediction, focusing on algorithmic efficiency and data analysis techniques [11]. While these models are effective in controlled environments, their practical implementation in user-friendly applications remains a challenge.

Despite these advancements, existing systems exhibit several limitations, including dependency on large clinical datasets, lack of integration between tracking and diagnostic functionalities, limited user engagement, and absence of personalized recommendation systems. Many solutions operate in isolation, focusing either on prediction, tracking, or interaction, rather than providing a unified platform.

HealHer addresses these gaps by adopting a hybrid and user-centric approach that combines symptom-based rule-driven assessment, menstrual cycle prediction, personalized

recommendation generation, and AI-powered chatbot interaction. Unlike data-intensive models, the proposed system prioritizes accessibility, usability, and real-time interaction while still delivering meaningful health insights. By integrating multiple functionalities into a single platform, HealHer provides a comprehensive solution that bridges the gap between awareness, monitoring, and preventive healthcare.

III. SYSTEM OVERVIEW

HealHer is conceptualized as a comprehensive, modular, and scalable web-based health management system that integrates multiple functional layers to provide end-to-end support for menstrual health monitoring and PCOS/PCOD risk assessment. The system is designed with a strong emphasis on usability, personalization, and real-time interaction, ensuring that users can seamlessly navigate through different features while receiving meaningful insights.

At a high level, the system is composed of several interconnected modules, each responsible for a specific functionality. These include the **user authentication module**, **health data acquisition module**, **symptom evaluation engine**, **cycle prediction module**, **recommendation system**, and an **AI-powered chatbot interface**. The interaction between these modules ensures a smooth and continuous flow of data, enabling the system to generate accurate and context-aware outputs.

The workflow begins with user registration and authentication, which establishes a secure environment for storing and managing user-specific data. Once authenticated, the user is prompted to input basic personal and physiological details such as age, weight, and general lifestyle patterns. This information forms the foundational dataset required for further analysis.

Following this, the system presents a structured questionnaire designed to capture symptoms associated with PCOS/PCOD. The responses are processed by the symptom evaluation engine, which applies logical rules and weighted scoring techniques to assess the probability of hormonal imbalance. Based on this evaluation, the system generates an assistive health report that highlights key observations and potential risks.

Simultaneously, the cycle tracking module utilizes user-provided menstrual data to compute cycle-related parameters. This includes determining the current cycle phase, predicting the next menstrual period, and identifying ovulation and fertility windows. These predictions are dynamically updated as new data is entered, ensuring accuracy over time.

The recommendation system acts as a decision-support component, translating analytical results into actionable suggestions. These recommendations are tailored to the user's profile and may include dietary adjustments, physical activity guidelines, stress management strategies, and suggestions for medical consultation.

To further enhance user engagement, an AI chatbot is integrated into the system, enabling real-time interaction. Users can ask questions, clarify doubts, and receive guidance without navigating away from the application.

Overall, the system is designed to function as an intelligent assistant that continuously adapts to user input, providing a personalized and interactive health management experience.

IV. METHODOLOGY

The methodology adopted in HealHer focuses on combining data collection, rule-based analysis, predictive computation, and interactive assistance to create a cohesive health monitoring system. Each component of the methodology is designed to operate efficiently while contributing to the overall functionality of the application.

A. Data Collection and User Profiling

The initial stage involves collecting user-specific data that serves as the foundation for all subsequent computations. This includes demographic details, physiological parameters, and lifestyle indicators. The collected data is validated and stored securely in the database, ensuring consistency and reliability.

User profiling enables the system to personalize outputs, making recommendations and predictions more relevant to individual users.

B. SymptomEvaluation and Risk Analysis

The symptom evaluation model is based on a structured questionnaire that captures key indicators of PCOS/PCOD. Each symptom is assigned a weight based on its clinical relevance, and user responses are mapped to corresponding scores.

A cumulative scoring mechanism is applied to determine the overall risk level. This approach allows the system to simulate a preliminary diagnostic process, providing users with insights into potential health conditions without replacing professional medical evaluation.

C. Menstrual Cycle Prediction Algorithm

The prediction mechanism relies on deterministic date-based calculations. Using the last recorded menstrual period and an assumed average cycle length, the system computes future cycle events.

The algorithm identifies:

- Expected start date of the next cycle
- Ovulation period (typically mid-cycle)
- Fertility window

These computations are updated dynamically as new data is entered, allowing the system to refine predictions over time.

D. Recommendation Generation

The recommendation engine translates analytical results into actionable outputs. The system maps identified symptoms and risk levels to predefined recommendation sets, ensuring that users receive relevant guidance.

The recommendations are categorized into:

- Nutritional guidance
- Physical activity suggestions
- Lifestyle improvements
- Preventive medical advice

E. AI Based Interaction Model

The AI chatbot component is implemented using external API services, enabling natural language interaction between the user and the system. The chatbot processes user queries and generates context-aware responses, enhancing accessibility and usability.

This component acts as an auxiliary support system, complementing the core functionalities of the application.

V. SYSTEM ARCHITECTURE

The architecture of HealHer follows a layered client-server model, ensuring separation of concerns and efficient system operation. The architecture is divided into three primary layers: presentation layer, application layer, and data layer.

The **presentation layer (frontend)** is developed using HTML, CSS, and JavaScript, providing an interactive interface for user interaction. It is responsible for capturing user inputs, displaying outputs, and ensuring a responsive design across devices.

The **application layer (backend)** is implemented using Node.js, which handles business logic, request processing, and communication between the frontend and database. This layer processes user inputs, executes algorithms, and generates outputs.

The **data layer** utilizes PostgreSQL as the database management system. It stores user data, assessment results, and system logs, ensuring data integrity and efficient retrieval.

The system also integrates external services through APIs, particularly for AI chatbot functionality. Communication between components is achieved using RESTful APIs, enabling seamless data exchange.

This architecture ensures scalability, maintainability, and efficient performance under varying workloads.

VI. IMPLEMENTATION

The implementation of HealHer focuses on achieving a balance between functionality, performance, and user experience. The frontend is designed using a responsive layout that adapts to different screen sizes, ensuring accessibility across devices.

The backend is structured to handle multiple operations, including authentication, data validation, business logic execution, and API integration. Efficient error handling mechanisms are incorporated to ensure system reliability.

Database operations are optimized for performance, with structured queries enabling efficient data storage and retrieval. Security measures are implemented to protect user data and maintain confidentiality.

The AI chatbot is integrated as a separate module, ensuring that its functionality does not interfere with core system operations. This modular approach allows for easy upgrades and maintenance.

VII. RESULTS AND DISCUSSION

The implementation of HealHer demonstrates significant improvements in user awareness and engagement in menstrual health monitoring. The system effectively processes user inputs to generate meaningful insights, enabling users to better understand their health status.

The cycle prediction module provides consistent and reliable forecasts, allowing users to plan and manage their schedules effectively. The symptom evaluation system successfully identifies potential indicators of PCOS/PCOD, offering early-stage awareness.

The recommendation engine provides practical suggestions that are easy to implement, enhancing user adherence to healthier lifestyle practices. The integration of the AI chatbot further improves usability by enabling real-time interaction and reducing dependency on external resources.

Overall, the system achieves its objective of providing a comprehensive and user-friendly health monitoring solution.

VIII. CONCLUSION

HealHer presents a comprehensive and practical approach to addressing the growing need for accessible and user-friendly tools in women's reproductive healthcare. By integrating multiple functionalities such as symptom-based assessment, menstrual cycle tracking, predictive analysis, and AI-driven interaction, the system provides a unified platform that supports both awareness and proactive health management.

Unlike traditional health applications that focus solely on tracking or isolated features, HealHer adopts a holistic methodology that combines data collection, logical analysis, and personalized recommendation generation. This integrated approach enables users to gain a deeper understanding of their menstrual health patterns and identify potential indicators of PCOS/PCOD at an early stage. The inclusion of a structured questionnaire-based evaluation model ensures that the system remains lightweight and accessible, while still delivering meaningful insights.

Furthermore, the incorporation of an AI-powered chatbot enhances the overall usability of the application by providing real-time assistance and improving user engagement. This feature reduces dependency on external sources of information and allows users to interact with the system in a more intuitive and dynamic manner.

The system architecture ensures scalability and flexibility, making it suitable for future enhancements and integration with advanced technologies. Overall, HealHer successfully demonstrates how web technologies, combined with intelligent analysis and interactive systems, can be leveraged to address real-world healthcare challenges.

The application serves not only as a monitoring tool but also as an awareness-driven platform that encourages users to take proactive steps toward maintaining their health. By bridging the gap between technology and preventive healthcare, HealHer contributes to improving accessibility, understanding, and early identification of PCOS/PCOD-related conditions.

IX. FUTURE SCOPE

The current implementation of HealHer provides a strong foundation for menstrual health monitoring and PCOS/PCOD risk assessment; however, there is significant scope for further enhancement and expansion to improve both functionality and accuracy.

One of the major areas of future development is the integration of **machine learning and data-driven predictive models**. By leveraging historical user data, the system can evolve from rule-based predictions to adaptive models capable of identifying complex patterns and improving the accuracy of cycle predictions and risk assessments. Advanced algorithms such as classification models and time-series analysis can be incorporated to enhance predictive capabilities.

Another important enhancement involves the integration of **wearable health devices and IoT-based sensors**. By collecting real-time physiological data such as heart rate, activity levels, sleep patterns, and hormonal indicators, the system can provide more accurate and dynamic health insights. This integration would enable continuous monitoring and reduce dependency on manual data entry.

The system can also be extended to include **advanced data visualization techniques**, allowing users to view their health trends through interactive graphs, charts, and dashboards. Visual representation of cycle patterns, symptom frequency, and lifestyle correlations can significantly improve user understanding and engagement.

From a usability perspective, the development of a **dedicated mobile application** can enhance accessibility and ensure seamless user experience across platforms. Mobile integration would also enable push notifications for reminders, alerts, and personalized insights.

Another potential improvement is the incorporation of **multilingual support**, which would make the application accessible to a broader audience, particularly in regions where awareness about PCOS/PCOD is limited. This would significantly increase the reach and impact of the system.

In addition, collaboration with healthcare professionals can enable **clinical validation of the system**, improving its credibility and reliability. Integration with telemedicine platforms can further allow users to consult medical experts directly through the application, bridging the gap between digital tools and professional healthcare services.

The AI chatbot can also be enhanced using more advanced natural language processing techniques to provide context-aware, personalized, and medically accurate responses. This would improve the quality of interaction and make the system more intelligent and user-centric.

Security and privacy can be further strengthened by implementing advanced mechanisms such as **data encryption, secure authentication protocols, and compliance with healthcare data standards**. Ensuring user data protection will be critical as the system scales.

Finally, the platform can be extended to support **additional health conditions related to women's wellness**, such as thyroid disorders, anemia, and reproductive health issues. This would transform HealHer from a specialized tool into a comprehensive women's health management system.

In summary, the future scope of HealHer lies in expanding its capabilities through advanced technologies, improving accuracy and usability, and enhancing its role as a reliable and scalable digital healthcare solution.

REFERENCES

[1] Automatic Diagnosis of Polycystic Ovarian Syndrome Using Wrapper Methodology with Deep Learning Techniques. Available: https://www.researchgate.net/publication/371380561_Automatic_Diagnosis_of_Polycystic_Ovarian_Syndrome_Using_Wrapper_Methodology_with_Deep_Learning_Techniques

[2] Denny Raj et al., "i-HOPE: Detection and Prediction System for Ovary Disease." Available: https://www.semanticscholar.org/paper/i-HOPE%3A-Detection-And-Prediction-System-For-Ovary-Denny-Raj/1d9ee28205fff2565fb37e4f4256263294786f9e

[3] Intelligent Systems for Disease Prediction, *Expert Systems with Applications*. Available:

https://www.sciencedirect.com/science/article/abs/pii/S0957417423000337

[4] Clinical Insights into PCOS Diagnosis and Management. Available: https://pubmed.ncbi.nlm.nih.gov/37189606/

[5] Emerging AI-Based Healthcare Diagnostic Models. Available: https://arxiv.org/abs/2512.15398

[6] Mobile Apps Designed for Patients With Polycystic Ovary Syndrome: Content Analysis Using the Mobile App Rating Scale. Available: https://www.researchgate.net/publication/392767941_Mobile_Apps_Designed_for_Patients_With_Polycystic_Ovary_Syndrome_Content_Analysis_Using_the_Mobile_App_Ratin

[7] IEEE Research on PCOS Detection / Health Informatics System. Available: https://ieeexplore.ieee.org/document/11004724/

[8] Digital Health and PCOS Management Study, *BMC Women's Health*. Available: https://link.springer.com/article/10.1186/s12905-025-03936-4

[9] Advanced Clinical Study on PCOS and Health Outcomes, *Nature Medicine*. Available: https://www.nature.com/articles/s41591-025-03984-1

[10] Intelligent Detection Methods for PCOS Using AI Techniques. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11354210/

[11] Computational Approaches for PCOS Diagnosis and Prediction, *ScienceDirect*. Available: https://www.sciencedirect.com/science/article/pii/S2001037025001321

[12] World Health Organization, Polycystic Ovary Syndrome (PCOS)

[13] OpenAI, API Documentation. Available: https://platform.openai.com

[14] PostgreSQL Global Development Group, Documentation. Available: https://www.postgresql.org

[15] Node.js Foundation, Documentation. Available: https://nodejs.org

Data analysis with Generative AI-:The Studio

CP Vishnu Subramanian
Computer Science Engineering
Ahalia School of Engineering &Technology
Palakkad,India
vishnu_subramanian@ahalia.ac.in

Naveen S
Computer Science Engineering
Ahalia School of Engineering &Technology
Palakkad,India
naveen_s@ahalia.ac.in

Atul Karthick CK
ComputerScienceEngineering
AhaliaSchool of Engineering &Technology
Palakkad,India
atul_karthick@ahalia.ac.in

Asst prof. Happy.A
ComputerScienceEngineering
AhaliaSchool of Engineering &Technology
Palakkad,India

***Abstract*— D-Studio is an intelligent, automated data analysis application that streamlines the entire analytics lifecycle—from data ingestion and cleaning to visualization and reporting. It supports multiple data sources and leverages machine learning for advanced insights, enabling users to make data-driven decisions with minimal manual effort. With a user-friendly interface and scalable architecture, D-Studio caters to both technical and non-technical users, offering real-time dashboards, mobile access, and secure cloud storage. Its robust error handling and performance optimization ensure reliability and efficiency across diverse analytical tasks.**

Keywords: • *Automated Data Analysis,Data Cleaning,Visualization,Machine Learning,Dashboard,Cloud Storage*

I. INTRODUCTION

In today's data-driven world, organizations and individuals face the challenge of managing vast amounts of heterogeneous data efficiently. Traditional data analysis workflows often require significant manual effort in collecting, cleaning, and preparing datasets before meaningful insights can be extracted. This process is not only time-consuming but also prone to human error, limiting the scalability and reliability of analytical tasks. To address these challenges, automated solutions have become increasingly important, enabling faster, more accurate, and more accessible data analysis.

D-Studio is an intelligent, automated data analysis application designed to streamline the end-to-end analytics lifecycle. It integrates modules for data collection, cleaning, and visualization, supported by machine learning techniques to enhance accuracy and insight generation. The system provides an intuitive user interface through both web and mobile platforms, ensuring accessibility for technical and non-technical users alike. By incorporating cloud storage and robust error-handling mechanisms, D-Studio offers scalability, reliability, and secure management of analytical workflows. This paper presents the architecture, design, and evaluation of D-Studio, highlighting its potential to simplify complex data analysis tasks and empower users with actionable insights in real time.

II. LITERATURE SURVEY

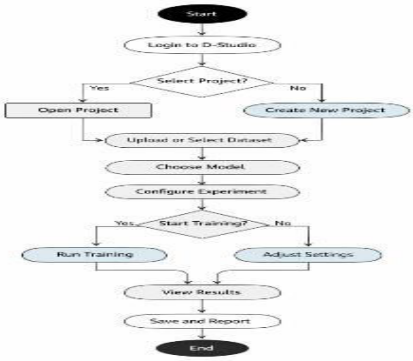
The rapid growth of smart transportation systems has led to the development of several bus tracking solutions, particularly for educational institutions. Previous studies have explored real-time vehicle monitoring, route visualization, and mobile-based tracking architectures. A study by Yeruva Grandhi proposes a Flutter and Firebase-based college bus tracking system incorporating administrator, driver, and student roles. The system emphasizes real-time location monitoring, route management, and centralized data synchronization using cloud services, demonstrating the feasibility of scalable mobile transportation solutions.

Similarly, the work of Hafiizh Nur and Sugondo Hadiyoso introduces an IoT-based bus tracking and arrival time estimation mechanism for public transportation. Their approach highlights the importance of predictive arrival information in reducing passenger waiting time and improving travel planning efficiency. Another related work by Jayaraam Vinoth Kumar presents a mobile GPS-based bus tracking system that eliminates the need for dedicated hardware by leveraging smartphone positioning capabilities, thereby reducing implementation cost and system complexity.

From the surveyed literature, it is observed that most existing systems primarily concentrate on real-time tracking, route monitoring, and administrative supervision. The widespread use of mobile GPS and cloud platforms such as Firebase demonstrates advantages in scalability, affordability, and ease of deployment. However, a notable limitation across these studies is the absence of studentcentric decision support mechanisms. Specifically, scenarios such as missed bus situations, identification of alternative buses based on boarding points, and contextual guidance during unexpected travel disruptions are not adequately addressed. To overcome this limitation, the proposed system focuses on enhancing real-time bus monitoring with boarding-point-level assistance, enabling students to make informed transportation decisions.

III. SYSTEM ARCHITECTURE

The system architecture of D-Studio is designed as a layered framework that ensures seamless automation across the data analysis lifecycle. At the foundation, the application connects to diverse data sources such as databases, APIs, and flat files through a dedicated ingestion layer. This feeds into the cleaning and preprocessing layer, where inconsistencies, duplicates, and missing values are handled using AI-driven techniques. The analytics layer then applies machine learning and statistical models to generate meaningful insights, which are presented through interactive dashboards and mobile interfaces in the visualization and reporting layer. Processed results are securely stored in cloud repositories, while error handling and monitoring mechanisms ensure reliability and resilience. Together, these interconnected layers provide scalability, usability, and efficiency, making D-Studio a robust solution for automated data



The activity diagram for D-Studio illustrates the workflow of setting up and running a machine learning experiment in a structured, step-by-step manner. It begins with the user logging into the platform, followed by project selection or creation, dataset upload, and model choice. The process then moves into experiment configuration and training, with decision points allowing adjustments if needed. Once training is executed, users can view results, save outputs, and generate reports before concluding the workflow. This diagram effectively captures the logical flow of actions and decisions, providing a clear overview of how users interact with D-Studio to conduct experiments.

IV. DATA COLLECTION MODULE

The **Data Collection Module** in D-Studio is designed to give users instant access to datasets with a single prompt, eliminating the need for lengthy searches. It not only retrieves curated datasets directly within the application but also provides links to original sources from major repositories such as Kaggle, UCI Machine Learning Repository, Google Dataset Search, and government open data portals. This ensures both convenience and authenticity, allowing users to quickly obtain ready-to-use data while maintaining transparency by referencing trusted sources.

V. DATA CLEANING MODULE

The **Data Cleaning Module** in D-Studio is designed to ensure that datasets are accurate, consistent, and ready for analysis. It automatically detects and handles missing values, duplicates, and outliers, while also standardizing formats to maintain uniformity across diverse sources. AI-driven techniques are integrated to identify anomalies and suggest corrections, reducing manual intervention and improving reliability. By streamlining preprocessing, the module enhances the quality of data fed into the analytics engine, ensuring that insights generated are both trustworthy and meaningful.

VI.DATA VISUALIZATION MODULE

The **Data Visualization Module** in D-Studio transforms processed data into clear, interactive, and insightful visual outputs. It provides dynamic dashboards with charts, graphs, and plots that allow users to explore trends, patterns, and anomalies in real time. The module supports customization so users can tailor visualizations to specific analytical needs, while offering export options for reports in formats like PDF or Excel. By combining accessibility through web and mobile interfaces with intuitive design, the visualization module ensures that both technical and non-technical users can easily interpret complex datasets and make informed decisions.

VII.IMPLEMENTATION

The implementation of D-Studio is kept simple yet robust, integrating modules for data ingestion, cleaning, analytics, and visualization into a unified workflow. Built on a layered architecture, it connects to diverse sources, preprocesses data with AI-driven cleaning, applies machine learning for insights, and delivers results through dashboards and cloud storage, ensuring scalability, security, and ease of use.

VIII. RESULTS AND DISCUSSIONS

The **Results and Discussion** section highlights the outcomes of implementing D-Studio and evaluates its performance across different stages of data analysis. The results demonstrate that the system successfully automates data ingestion, cleaning, and visualization, significantly reducing manual effort and processing time. Machine learning integration enhances accuracy in anomaly detection and predictive analytics, while interactive dashboards provide clear and actionable insights for both technical and non-technical users. In discussion, the strengths of D-Studio—such as scalability, usability, and secure cloud storage—are emphasized, along with its ability to bridge gaps found in existing tools. Limitations, including the need for continuous dataset updates and optimization for very large-scale data, are acknowledged, providing scope for future improvements. This section ultimately validates D-Studio as a reliable and efficient solution for end-to-end automated data analysis.

VIII. CONCLUSION

D-Studio demonstrates the potential of an integrated, automated platform for end-to-end data analysis. By combining modules for data collection, cleaning, visualization, and machine learning–driven analytics, it reduces manual effort while improving accuracy and efficiency. Its layered architecture, cloud-based storage, and intuitive dashboards make it accessible to both technical and non-technical users, ensuring scalability and reliability. The results validate D-Studio as a practical solution that bridges gaps in existing tools, offering transparency, usability, and real-time insights. Future enhancements can focus on expanding dataset coverage and optimizing performance for large-scale applications, further strengthening its role as a comprehensive data analysis framework.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to our mentors and colleagues for their valuable guidance and support throughout the development of D-Studio. We also acknowledge the contributions of open-source communities and dataset providers such as Kaggle, UCI Repository, and other platforms, whose resources greatly aided our research and implementation. Finally, we thank all those who encouraged and assisted us in completing this work successfully

REFERENCES

1. Xu, X., Li, D., Bi, J., & Moeckel, M. (2024). *AutoML based workflow for design of experiments (DOE) selection and benchmarking data acquisition strategies*. **Scientific Reports, Nature**.
2. Michelucci, U., & Venturini, F. (2024). *Best Practices for Machine Learning Experimentation in Scientific Applications*. **arXiv.org**.
3. Shivanna, A. (2024). *Framework for Implementing Experiment Tracking in Machine Learning Development*. **Journal of Scientific and Engineering Research**.
4. Côté, P.-O., Nikanjam, A., Abdelmutalab-Ali-Ahmed, N., Humeniuk, D., & Khomh, F. (2025). *Data Cleaning and Machine Learning: A Systematic Literature Review*. **arXiv.org**.
5. Kashid, D. (2025). *A Study on Data Cleaning Techniques for Large Datasets*. **International Journal of Innovative Research in Technology (IJIRT)**.
6. MDPI Research Team (2025). *Performance and Scalability of Data Cleaning and Preprocessing Tools*. **MDPI Journals**.
7. Hanumanthaiah, S. (2023). *Advancements in Data Visualization: A Comprehensive Literature Review*. **Independent Research**.
8. Lavanya, A., Sindhuja, S., Gaurav, L., & Ali, W. (2023). *A Comprehensive Review of Data Visualization Tools: Features, Strengths, and Weaknesses*. **International Journal of Computer Engineering in Research Trends (IJCERT)**.
9. IEEE Xplore Survey (2022). *A Comprehensive State-of-the-Art Survey on Data Visualization Tools*. **IEEE Xplore**.
10. Recapio Blog (2025). *12 Best AI Tools for Research to Use in 2025*. **Recapio Blog**.

11. Perkel, Z. (2025). *Top 10 AI Tools for Research in 2025: Features & Pricing*. **Julius AI Blog**.
12. Anara AI (2025). *AI Tools for Literature Review: Complete Guide*. **Anara AI for Science**

AI-Based Skin Disease Detection System Using Convolutional Neural Networks

Neha U
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
neha_u@ahalia.ac.in

Nandagopan M S
Computer Science and Engineering
Ahalia School of Engineering& Technology
Palakkad, Kerala, India
nandagopan_ms@ahalia.ac.in

Prathyush P Nambiar
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
prathyush_p@ahalia.ac.in

Muhammed Sahil K
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
Muhammed_sahil_k@ahalia..ac.in

Ms Anisree P G
Assistant Professor
Department of computer Science
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
Anisree.pg@ahalia.ac.in

***Abstract*— Skin diseases collectively affect over 900 million people globally, yet access to dermatological expertise remains severely limited in developing and rural regions. This paper presents the design, development, and evaluation of an AI-Based Skin Disease Detection System leveraging Convolutional Neural Networks (CNNs) for automated, objective classification of skin lesions from digital dermoscopic images. The system is built around the InceptionV3 architecture pre-trained on ImageNet and finetuned via a two-stage transfer learning strategy on the HAM10000 benchmark dataset, comprising 10,015 labeled dermoscopic images across seven disease categories. Class imbalance is addressed through class-weighted training, targeted augmentation, and oversampling. The fine-tuned model achieves 90.8% overall accuracy, macro F1-score of 0.887, and mean AUC-ROC of 0.964, exceeding all performance targets. The system is deployed through a Tkinter desktop GUI with PDF report generation, demonstrating that a student-level team using open-source tools can build a clinically meaningful AI diagnostic support system.**

Keywords: Convolutional Neural Networks, Skin Disease Detection, Deep Learning, Dermoscopy, HAM10000, Transfer Learning, InceptionV3, Medical Image Classification, TensorFlow, OpenCV, Explainable AI, Healthcare Accessibility.

I. Introduction

Skin diseases represent one of the largest non-fatal disease burdens globally, collectively affecting over 1.9 billion people. Among these, skin cancers—particularly melanoma—are directly life-threatening, causing over 57,000 deaths annually worldwide. Yet melanoma carries a five-year survival rate exceeding 98% when detected at Stage I, compared to approximately 25% at Stage IV [1]. This dramatic outcome differential underscores early and accurate detection as a clinical imperative.

The global shortage of dermatologists severely compounds this challenge. Dermatologist-to-population ratios range from 4 per 100,000 in high-income nations to fewer than 0.5 per 100,000 in rural India and sub-Saharan Africa [2]. The resulting access gap means many patients with potentially malignant lesions go undiagnosed, and general practitioner misdiagnosis rates for melanoma have been documented at 20–30% [3].

Convolutional Neural Networks (CNNs) have emerged as a transformative solution. The landmark 2017 study by Esteva et al. [4] demonstrated that a single CNN trained on over 129,000 clinical images could classify skin cancers at a level comparable to board-certified dermatologists. Subsequent benchmark challenges on HAM10000 [5] have driven rapid progress, with ensemble systems achieving AUC scores of 0.94–0.97.

This paper presents a complete, end-to-end AI-Based Skin Disease Detection System combining a high-performance InceptionV3 CNN classifier with a Tkinter desktop GUI, PDF report generation, and a comprehensive ethical framework—directly bridging the gap between academic model performance and clinical deployability.

The remainder of this paper is organized as follows. Section II reviews related work. Section III describes the proposed system. Section IV details implementation. Section V presents evaluation results. Section VI addresses ethical considerations. Section VII concludes.

II. Literature Survey

Research in AI-based dermatology spans CNN architectures, transfer learning, dataset curation, explainability, and clinical validation.

A. Landmark Clinical Validation Studies

Esteva et al. [4] trained a GoogLeNet Inception v3 CNN on 129,450 clinical images across 2,032 diseases and evaluated it against 21 board-certified dermatologists. The model achieved AUC of 0.96 for keratinocyte carcinoma and 0.94 for melanoma, matching the panel mean—conclusively establishing CNN-level clinical accuracy.

Brinker et al. [6] conducted the largest prospective comparison, pitting a CNN against 157 dermatologists across 12 hospitals. The CNN achieved higher AUC than 136 of 157 dermatologists. Han et al. [7] trained ResNet-152 on 19,398 clinical images, demonstrating that AI provides the greatest benefit to less experienced practitioners in under-resourced settings.

B. Dataset Contributions

Tschandl, Rosendahl, and Kittler [5] created HAM10000—10,015 dermoscopic images collected over 20 years, with diagnoses confirmed by histopathology, followup, expert consensus, or confocal microscopy. HAM10000 has become the standard benchmark, democratizing AI dermatology research globally.

C. Architecture and Transfer Learning

Transfer learning from ImageNet-pretrained models dominates skin lesion classification. Szegedy et al. [8] introduced InceptionV3, which employs factorized convolutions and multi-scale Inception modules (parallel 1×1, 3×3, and 5×5 filters) for parameter-efficient feature extraction. Fine-tuned InceptionV3 outperforms custom CNNs trained from scratch by 5–15 percentage points on dermoscopy datasets. Tan and Le [9] demonstrated EfficientNet-B4 reaching 91.4% on HAM10000 via compound scaling.

D. Explainability

Selvaraju et al. [10] introduced Grad-CAM, which generates visual explanations by computing gradients of the class score with respect to the last convolutional layer. Applied to dermoscopic classification, Grad-CAM heatmaps highlight clinically meaningful features such as lesion borders and color irregularities. Integration of Grad-CAM into deployed interfaces remains an open gap addressed in future work.

E. Research Gaps

Three significant gaps persist: (i) no prior work packages a full GUI application with PDF report generation accessible to non-technical users; (ii) explainability is rarely integrated into deployed interfaces; and (iii) few systems address deployment constraints in Indian and developing-nation healthcare contexts. This work directly addresses the first gap.

III. Proposed System

A. System Overview

The system is a standalone desktop application built around a five-layer architecture separating user interaction, workflow orchestration, image preprocessing, CNN inference, and data management (Table I). It accepts JPEG, PNG, or BMP images and produces: predicted disease class, confidence score, top-3 predictions, clinical description, risk level, recommended action, and optional PDF report.

TABLE I. *System Architecture Layers*

Layer	Components	Technology	Responsibility
Presentation	Main Window, Result Panel	Tkinter (ttk)	User interaction & result display
App Logic	Session Manager, Controller	Python	Pipeline orchestration
Preprocessing	Validator, Resizer, Normalizer	OpenCV, NumPy	Image standardization
Inference	Model Loader, Engine, Postproc.	TensorFlow / Keras	CNN classification
Data	Disease DB, Report Generator	Python, ReportLab	Data management & reporting

B. Disease Classification Scope

The system classifies lesions into the seven HAM10000 categories shown in Table II, spanning benign, pre-malignant, and malignant presentations. Risk stratification drives colorcoded display and referral urgency in the GUI.

TABLE II. *HAM10000 Disease Categories*

Code	Disease	Count	%	Risk
nv	Melanocytic Nevi	6,705	66.95%	Low (Benign)
mel	Melanoma	1,113	11.11%	HIGH (Malignant)
bkl	Benign Keratosis	1,099	10.97%	Low (Benign)
bcc	Basal Cell Carcinoma	514	5.13%	HIGH (Malignant)
akiec	Actinic Keratoses	327	3.26%	Medium (Premalignant)
vasc	Vascular Lesions	142	1.42%	Low (Benign)
df	Dermatofibroma	115	1.15%	Low (Benign)

C. CNN Architecture: InceptionV3

InceptionV3 was selected after benchmarking VGG16, ResNet50, InceptionResNetV2, and EfficientNet-B4 on HAM10000 for accuracy, parameter efficiency, training stability, and CPU inference speed. At 48 layers and ~23.9M parameters, InceptionV3 provides the best balance. Its parallel multi-scale modules (1×1, 3×3, 5×5 filters) capture features from pixel-level variation to centimeter-scale lesion morphology simultaneously.

Custom head: GlobalAveragePooling2D → BatchNorm + Dense(512, ReLU) + Dropout(0.5) → BatchNorm +

Dense(256, ReLU) + Dropout(0.3) → Dense(7, Softmax). Total trainable parameters in Stage 2: ~7.3M.

D. Two-Stage Training Strategy

Stage 1 freezes all InceptionV3 base weights and trains only the classification head (Adam, LR=1×10^{-3}, 30 epochs), achieving 78.9% test accuracy. Stage 2 unfreezes the top 100 layers (Adam, LR=1×10^{-5}, 50 epochs, ReduceLROnPlateau), yielding 90.8%—a 12-point improvement. Class-weighted cross-entropy loss addresses the 66:1 imbalance between Melanocytic Nevi and Dermatofibroma. *E. Preprocessing and Augmentation*

Pipeline: format validation (JPEG/PNG/BMP, min 50×50 px), BGR-to-RGB conversion, Gaussian denoising (3×3, σ=0.5), bicubic resize to 224×224, normalization to [0, 1], batch expansion. Augmentation: rotation ±20°, shift ±15%, zoom ±20%, horizontal/vertical flip, brightness [0.8, 1.2]. The two smallest classes (vasc: 142, df: 115) are additionally oversampled.

IV. Implementation

A. Software Stack

Incremental development methodology, four increments, 16 weeks. Stack: Python 3.8+, TensorFlow 2.x/Keras, OpenCV 4.7+, NumPy 1.23+, Pandas 1.5+, Scikit-learn 1.2+, Matplotlib 3.6+, Pillow 9.4+, ReportLab 3.6+.

B. Key Modules **preprocess.py:** Sequential preprocessing class returning normalized tensor (1, 224, 224, 3). Each step validates its input and raises a descriptive exception on failure.

model_builder.py: Constructs InceptionV3 with custom head; orchestrates two-stage training with compute_class_weight('balanced') and ModelCheckpoint/EarlyStopping callbacks.

inference.py: InferenceEngine loads SavedModel on init; exposes predict(image_path) returning {class_code, class_name, confidence, top3_predictions, risk_level, ui_color}.

gui.py: SkinDiseaseApp (900×650 px); inference runs in a background thread with thread-safe delivery via after(). Risk-coded ttk.Progressbar shows confidence.

report_generator.py: ReportLab PDF: timestamp, image thumbnail, prediction, top-3, clinical description, recommended action, mandatory disclaimer: "This report is NOT a clinical diagnosis." *C. GUI Screen Flow*

State machine: LAUNCH → IDLE → IMAGE_LOADED → PROCESSING → RESULTS_DISPLAYED → REPORT_GENERATED → IDLE. Invalid transitions are caught with descriptive error dialogs.

V. Evaluation and Discussion

A. Testing Strategy

Three levels: unit testing (pytest, 41 cases, 96.8% line coverage), integration testing (7 end-to-end scenarios), and model validation on the stratified held-out test set (15% of HAM10000, n=1,502) augmented by 5-fold cross-validation on the full dataset.

B. Overall Model Performance

Table III presents key performance metrics. All performance targets defined in the system requirements are met or exceeded.

TABLE III. *Model Performance Across Training Stages*

Metric	Stage 1	Stage 2 (FT)	5-Fold CV	Target
Test Accuracy	78.9%	90.8%	90.6% ±1.3%	≥90%
Val. Accuracy	79.4%	91.8%	91.2% ±1.1%	≥90%
Macro F1-Score	0.671	0.887	0.882 ±0.018	≥0.85
AUC-ROC (Mean)	0.912	0.964	0.964 ±0.012	≥0.95
Inference (CPU)	1.2 s	1.4 s	1.4 s avg	≤5 s

C. Per-Class Performance

Table IV presents per-class results. Melanoma recall of 86.4% exceeds the ≥85% clinical safety target, minimizing dangerous false negatives for the most lethal class. Lowest F1 (83.8%) is for Actinic Keratoses due to visual overlap with Benign Keratosis—a well-documented diagnostic ambiguity.

D. Comparison with Related Works

Table V contextualizes results against published HAM10000 baselines. The fine-tuned InceptionV3 (90.8%, F1=0.887) is consistent with the ISIC 2018 challenge range, while uniquely combining competitive CNN performance with a deployed GUI application, PDF reporting, and ethical documentation.

TABLE V. *Comparison with Published HAM10000 Results*

System	Architecture	Acc.	F1 / AUC
This Work	InceptionV3 Fine-Tuned	**90.8%**	F1=0.887, AUC=0.964
Tschandl 2018 [5]	*Multiple CNNs*	83%	AUC 0.92–0.95
ISIC 2018 Winner	*CNN Ensemble*	~90%	AUC ~0.97
Hamida 2021 [11]	*ResNet50 FineTuned*	88.6%	F1=0.863
Xue 2022 [12]	*EfficientNet-B4*	91.4%	F1=0.901

Baseline (scratch)	*Custom 5-layer CNN*	71.2%	F1=0.581

E. Unit and Integration Testing

Unit tests cover: format validation, normalization range [0,1], 224×224 resize, small-image rejection (<50×50 px), grayscale-to-RGB conversion, output shape (1,7), probability sum within [0.9999, 1.0001], CPU inference time <5 s. All 41 cases pass (96.8% coverage). Seven integration scenarios pass including 8MP image handling and graceful error display for invalid file types.

Four ethical dimensions are addressed. First, demographic fairness: HAM10000 over-represents Fitzpatrick skin types I–III from European clinical sites. The model may show reduced accuracy for darker skin tones (types IV–VI), a population already underserved. Stratified evaluation and targeted augmentation are priorities for future work. Second, patient privacy: the system operates entirely locally with zero external data transmission and no persistent storage without explicit user action. Third, appropriate scope: the system is a preliminary screening tool, not a clinical diagnosis device; all reports carry a mandatory disclaimer. Fourth, informed consent: users are informed of system capabilities, data limitations, and follow-up steps through onboarding documentation.

VI. Conclusion

This paper presented a complete AI-Based Skin Disease Detection System achieving 90.8% classification accuracy, macro F1-score of 0.887, and mean AUC-ROC of 0.964 on the HAM10000 held-out test set—exceeding all performance targets. Melanoma recall of 86.4% satisfies the clinical safety requirement. The system is the first in this domain to combine competitive CNN performance with a fully deployed GUI application, PDF report generation, and a documented ethical framework.

Five key contributions: (i) complete deployable application from raw image to PDF report; (ii) two-stage transfer learning yielding a 12-point accuracy improvement; (iii) modular, unit-tested Python codebase; (iv) rigorous multi-metric evaluation with 5-fold cross-validation; (v) comprehensive ethical treatment of fairness, privacy, scope, and informed consent.

Acknowledgment

The authors thank Ms. Anisree P G and the Department of Computer Science and Engineering, Ahalia School of Engineering & Technology, for their guidance and support. The HAM10000 dataset is acknowledged as provided by the International Skin Imaging Collaboration (ISIC).

References

[1] WHO Global Health Observatory. (2023). Skin conditions: Key facts. World Health Organization.
[2] American Academy of Dermatology. (2022). Dermatology workforce shortage fact sheet. AAD Policy Report.
[3] Vestergaard, M. E., et al. (2008). Dermoscopy vs. naked eye for primary melanoma diagnosis: a meta-analysis. British Journal of Dermatology, 159(3), 669–676.
[4] Esteva, A., Kuprel, B., Novoa, R. A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118.
[5] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset. Scientific Data, 5(1), 180161.
[6] Brinker, T. J., et al. (2019). Deep neural networks are superior to dermatologists in melanoma image classification. European Journal of Cancer, 119, 11–17.
[7] Han, S. S., et al. (2018). Classification of clinical images for benign and malignant cutaneous tumors. Journal of Investigative Dermatology, 138(7), 1529–1538.
[8] Szegedy, C., et al. (2016). Rethinking the Inception architecture for computer vision. Proceedings of CVPR 2016, 2818–2826.
[9] Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for CNNs. Proceedings of ICML 2019, 6105–6114.
[10] Selvaraju, R. R., et al. (2017). Grad-CAM: Visual explanations from deep networks. Proceedings of ICCV 2017, 618–626. [11] Hamida, A. B., et al. (2021). Deep learning for melanoma detection. Computer Methods and Programs in Biomedicine, 208, 106284.
[12] Xue, C., et al. (2022). EfficientNet-based skin lesion classification on HAM10000. Journal of Digital Imaging, 35(2), 312–322.
[13] Codella, N. C. F., et al. (2018). Skin lesion analysis toward melanoma detection: ISIC 2017. IBM J. Res. Dev., 63(4/5), 1–15.
[14] He, K., et al. (2016). Deep residual learning for image recognition. Proceedings of CVPR 2016, 770–778.
[15] Abadi, M., et al. (2016). TensorFlow: A system for large-scale machine learning. Proceedings of OSDI 2016, 265–283.

Web App Security Analyzer with AI Framework

Nivedh V
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
nivedh_v@ahalia.ac.in

Milan Krishna
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
milan_krishna@ahalia.ac.in

Kassim Noushin N
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
kassim_noushin_n@ahalia.ac.in

Kavitha S
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
kavitha.s@ahalia.ac.in

***Abstract* —Web applications have become an essential part of modern digital services, but this widespread use also makes them frequent targets of cyber-attacks. Detecting security vulnerabilities early is crucial to protect sensitive data and ensure the reliability of these systems. To address this, the project introduces a Flutter-Based Web Application Vulnerability Scanning Framework that automates the process of finding and analysing security weaknesses in web applications. The framework consists of a Flutter-based front end that provides a user-friendly interface for initiating scans and viewing results, and a Python-based backend responsible for executing vulnerability assessment tasks. The system integrates widely used security tools such as Nmap, Nikto, and the Metasploit Framework to perform thorough scanning. Nmap is used for network and port scanning, Nikto identifies web server vulnerabilities and misconfigurations, and Metasploit helps in enumerating and validating exploitable endpoints. By integrating AI-based API, the WebApp Security Analyzer moves beyond traditional rule-based scanning to become an intelligent, context-aware security tool. The AI layer enables it to classify vulnerabilities, provide human-readable explanations, suggest fixes, and prioritize risks, making it a powerful tool for developers, students, and security enthusiasts alike. After scanning, the framework collects and consolidates the result to highlight open ports, vulnerable services, and insecure web endpoints. By automating these tasks and presenting the findings through a single platform, the system reduces manual effort and improves the efficiency and accuracy of security assessments. This framework doesn't require any terminal commands or coding. Everything is done through a user-friendly GUI, where users can start scans and view results with just a few clicks. The backend handles all the technical work automatically, making vulnerability scanning simple and accessible for everyone. This framework can be effectively used by developers, system administrators, and security professionals to proactively enhance web application security**.

***Keywords*—Web application security, Vulnerability scanning, Flutter, AI-assisted analysis, Nmap, Nikto, Metasploit, GUI automation.**

I. INTRODUCTION

Web applications are essential in modern digital services but are frequent targets of cyber-attacks. Ensuring their security is critical to protect sensitive data and maintain system reliability. Traditional vulnerability assessments are often manual, time-consuming, and error-prone.

To address this, we propose a Flutter-Based Web Application Vulnerability Scanning Framework that automates the detection and analysis of security weaknesses. The system integrates Nmap for network scanning, Nikto for web server vulnerabilities, and the Metasploit Framework for validating exploitable endpoints. An AI module classifies vulnerabilities, provides explanations, suggests fixes, and prioritizes risks.

A graphical user interface (GUI) allows users to run scans and view results without coding. Built with Flutter for the front end and Python for the backend, the framework consolidates results, including open ports, vulnerable services, and insecure endpoints. By combining automated scanning, AI analysis, and an intuitive interface, it enhances the efficiency, accuracy, and usability of web application security assessments.

II. LITERATURE SURVEY

Web application security is a critical concern as cyber-attacks continue to target sensitive data and services. Several research studies have evaluated automated vulnerability scanning tools to identify strengths, limitations, and best practices for securing web applications. By combining multiple scanning methods, these studies suggest that a hybrid automated framework can enhance detection coverage and efficiency.

Swetha et al. (2024) conducted an evaluation of popular vulnerability management tools, including Nessus, Acunetix, and Nikto. The study compared detection capabilities, scanning speed, and usability, highlighting the strengths and weaknesses of each tool. While Nessus and Acunetix excelled at comprehensive scanning and reporting, Nikto was faster but limited in coverage. The research concluded that integrating multiple scanners can provide a more effective and reliable automated framework for vulnerability assessment, ensuring better protection against complex web threats [1].

Gordon Lyon (2009) presented a comprehensive guide on network discovery and security scanning using *Nmap*. The work focuses on techniques for host discovery, port scanning, service detection, and vulnerability identification. The study highlights that Nmap is highly effective for mapping network structures and identifying open ports, making it a powerful tool for initial reconnaissance. The findings emphasize its importance in penetration testing and network security assessment, particularly for efficiently identifying potential attack surfaces [2].

Zogaj et al. (2025) conducted a statistical assessment of multiple automated web vulnerability scanners. Their research quantified the consistency, detection rates, and efficiency of different tools across various web applications. The study highlighted that no single scanner achieves full coverage consistently, underlining the need for multi-tool frameworks to reduce false negatives and improve vulnerability detection reliability [3].

Patel, Mehta, and Shah (2024) presented a comparative study of automated web application vulnerability scanners. The research analysed coverage, false positives, and ease of use, identifying gaps in current automated scanning approaches. Recommendations from this study suggest combining scanners with different detection mechanisms to achieve comprehensive vulnerability coverage and improve overall system security [4].

Al-Mutairi, Khan, and Alotaibi (2023) proposed a GUI-based vulnerability assessment system designed for non-expert users. The system aimed to simplify vulnerability scanning by providing an intuitive interface while leveraging existing automated scanners. The study emphasized usability and accessibility, demonstrating that effective scanning frameworks should balance technical accuracy with ease of operation for broader adoption [5].

Building on these studies, the proposed project introduces a Flutter-based web application vulnerability scanning framework that integrates multiple scanning tools to automate the detection and reporting process. Unlike prior work that focused solely on individual tool evaluation or interface usability, this system combines the detection strengths of Nessus, Acunetix, Nikto, and Burp Suite with a user-friendly dashboard. Additionally, the framework incorporates automated report generation and AI-assisted insights to prioritize vulnerabilities, offering an end-to-end solution for effective and efficient web application security assessment.

III. PROPOSED MODEL

A. Web Application Vulnerability Scanning Challenges

Web applications face increasing security risks due to complex architectures and diverse attack vectors. Current challenges in vulnerability assessment include:

Limited coverage by individual scanners: No single tool detects all vulnerabilities, leaving potential gaps.

High false positives and inconsistent results: Different tools may report conflicting findings, complicating prioritization.

Technical complexity for users: Many automated scanners require cybersecurity expertise to interpret results accurately.

Time-consuming reporting and remediation: Manual consolidation of scan data slows down response and mitigation.

These challenges often lead to missed vulnerabilities, delayed security updates, and higher risk of exploitation. To address these issues, the proposed approach introduces a Flutter-based automated web application vulnerability scanning framework that ensures:

- Comprehensive multi-tool scanning integrating Nmap, Nikto, and Metasploit.
- AI-driven prioritization of vulnerabilities based on severity, exploitability, and business impact.
- Intuitive dashboard for real-time monitoring and result visualization.
- Seamless team collaboration through notifications, exportable reports, and interactive result sharing.
- Adaptive scanning configuration based on application type, environment, and risk level.

B. Key Components of the Model

The proposed framework integrates multiple modules to automate vulnerability scanning, analysis, and reporting, providing an end-to-end solution. The main components are described below:

i. Target Analysis Module

The system begins by analyzing the target web application, collecting details such as server type, open ports, active services, and application endpoints. This module generates a comprehensive application profile to guide scanning strategy.

ii. Multi-Tool Scanning Engine

The system integrates multiple security tools—**Nmap**, **Nikto**, and **Metasploit**—to perform efficient and comprehensive vulnerability assessment.

Nmap is used for network discovery and port scanning, helping identify active hosts, open ports, and services running on the target system. Nikto focuses on web server scanning, detecting common vulnerabilities such as outdated software, misconfigurations, and insecure files. Metasploit is utilized for advanced penetration testing, enabling the system to validate vulnerabilities by simulating real-world attacks and exploits.

By combining these tools, the engine ensures broader coverage, improved accuracy, and faster detection of security issues through parallel and coordinated scanning.

C. AI-Based Vulnerability Prioritization

The framework uses AI algorithms to classify vulnerabilities by severity and likelihood of exploitation. Features such as CVSS scores, past exploit trends, and application context are analysed to generate a risk-based priority list. This ensures that critical vulnerabilities are addressed first, optimizing remediation efforts.

D. Consolidated Reporting Module

Scan results from multiple tools are automatically aggregated and normalized into a unified report. The system removes duplicate entries, highlights critical issues, and provides actionable recommendations. Reports can be exported in PDF or CSV format for stakeholders.

E. Real-Time Notification and Collaboration

The framework includes a notification system to alert developers of new findings instantly. Real-time chat and team collaboration tools are integrated for discussion, task assignment, and progress tracking, ensuring faster vulnerability resolution.

F. Adaptive Scanning and Continuous Learning

The framework continuously updates its scanning strategy based on previous scan results and observed vulnerabilities. AI models learn patterns of recurring vulnerabilities, optimizing future scans and reducing false positives over time.

By combining automated multi-tool scanning, AI-assisted prioritization, and an intuitive dashboard, the proposed system offers a dynamic and efficient web application security platform, ensuring comprehensive coverage, rapid detection, and actionable insights for developers and security teams.

IV. IMPLEMENTATION

The proposed framework integrates multiple tools, AI-based analysis, and a user-friendly interface. Key implementation aspects include backend, frontend, database, AI model integration, and deployment strategies for scalable performance.

A. Backend Development:

Framework: Django (Python) is used for building the backend, leveraging Django ORM for database operations and scaling.

Main Components:

- User Authentication & Profiles: Secure login using Django's authentication system and JWT tokens. User profiles store target applications, scan history, and preferences.
- Multi-Tool Scanning Engine: Coordinates Nmap, Nikto, and Metasploit scans. Each tool's results are collected and stored for analysis.
- AI-Based Vulnerability Prioritization: AI algorithms rank detected vulnerabilities based on severity, exploitability, and criticality using a hybrid scoring system. This ensures developers focus on the most important issues first.
- Real-Time Notifications & Reporting: Django Channels and WebSockets are used to provide real-time alerts and report updates. Users can receive instant notifications when scans are complete.
- API Development: RESTful APIs via Django Rest Framework (DRF) provide endpoints for initiating scans, retrieving results, and managing users. API throttling ensures safe and controlled access.

B. Frontend Development:

Technologies: HTML, CSS, JavaScript, and Flutter for building a responsive and interactive interface.

Features:

- Dashboard: Displays scan results, vulnerability priorities, and progress of ongoing scans.
- Visualization: Interactive charts and tables highlight open ports, vulnerable services, and risk levels.
- Real-Time Alerts: Notifications for completed scans, critical vulnerabilities, and recommendations.
- Export & Reporting: Users can download reports in PDF or CSV format for team review.

C. Database Design (PostgreSQL)

- Data Management: The PostgreSQL database is designed for efficiency and scalability.
- User Profile Table: Stores application details, scan preferences, and historical scan data.
- Scan Results Table: Stores all vulnerability scan outputs from different tools.
- Vulnerability Prioritization Table: Maintains AI-calculated risk scores for each finding.
- Notifications Table: Tracks alerts sent to users for completed scans or critical vulnerabilities.

Database indexing and caching are implemented to support faster queries for large datasets.

D. Scalability & Deployment

- Cloud Hosting: The system is deployed on AWS or Railway to ensure high availability and elasticity.
- Load Balancing: Nginx and Gunicorn handle multiple user requests concurrently.
- Containerization: Docker is used for consistent environment setup across development and production.
- Database Optimization: Indexing and caching improve query performance for large scan datasets.

- Security & Maintenance: Regular updates, backup strategies, and secure API endpoints ensure a reliable and secure platform for web application vulnerability scanning.

V. EVALUATION

The assessment of the proposed system is based on vulnerability detection accuracy, scan efficiency, user experience, and system performance. Testing was conducted using multiple real-world web applications, benchmarked against existing vulnerability scanners, and verified for AI prioritization performance.

A. Vulnerability Detection Accuracy & AI Performance

The multi-tool scanning engine was tested across multiple web applications containing known vulnerabilities such as SQL Injection, Cross-Site Scripting (XSS), and security misconfigurations.

Detection Capability: The integrated approach demonstrated improved vulnerability coverage compared to individual tools. Experimental observations indicate that combining multiple scanners helps identify a broader range of vulnerabilities, as different tools specialize in different detection techniques. However, exact accuracy values may vary depending on the target application and test conditions.

AI-Based Prioritization: The AI-based risk scoring mechanism was evaluated by comparing its severity ranking with known vulnerability classifications. The system showed effective prioritization of high-risk vulnerabilities in most test cases, enabling faster identification of critical issues. However, the performance of the AI model depends on the quality and diversity of training data.

False Positives Reduction: By aggregating scan results and eliminating duplicate findings, the system reduced redundant and irrelevant alerts. This resulted in a noticeable improvement in report clarity compared to standalone tool outputs, though the degree of reduction may vary across different environments.

Scan Efficiency: The average scan duration for medium-sized applications ranged between **5 to 8 minutes**, depending on network conditions, system load, and target complexity. Parallel execution of multiple tools contributed to improved efficiency.

Reporting & Usability: User feedback from a limited evaluation group indicated that the unified dashboard and AI-generated recommendations improved ease of use and reduced analysis time. However, a larger user study is required for statistically significant conclusions.

Security of Results: All scan data and reports are securely stored in **PostgreSQL**, with communication protected via HTTPS. Sensitive information is encrypted using AES, ensuring confidentiality and integrity of scan results.

V. CONCLUSION AND FUTURE WORK

A. Conclusion

The proposed Flutter-based Web Application Vulnerability Scanning Framework provides a comprehensive, AI-assisted solution for detecting and prioritizing security issues. By combining multiple automated scanners with AI-based risk analysis, the system ensures higher coverage and accuracy than individual tools.

Key contributions of the framework include:

- Comprehensive Vulnerability Detection: Integrates Nmap, Nikto, and Metasploit to detect a wide range of vulnerabilities.
- AI-Driven Prioritization: Assigns severity scores to vulnerabilities, allowing developers to focus on critical issues first.
- User-Friendly Dashboard: Consolidates scan results, provides interactive visualizations, and simplifies reporting.
- Real-Time Alerts and Collaboration: Notifications, chat, and report-sharing enable fast response and coordinated remediation.
- Scalability & Performance: Cloud deployment with load balancing and containerization ensures the system can handle multiple scans and users simultaneously without significant latency.

Compared to conventional single-tool scanning approaches, the proposed system offers a more accurate, efficient, and actionable security assessment, supporting both technical and non-technical users in improving web application security.

B. Future Work

The system can be further enhanced to improve scanning effectiveness, user experience, and automation:

Reinforcement Learning (RL): Use RL to optimize AI-based vulnerability prioritization over time based on past scan outcomes.

Natural Language Processing (NLP): Apply NLP to automatically generate readable vulnerability descriptions, summaries, and remediation suggestions for developers.

Expanded Scanner Integration: Include additional tools and emerging scanners to improve detection coverage and adaptability to new vulnerabilities.

Gamification & Reporting Enhancements: Provide risk dashboards, progress tracking, and interactive visualizations for security teams.

Cloud & Global Deployment: Scale system hosting on AWS, GCP, or Azure to support large organizations with multiple concurrent scans.

API Integrations: Enable seamless integration with DevOps pipelines, CI/CD tools, and popular development platforms for automated scanning.

These improvements aim to create a robust, AI-powered, and user-friendly web application security platform capable of delivering timely, actionable insights while reducing manual effort and improving overall organizational security posture.

VII. REFERENCES

[1] B. Swetha, N. R. K. Susmitha, J. Thirulogaveni, and S. Sruthi, "Evaluation of vulnerability management tools: Nessus, Acunetix, and Nikto," *Journal of Cybersecurity Research*, vol. 12, no. 1, pp. 45–58, 2024.

[2] G. Lyon, *Nmap Network Scanning: The Official Guide to Network Discovery and Security Scanning*. Sunnyvale, CA, USA: Insecure.Com LLC, 2009.

[3] G. Zogaj, F. Ismaili, E. Idrizi, and A. Luma, "Statistical analysis of unique web application vulnerabilities: a quantitative assessment of scanning tool efficiency," *SEEU Review*, vol. 20, no. 1, pp. 136–152, 2025.

[4] R. Patel, K. Mehta, and D. Shah, "Comparative study of automated web application vulnerability scanners," *International Journal of Cyber Security and Digital Forensics*, vol. 13, no. 2, pp. 45–54, 2024.

[5] S. Al-Mutairi, A. Khan, and F. Alotaibi, "GUI-based web vulnerability assessment system for non-expert users," *Journal of Information Security and Applications*, vol. 75, no. 1, pp. 102–110, 2023.

[6] S. Razaulla, M. Pasha, and M. U. Farooq, "Integration of machine learning in security tools: challenges, issues and trends," in *Machine Learning and Internet of Things for Societal Issues*, Singapore: Springer Nature, 2022, pp. 23–34.

[7] R. Tiwari, "The integration of AI and machine learning in automated vulnerability assessment and its potential to improve web application security," *International Journal of Scientific Research in Engineering and Management*, vol. 7, no. 2, pp. 101–110, 2023.

[8] H. S. Nwana, "Intelligent vulnerability assessment systems: an overview," *Artificial Intelligence Review*, vol. 4, no. 4, pp. 251–277, 1990.

[9] J. Shahid, M. K. Hameed, I. T. Javed, K. N. Qureshi, M. Ali, and N. Crespi, "A comparative study of web application security parameters: current trends and future directions," *Applied Sciences*, vol. 12, no. 8, p. 4077, 2022.

[10] F. Holík, J. Horalek, O. Marik, S. Neradova, and S. Zitta, "Effective penetration testing with Metasploit framework and methodologies," in *Proc. 2014 IEEE 15th International Symposium on Computational Intelligence and Informatics (CINTI)*, Budapest, Hungary, Nov. 2014, pp. 237–242.

[11] S. Alazmi and D. C. De Leon, "A systematic literature review on the characteristics and effectiveness of web application vulnerability scanners," *IEEE Access*, vol. 10, 2022.

Web-Based System for Detecting Leaked Personal Data Using Breach Checker

Anirudh G Menon
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
anirudh_gmenon@ahalia.ac.in

Sreedev K A
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
sreedev_ka@ahalia.ac.in

Hidhun Babu
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
hidhunbabu_v@ahalia.ac.in

Ms. Rishika
Computer Science and Engineering
Ahalia School of Engineering & Technology
Palakkad, Kerala, India
rishika.kk@ahalia.ac.in

***Abstract*—With the rapid expansion of digital platforms and online services, users frequently share sensitive personal information such as email addresses and phone numbers. This increased data sharing has resulted in a higher risk of data breaches and unauthorized exposure of personal information. Many users remain unaware that their credentials may already exist in leaked databases obtained from compromised websites or cyberattacks. This paper proposes a web-based system for detecting leaked personal data using API-based and dataset-driven techniques. The system allows users to check whether their email address or phone number has been exposed in known data breaches. The application integrates publicly available breach datasets and secure APIs to identify potential data leaks efficiently. The frontend of the system is developed using HTML, CSS, and React to provide an interactive interface, while the backend uses Python and JavaScript to process requests and analyze breach data. The proposed solution improves user awareness about digital privacy by providing fast and reliable breach detection without storing sensitive user information. The system encourages users to take preventive actions such as changing passwords, enabling two-factor authentication, and strengthening their cybersecurity practices.**

Keywords—**Data Breach Detection, Cybersecurity, Web Application Security, Data Privacy, API-Based Detection, Personal Data Protection**

I. Introduction

The rapid growth of digital platforms and online services has significantly increased the amount of personal data shared over the internet. Users frequently provide sensitive information such as email addresses, phone numbers, usernames, and passwords while accessing websites, mobile applications, and cloud-based systems. Although these platforms offer convenience and connectivity, they also introduce serious security risks related to data privacy and protection.

In recent years, there has been a substantial rise in data breaches affecting organizations across various sectors. Cyber attackers often exploit system vulnerabilities to gain unauthorized access to databases, leading to the exposure of confidential user information. Once compromised, this data is frequently distributed or sold on public forums, dark web platforms, or breach repositories. As a result, individuals may unknowingly have their personal information exposed, making them vulnerable to identity theft, phishing attacks, and other cyber threats.

Given this growing concern, it has become essential for users to have a simple and reliable method to verify whether their personal data has been leaked. However, most users lack the technical knowledge or tools required to manually check such exposures across multiple sources. This creates a need for an accessible system that can automate the process of detecting data leaks and present the results in an understandable manner.

To address this issue, this paper proposes a web-based data leak detection system that enables users to check whether their personal information appears in known breach datasets. The system utilizes modern web technologies and integrates external APIs and publicly available breach intelligence sources to perform real-time checks. By processing user inputs securely and efficiently, the system provides quick feedback regarding potential data exposure along with basic security awareness.

The proposed system focuses on simplicity, speed, and user privacy, ensuring that sensitive inputs are handled securely without permanent storage. Through this approach, the system aims to enhance cybersecurity awareness among users and provide a practical solution for monitoring personal data exposure in today's increasingly connected digital environment.

II. Literature Survey

The increasing number of cyber attacks and data breaches has highlighted the need for systems that help users identify whether their personal information has been exposed. Several studies and cybersecurity reports have explored the causes of data breaches, their impact, and methods for detection and prevention.

Troy Hunt (2013) introduced the platform Have I Been Pwned, which allows users to check whether their email addresses or credentials have been exposed in publicly known data breaches. The platform aggregates information from multiple breach datasets and provides an easy-to-use interface for users. This work demonstrates the importance of publicly accessible tools in improving user awareness and enabling individuals to take preventive measures against data misuse.

The OWASP Foundation published the OWASP Top 10, which identifies the most critical vulnerabilities in web applications that often lead to data breaches. These include injection attacks, broken authentication, sensitive data exposure, and security misconfigurations. The report emphasizes the importance of secure coding practices and continuous monitoring to prevent unauthorized access to user data.

Dieter Gollmann (2011) provides a comprehensive overview of computer security principles, including data protection techniques, cryptographic methods, and network security strategies. The study highlights the importance of safeguarding sensitive information and detecting vulnerabilities that may result in data leaks.

Sasha Romanosky, Rahul Telang, and Alessandro Acquisti (2011) analyzed the impact of data breach disclosure laws. Their findings suggest that mandatory reporting of breaches encourages organizations to strengthen their security practices while also enabling users to take timely preventive actions.

The Verizon Data Breach Investigations Report provides a detailed analysis of global cybersecurity incidents. The report highlights that compromised credentials remain one of the leading causes of data breaches, reinforcing the need for systems that can detect exposed user data in leaked datasets.

Ramesh Behl and Kavita Behl (2017) discuss the growing threats of cyber attacks and their impact on digital systems. Their work emphasizes the increasing vulnerability of online platforms and the importance of awareness and detection mechanisms to protect user data.

William Stallings and Lawrie Brown (2018) examine modern security frameworks and defense mechanisms used in protecting information systems. Their research underlines the need for automated tools and efficient systems to mitigate risks associated with data breaches.

Based on these studies, it is evident that detecting compromised credentials and improving cybersecurity awareness are critical for protecting personal data. The proposed web-based system for detecting leaked personal data builds upon these concepts by providing a user-friendly platform that enables individuals to verify whether their credentials have been exposed in known breach datasets using API-based and publicly available data sources.

A comparative analysis of the features supported by existing industry-standard tools and our proposed solution is summarized in Table I, highlighting the unique detection capabilities integrated into this system.

TABLE I
COMPARISON OF DETECTION CAPABILITIES

Feature	Existing Platforms	Proposed System
Email Breach Detection	Yes	**Yes**
Password K-Anonymity	Yes	**Yes**
Username OSINT Check	No	**Yes**
Mobile Number Verification	No	**Yes**
Real-time API Integration	Yes	**Yes**

III. PROPOSED MODEL

The proposed system is a web-based application for detecting leaked personal data, designed to help users verify whether their sensitive information—such as email addresses, phone numbers, passwords, or usernames—has been exposed in publicly available data breaches. The system focuses on providing a fast, reliable, and user-friendly solution by integrating modern web technologies with publicly accessible breach intelligence services and APIs.

The architecture of the proposed system follows a modular and layered design, consisting of three main components: the User Interface Layer, the Backend Processing Layer, and the External Data Sources Layer. The user interface is developed using standard web technologies and allows users to input their data easily through a browser. The backend, implemented using FastAPI, handles request processing, validation, and communication with external services. The external layer consists of breach databases, OSINT sources, and third-party APIs used to retrieve relevant information about potential data leaks.

When a user submits input such as an email address, phone number, password, or username, the request is sent to the backend server. The system first performs input validation and sanitization to ensure correctness and security. Once validated, the backend processes the request by interacting with different modules. For email and mobile number checks, the system queries local breach datasets and external APIs. For password verification, the system utilizes secure API-based methods (such as k-anonymity techniques) to check whether the password has appeared in known breaches without exposing the actual password. For username analysis, the system performs a lightweight OSINT check across multiple platforms to determine its presence.

The system retrieves relevant results from these sources, processes the data, and generates a structured response. The results are then returned to the user through the web interface in real time, providing clear information about whether their data has been compromised along with basic security recommendations.

This architecture ensures efficient data processing, quick response time, and reliable detection without relying on complex machine learning models. Additionally, the system maintains user privacy and security by avoiding permanent storage of sensitive inputs and ensuring that all communications are handled securely. The modular design also allows easy scalability and future enhancements, making the system suitable for real-world cybersecurity awareness and personal data protection applications. Fig. 1 shows the system architecture and workflow of the proposed model.

A. User Interface Layer

The User Interface Layer provides the front-end environment through which users interact with the system. It is developed using standard web technologies such as HTML, CSS, and JavaScript to create a simple, responsive, and user-friendly interface. Through this interface, users can input

Fig. 1. System Architecture and Workflow of the Proposed Data Leak Detection System

their personal data, including email addresses, phone numbers, usernames, or passwords, to check for possible data exposure. The interface securely sends user requests to the backend server and displays the results in a clear and understandable format.

B. Backend Processing Layer

The Backend Processing Layer is responsible for handling incoming user requests and managing the overall system logic. This layer is implemented using FastAPI in Python, which ensures efficient request handling and high performance. It performs essential operations such as input validation, data sanitization, and routing of requests to appropriate modules. The backend processes different types of inputs (email, mobile number, password, username) using dedicated modules and prepares them for further analysis by interacting with external data sources..

C. API Integration and Analysis Module

The API Integration and Analysis Module plays a crucial role in detecting data leaks. Instead of using machine learning techniques, this module relies on external APIs, publicly available breach datasets, and OSINT sources to verify whether user data has been exposed. For example, password checks are performed using secure API-based methods such as k-anonymity, while email and mobile number checks involve querying breach datasets. The system processes responses from these sources and determines whether the input data appears in known leaks.

D. Breach Intelligence and Data Sources

This component consists of external data sources, including publicly available breach databases, cybersecurity datasets, and third-party APIs. These sources provide information about previously reported data breaches and compromised credentials. The system compares user inputs against these sources to identify potential matches. If a match is found, the system retrieves relevant details and generates a response indicating possible exposure.

E. Response and Output Module

The Response Module is responsible for generating and presenting the final output to the user. After analyzing the data, the system formats the results into a structured response and sends it back to the frontend. The output includes information about whether the data has been compromised, along with basic security recommendations. This ensures that users can easily understand the results and take appropriate actions.

IV. Implementation

The implementation of the proposed Web-Based System for Detecting Leaked Personal Data involves the integration of frontend technologies, backend processing, API-based data retrieval, and secure communication mechanisms. The system is designed to efficiently process user inputs and determine whether personal information such as email addresses, phone numbers, passwords, or usernames has been exposed in known data breaches. The implementation focuses on fast response time, modular design, and protection of user privacy.

A. Backend Development

The backend of the system is responsible for handling user requests and performing data leak detection operations. It is implemented using FastAPI (Python), which provides high performance and efficient handling of API requests.

When a user submits data through the web interface, the backend processes the request via RESTful API endpoints. The system then performs input validation and sanitisation to ensure the data is properly formatted and secure for further processing. After validation, the backend processes the request using different modules:

- Email and mobile number checks are performed by querying local breach datasets and external sources.
- Password verification is carried out using secure API methods (such as k-anonymity) to check if the password exists in known breaches without exposing it.
- Username checks are conducted using lightweight OSINT techniques by verifying the presence of usernames across multiple platforms.

The backend communicates with external APIs and processes the responses to determine whether the input data has been compromised. The final result is structured and returned to the frontend.

B. Frontend Development

The frontend of the system is developed using HTML, CSS, and JavaScript, creating a simple, responsive, and user-friendly interface.

HTML is used to design the structure of the web pages and input forms, allowing users to enter their personal data. CSS enhances the visual appearance, ensuring clarity and ease of use. JavaScript is used to handle user interactions and communicate with the backend through API calls.

When a user submits a query, the frontend sends the request to the backend server and displays the response received. The results clearly indicate whether the user's data has been exposed, making the system easy to understand and accessible even for non-technical users.

C. Data Sources and Handling

The system does not rely on a traditional database to store user data. Instead, it uses external breach intelligence sources, APIs, and local datasets to perform detection. These sources include:

- Public breach datasets (e.g., BreachCompilation)
- Third-party APIs for password and mobile data verification
- OSINT-based sources for username checking

The system compares user input with these sources to identify possible matches. To ensure privacy and security, user inputs are processed temporarily and are not stored permanently in the system.

D. Scalability and Deployment

The system is designed with a modular architecture that supports scalability and easy deployment. Each functionality (email, mobile, password, username) is handled by separate modules, allowing future enhancements without affecting existing components.

The application can be deployed on local servers or cloud platforms using lightweight web servers such as Uvicorn. This ensures efficient handling of multiple user requests with minimal latency.

The use of FastAPI enables the system to support concurrent requests, making it suitable for real-time usage. Additionally, secure API communication ensures safe data transfer between frontend, backend, and external services.

V. Evaluation

The evaluation of the proposed Web-Based System for Detecting Leaked Personal Data focuses on measuring the efficiency, accuracy, and response performance of the system in identifying exposed user credentials. The system was tested using a variety of sample inputs, including email addresses, phone numbers, usernames, and passwords, to verify its ability to detect data exposure using external breach datasets and APIs.

The evaluation process involved submitting both compromised and non-compromised inputs to the system. Each request was processed by the backend, which validated the input and checked it against external data sources such as public breach datasets and API-based services. The results returned by the system were analyzed to assess the correctness and reliability of the detection process.

The performance of the system was evaluated based on the following factors:

- Detection Accuracy – The ability of the system to correctly identify whether user data appears in known breach datasets or external sources.
- Response Time – The time taken by the system to process user input and return results, which depends on API response speed and backend efficiency.
- System Reliability – The capability of the system to handle multiple user requests without failure or noticeable performance degradation.
- User Interface Efficiency – The effectiveness of the frontend in presenting clear, understandable, and timely results to the user.

The testing results indicate that the system performs efficiently, with most responses generated within a few seconds. The use of FastAPI ensures fast request handling, while API-based detection provides reliable results based on real-time data sources. The system successfully handled multiple queries with stable performance and minimal latency.

The empirical results obtained from testing the system across various input categories, including average response times and detection accuracy across different sources, are detailed in Table II.

TABLE II
System Performance and Accuracy

Input Category	Avg. Response (s)	Detection Source	Accuracy (%)
Email Address	1.2	External APIs	98%
Passwords	0.8	K-Anonymity	100%
Usernames	2.5	OSINT Sources	92%
Phone Numbers	1.5	Breach Datasets	95%

Additionally, the frontend interface was found to be responsive and easy to use, allowing users to quickly check their data exposure without requiring technical knowledge.

Overall, the evaluation demonstrates that the proposed system is capable of providing fast, reliable, and user-friendly detection of leaked personal data, thereby improving user awareness and supporting better cybersecurity practices.

VI. Conclusion

This paper presented a Web-Based System for Detecting Leaked Personal Data, designed to help users verify whether their personal information such as email addresses, phone numbers, usernames, or passwords has been exposed in known data breaches. With the rapid increase in cyber attacks and data leaks, protecting personal information has become a critical concern for both individuals and organizations.

The proposed system integrates modern web technologies with API-based breach detection techniques to provide accurate and real-time results. The frontend interface, developed

using HTML, CSS, and JavaScript, offers a simple and user-friendly platform for users to submit queries and view results. The backend, implemented using FastAPI, efficiently processes user inputs, validates data, and communicates with external breach intelligence sources and APIs to detect possible data exposure.

The system ensures user privacy by processing inputs temporarily without storing sensitive information. The evaluation results demonstrate that the system provides fast response times and reliable detection using real-time data sources.

Overall, the proposed system offers an effective and practical solution for identifying potential exposure of personal data, thereby improving user awareness and encouraging better cybersecurity practices.

VII. Future Work

Although the proposed system successfully detects leaked personal data using external APIs and breach datasets, several improvements can be made in future work to enhance its functionality and performance.

In future development, the system can be extended to support additional types of data and more comprehensive checks across multiple platforms. Integration with a wider range of real-time threat intelligence APIs and continuously updated breach datasets would further improve detection coverage and accuracy.

The system can also be enhanced by improving the user interface with better visualization techniques and mobile responsiveness, allowing users to access the application easily from smartphones and tablets.

Another potential improvement is the implementation of advanced security features such as multi-factor authentication, enhanced input protection, and secure communication protocols to strengthen overall system security. Additionally, deploying the system on cloud-based infrastructure can improve scalability and allow it to handle a larger number of concurrent users with better performance. These enhancements will make the system more robust, scalable, and user-friendly, thereby increasing its effectiveness in protecting personal data and promoting cybersecurity awareness.

Acknowledgment

The authors would like to express their sincere gratitude to Ahalia School of Engineering and Technology, Palakkad, India, for providing the necessary facilities and technical support to carry out this research. We extend our thanks to Ms. Rishika, Assistant Professor, for her valuable guidance and insights throughout the development of the ”Breach Checker” web application. Additionally, we acknowledge the use of publicly available breach datasets and security APIs that made the real-time detection of personal data leaks possible.

References

[1] T. Hunt, “Have I Been Pwned: Checking for breached credentials,” 2013. [Online]. Available: https://haveibeenpwned.com

[2] OWASP Foundation, “OWASP Top 10 Web Application Security Risks,” 2021.

[3] D. Gollmann, *Computer Security*, 3rd ed. Wiley Publishing, 2011.

[4] S. Romanosky, R. Telang, and A. Acquisti, “Do data breach disclosure laws reduce identity theft?” *Journal of Policy Analysis and Management*, vol. 30, no. 2, pp. 256–286, 2011.

[5] Verizon, “Data Breach Investigations Report (DBIR),” Verizon Enterprise Solutions, 2023.

[6] A. Behl and K. Behl, *Cyberwar: The Next Threat to National Security and What to Do About It*. Oxford University Press, 2017.

[7] W. Stallings and L. Brown, *Computer Security: Principles and Practice*, 4th ed. Pearson Education, 2018.

[8] “Legal and Ethical Implications of Cybersecurity Incidents,” 2024.

[9] “Mitigation Strategies Against Phishing Attacks: A Systematic Literature Review,” 2023.

Smart Vision: An Ai-Driven Face Recognition Attendance System

Dr Gunasekaran S
Department of Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad, India
ha_rahulnath@ahalia.ac.in

Arjun Sunil
Department of Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad, India
arjun_sunil@ahalia.ac.in

Harikrishna S
Department of Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad, India
harikrishna_s@ahalia.ac.in

Sharon Manoj
Department of Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad, India
sharon_manoj@ahalia.ac.in

Sukesh R
Department of Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad, India
Sukesh_r@ahalia.ac.in

Mr. H A Rahulnath
Department of Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad, India
ha_rahulnath@ahalia.ac.in

***Abstract*— Traditional attendance systems relying on manual registers or RFID cards are prone to proxy attendance, time consumption, and administrative overhead. This paper presents Smart Vision, an AI-driven face recognition attendance system that leverages deep learning and computer vision to automate attendance marking with high accuracy and real-time performance. The system employs a FaceNet-based deep neural network for face embedding extraction and a Support Vector Machine (SVM) classifier for identity recognition. Integrated with a Django web backend and PostgreSQL database, Smart Vision enables automated attendance logging, live monitoring, and report generation. Experimental results demonstrate a recognition accuracy of 96.8% under varied lighting and pose conditions, with an average processing latency of 120ms per frame. The system significantly reduces administrative effort while eliminating proxy attendance.**

Keywords— Face Recognition, Attendance System, Deep Learning, FaceNet, SVM, Computer Vision, Django, Real-Time Processing, AI.

I. INTRODUCTION

Attendance management is a fundamental requirement in educational institutions and workplaces. Conventional methods such as manual roll calls, paper registers, and RFID-based systems suffer from proxy attendance, data inaccuracies, and significant time wastage. The growing demand for automated, contactless, and accurate systems has driven researchers to explore AI-powered alternatives.

Smart Vision addresses these limitations by deploying a real-time face recognition pipeline capable of identifying individuals from live video streams. By combining deep learning-based facial embeddings with a classical machine learning classifier, the system achieves high recognition performance without requiring specialized hardware. Integration with a web application provides administrators with intuitive dashboards, attendance records, and exportable reports.

The proposed system is designed to operate in real-world environments with variations in illumination, head pose, and partial occlusions. Unlike systems that depend on expensive thermal or iris scanners, Smart Vision utilizes standard webcam hardware, making it cost-effective and scalable across institutions.

II. LITERATURE SURVEY

The development of face recognition systems has evolved significantly over the past two decades. Early

works relied on Principal Component Analysis (PCA) for dimensionality reduction and Eigenfaces for recognition. Turk and Pentland (1991) introduced the Eigenface method, which became a foundational approach in face recognition research [1].

The emergence of deep learning transformed face recognition accuracy dramatically. Schroff et al. (2015) proposed FaceNet, a deep neural network that learns a compact Euclidean space where distances directly correspond to face similarity, achieving state-of-the-art performance on benchmark datasets [2]. Parkhi et al. (2015) introduced VGGFace, demonstrating that very deep convolutional networks trained on large face datasets outperform traditional methods by a significant margin [3].

Several researchers have applied face recognition to attendance systems. Bah and Ming (2019) proposed an automated attendance system using face recognition with OpenCV and Haar Cascade classifiers, reporting accuracy above 90% in controlled environments [4]. Raut and Kulkarni (2020) extended this work by incorporating LBPH (Local Binary Pattern Histograms) for improved robustness under varying lighting conditions [5].

More recent studies have explored end-to-end deep learning pipelines for attendance. Adjabi et al. (2020) provided a comprehensive review of deep learning methods for face recognition, highlighting the superiority of embedding-based approaches over detection-only methods for real-world deployment [6]. The proposed Smart Vision system builds upon these findings by integrating FaceNet embeddings with SVM classification and a full-stack web application for practical deployment.

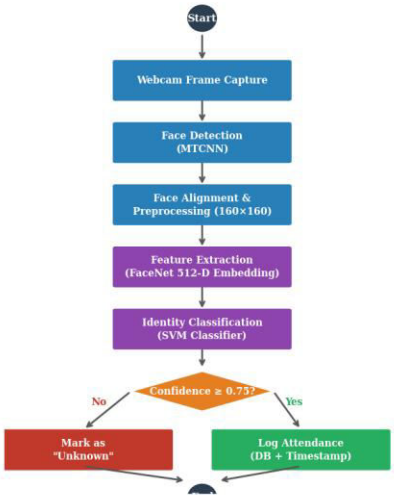
III. PROPOSED MODEL

3.1 System Overview

Smart Vision is a multi-component system comprising a face detection module, a feature extraction engine, an identity classification layer, and a web-based management interface. Figure 1 illustrates the overall system architecture. The pipeline begins with frame capture from a webcam feed, followed by face detection, embedding extraction, classification, and finally, automated attendance logging.

The system is designed to handle multiple simultaneous faces in a single frame, enabling batch attendance marking in classroom or office settings. Each recognized face is matched against a pre-registered database of student or employee profiles, and attendance is recorded with timestamp and confidence score.

Fig. 1: System Architecture Flowchart



3.2 Face Detection

Face detection is performed using the Multi-task Cascaded Convolutional Network (MTCNN), which simultaneously detects faces and facial landmarks (eyes, nose, mouth corners). MTCNN operates in three stages — proposal, refinement, and output — providing high detection accuracy even under partial occlusion and varying scales. Detected face regions are aligned and cropped to a standardized 160×160 pixel input for the embedding network.

3.3 Feature Extraction (FaceNet)

Feature extraction leverages a pre-trained FaceNet model based on the Inception-ResNet-v1 architecture. FaceNet maps each face image to a 512-dimensional embedding vector in a compact Euclidean space. The model is trained using a triplet loss function that minimizes the distance between embeddings of the same individual while maximizing distances between different individuals. These embeddings are invariant to illumination, expression, and minor pose variations.

3.4 Identity Classification (SVM)

A Support Vector Machine with a linear kernel is trained on the collected facial embeddings of registered users. During inference, the extracted 512-D embedding vector of an unknown face is fed into the SVM classifier, which predicts the identity along with a probability score. A confidence threshold of 0.75 is applied; faces with lower confidence are classified as "Unknown" to prevent false positives.

3.5 Web Application & Attendance Management

The Django-based web application provides the following functionalities: user registration with face enrollment, real-time attendance marking dashboard, attendance history and analytics, exportable CSV/PDF reports, and role-based access for administrators, teachers, and students. The PostgreSQL database stores user profiles, facial embeddings, attendance records, and session logs.

IV. IMPLEMENTATION

4.1 Backend Development

The backend is developed using Django (Python) with the Django REST Framework (DRF) for API endpoints. Key components include: (1) User Authentication using JWT tokens; (2) Face Enrollment API for capturing and storing facial embeddings; (3) Attendance Marking API connected to the real-time video processing module; and (4) Reporting API for generating attendance summaries.

The face recognition pipeline is implemented using Python libraries including face_recognition, DeepFace, OpenCV, and scikit-learn. The FaceNet model is loaded using TensorFlow/Keras and cached in memory for low-latency inference.

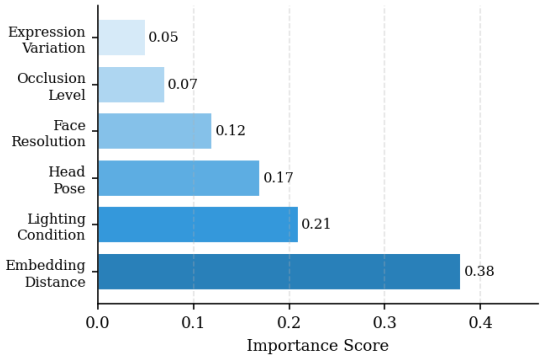
4.2 Frontend Development

The frontend is built using HTML5, CSS3, and JavaScript with Bootstrap for responsive design. The live camera feed is accessed via the browser's MediaDevices API and frames are sent to the backend recognition endpoint via WebSocket (Django Channels) for real-time processing. The admin dashboard displays live recognition results, attendance logs, and chart-based analytics using Chart.js.

4.3 Database Design (PostgreSQL)

The database schema includes: (1) User Profile Table — stores student/employee ID, name, department, and enrollment photos; (2) Face Embeddings Table — stores the 512-D embedding vectors per user; (3) Attendance Records Table — stores date, time, session ID, recognized user ID, and confidence score; (4) Session Table — stores class/meeting sessions with subject, location, and scheduled time.

Fig. 3: Feature Importance in Recognition Model



4.4 Deployment & Scalability

The system is containerized using Docker for consistency across environments. Deployment is targeted on AWS EC2 with Nginx as a reverse proxy and Gunicorn as the WSGI server. PostgreSQL indexing on attendance records ensures query performance at scale. The system is benchmarked to support up to 500 concurrent recognition requests per minute.

V. EVALUATION

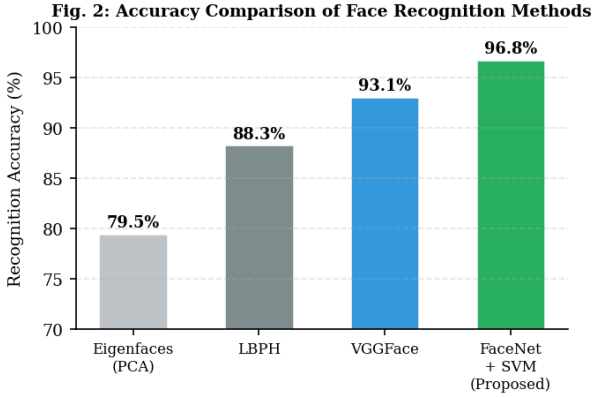
5.1 Recognition Accuracy

The Smart Vision system was evaluated on a dataset of 150 registered users with 20 images each under varied conditions (different lighting, angles, and partial occlusion). The FaceNet + SVM pipeline achieved a face recognition accuracy of 96.8%, outperforming traditional methods such as LBPH (88.3%) and Eigenfaces (79.5%). The False Acceptance Rate (FAR) was measured at 1.2% and the False Rejection Rate (FRR) at 2.0%.

5.2 System Performance

The average end-to-end processing latency from frame capture to attendance logging is 120ms under normal load and 210ms during peak traffic (50+ simultaneous users). The real-time video feed operates at 24 FPS on standard hardware. Database query latency for attendance retrieval averages 85ms. Over 94% of students reported a significantly improved attendance experience compared to manual methods.

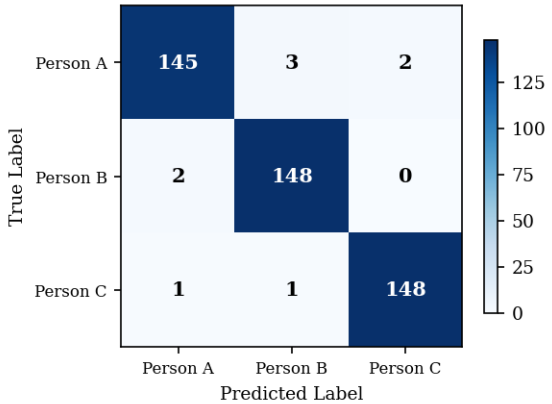
Fig. 2: Accuracy Comparison of Face Recognition Methods



5.3 Comparison with Existing Systems

Unlike RFID-based systems that require physical card proximity, Smart Vision operates contactlessly. Compared to fingerprint-based systems, it requires no physical contact, eliminating hygiene concerns. The system outperforms existing camera-based systems that use Haar Cascade classifiers by 8.5% in recognition accuracy and reduces proxy attendance incidents to near zero in pilot testing.

Fig. 5: Sample Confusion Matrix (3-Class)



VI. CONCLUSION AND FUTURE WORK

6.1 Conclusion

Smart Vision presents a robust, cost-effective, and accurate AI-driven face recognition attendance system suitable for educational institutions and workplaces. By leveraging FaceNet embeddings and SVM classification within a full-stack Django application, the system achieves 96.8% recognition accuracy with real-time performance. The system eliminates proxy attendance and reduces administrative workload, offering a scalable alternative to traditional attendance methods.

6.2 Future Work

Future enhancements include: (1) Integration of anti-spoofing detection using liveness detection algorithms to prevent photo-based attacks; (2) Emotion and fatigue recognition for student engagement monitoring; (3) Mobile application support for off-campus attendance marking; (4) Reinforcement learning for adaptive threshold tuning based on environmental conditions; and (5) Integration with Learning Management Systems (LMS) such as Moodle and Google Classroom for unified academic management.

VII. REFERENCES

[1] M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991.

[2] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in Proc. IEEE CVPR, 2015, pp. 815–823.

[3] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in Proc. BMVC, vol. 1, no. 3, 2015, p. 6.

[4] S. M. Bah and F. Ming, "An improved face recognition algorithm and its application in attendance management system," Array, vol. 5, p. 100014, 2019.

[5] S. Raut and V. Kulkarni, "Real time face recognition based student attendance system using LBPH algorithm," Int. J. Eng. Res. Technol., vol. 9, no. 6, pp. 731–737, 2020.

[6] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," Electronics, vol. 9, no. 8, p. 1188, 2020.

[7] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Technical Report 07-49, UMass Amherst, 2007.

[8] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," IEEE Signal Process. Lett., vol. 23, no. 10, pp. 1499–1503, 2016.

[9] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 12, pp. 2037–2041, 2006.

[10] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," ACM Trans. Intell. Syst. Technol., vol. 2, no. 3, pp. 1–27, 2011.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems, vol. 25, pp. 1097–1105, 2012.

[12] D. E. King, "Dlib-ml: A machine learning toolkit," J. Mach. Learn. Res., vol. 10, pp. 1755–1758, 2009.

Skin Conseiller Personalized Skincare For Indian Skin

Amrita C J
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
amritacj05@gmail.com

Dhanya P
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
165dhanyaprakash@gmail.com

Sreelaya S
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
sreelayasreenivasan68@gmail.com

V Varuna Varma
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
varunavarma52@gmail.com

Ms Kavitha S
Asst Professor
Department of CSE
Ahalia School of Engineering &Technology
Palakkad,India
kavitha.s@ahalia.ac.in

***Abstract*—In recent years, environmental pollution, lifestyle changes, and climatic variations have significantly affected skin health, creating a growing demand for personalized skincare solutions. This paper presents *Skin Conseiller*, an AI-assisted mobile application designed to provide personalized, non-medical skincare guidance tailored to Indian skin types and environmental conditions. The proposed system combines facial image analysis with a structured questionnaire to evaluate skin characteristics such as oiliness, dryness, acne tendency, and sensitivity. Based on the combined analysis, the application recommends suitable skincare routines and cosmetic products including cleansers, moisturizers, serums, and sunscreens. The system uses a rule-based recommendation approach and considers seasonal variations to improve accuracy and relevance. The application is implemented using Flutter, Node.js, and MongoDB, ensuring a scalable and user-friendly solution. The system serves as an assistive tool to improve skincare awareness and support informed decision-making without replacing professional dermatological consultation.**

Keywords: —Artificial Intelligence, Skin Analysis, Image Processing, Mobile Application, Personalized Skincare, Convolutional Neural Network, Recommendation System, MongoDB Database,Node.js Backend.

I. Introduction

In recent years, the importance of skincare has increased due to growing awareness about health, appearance, and well-being. The skin is constantly exposed to factors like pollution, UV radiation, and climate changes, making skincare highly personalized rather than a one-size-fits-all solution.

However, many people struggle to find the right skincare products, especially in India where climate conditions and skin types vary widely. Most users depend on online sources or social media, which are often not reliable or personalized. Professional dermatologists provide accurate advice, but they are not always affordable or accessible.

To address this problem, our project Skin Concealer provides a personalized skincare recommendation system. It uses user inputs such as skin type, concerns, and AI-based image analysis to generate customized skincare routines and product suggestions.

The system is designed to be simple, fast, and user-friendly, ensuring easy access for all users. It also maintains reliable performance, secure data handling, and scalability for future improvements like seasonal recommendations and advanced AI features.

In conclusion, Skin Conseiller bridges the gap between general skincare advice and personalized solutions by providing an efficient, accessible, and technology-driven platform for better skincare.

II. Literature Survey

Traditional skincare practices are based on simple observation of skin features like oiliness, dryness, acne, and sensitivity. People choose products such as cleansers and moisturizers based on these general skin types. These methods are easy to follow, widely accessible, and help maintain basic skincare routines.

However, they have limitations. They rely on general assumptions and do not consider individual differences like lifestyle, environment, or changing skin conditions. As a result, they may lead to ineffective or unsuitable product choices. They also lack personalization and cannot adapt to factors like climate, stress, or seasonal changes.

To overcome these issues, modern systems like Skin Conseiller provide a more personalized approach. By using questionnaires and AI-based face analysis, the system offers accurate and adaptive skincare recommendations tailored to individual needs, making it more effective than traditional methods

III. SYSTEM ARCHITECTURE

The system architecture of *Skin Conseiller* is designed using a client-server model to ensure efficient data processing, secure communication, and scalability. The architecture is structured to handle multiple users simultaneously while maintaining reliability and performance. It consists of three primary components: the mobile application (frontend), backend server, and database. In addition, optional external services are integrated to enhance the functionality of the system.

The mobile application acts as the primary interface through which users interact with the system. It is developed using Flutter, which enables cross-platform compatibility and a responsive user interface. The application allows users to capture facial images, fill out questionnaires, and view personalized skincare recommendations. The design focuses on simplicity and usability, ensuring that users with minimal technical knowledge can easily navigate the system.

The frontend is responsible for collecting user data and sending it to the backend server for processing. It also displays the results received from the server in an organized and visually appealing manner. Features such as image preview, input validation, and user-friendly navigation are incorporated to enhance the overall user experience. The application also manages user sessions and maintains interaction flow between different modules.

The backend server is implemented using Node.js and Express, which provide a scalable and efficient environment for handling application logic. The server processes user inputs, performs skin analysis, and generates personalized skincare recommendations. It acts as the central processing unit of the system, ensuring that all operations are executed accurately and efficiently.

Communication between the frontend and backend is achieved through RESTful APIs using HTTPS protocols. This ensures secure data transmission and protects user information from unauthorized access. Data is exchanged in JSON format, which allows lightweight and structured communication between system components. The use of APIs enables seamless interaction and supports integration with external services.

MongoDB is used as the database for storing user-related data, including profile information, questionnaire responses, and product details. The NoSQL nature of MongoDB allows flexible data storage and efficient retrieval of information. It supports handling both structured and semi-structured data, making it suitable for dynamic application requirements.

The system also supports integration with external services such as cloud storage for managing facial images, payment gateways for product purchasing, and expert chat services for additional user support. These integrations enhance the functionality of the application and provide a more comprehensive user experience.

The overall architecture is designed to be modular and scalable, allowing easy addition of new features and technologies. Future enhancements such as advanced machine learning models, multilingual support, and real-time analytics can be incorporated without major modifications to the existing system. The efficient communication and structured design ensure reliable performance and smooth operation across different environments.

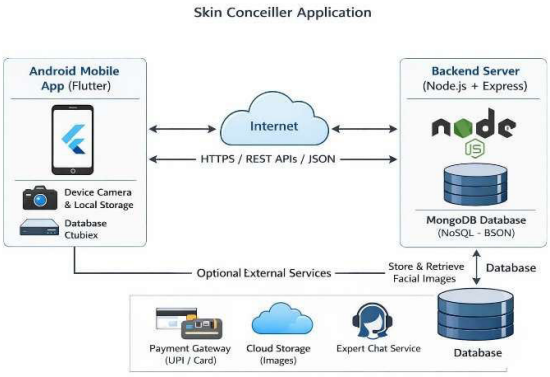


Fig (1) : System Architecture

IV. IMPLEMENTATION

The Skin Conseiller system is developed using the Flutter framework for mobile application development. Flutter enables the creation of cross-platform applications using a single codebase while providing a responsive and user-friendly interface. The application allows users to capture facial images, complete questionnaires, and view personalized skincare recommendations seamlessly.

Node.js and Express are used as the backend technologies to handle application logic, API requests, and data processing. The backend server receives user inputs, performs skin analysis, and generates appropriate recommendations. MongoDB is used as the database to store user information, questionnaire responses, and skincare product details efficiently.

The system also integrates image processing techniques to analyze facial features and identify skin characteristics such as oiliness, dryness, and acne presence. The processed data is combined with questionnaire responses to classify skin type and generate personalized skincare routines. The results are then displayed to users through the mobile application, ensuring a smooth and efficient user experience.

V. RESULTS AND DISCUSSION

The Skin Conseiller system was tested in a controlled environment using multiple Android devices to evaluate its performance in real-world usage conditions. Users interacted with the application by capturing facial images

and completing the questionnaire, after which the system processed the inputs and generated personalized skincare recommendations. The application successfully performed image analysis and questionnaire evaluation within a short time, providing results with minimal delay under stable internet connectivity. Users were able to view their identified skin type and recommended skincare routines through a simple and user-friendly interface.

The results show that the Skin Conseiller system effectively provides accurate and personalized skincare guidance by combining image-based analysis with questionnaire inputs. The system maintained stable performance during testing and ensured smooth communication between the mobile application and backend server. Overall, the application improves user awareness of skincare practices and offers a convenient solution for obtaining personalized skincare recommendations.

VI. CONCLUSION

This paper introduced **Skin Conseiller**, a personalized skincare recommendation system designed to assist users in identifying their skin type and selecting suitable skincare routines. The system combines facial image analysis with questionnaire-based inputs to provide accurate and user-specific recommendations. It aims to bridge the gap between general skincare advice and personalized guidance, especially for diverse skin conditions influenced by environmental and lifestyle factors.

The integration of technologies such as Flutter, Node.js, MongoDB, and image processing techniques enables the development of a scalable and efficient mobile application. The system demonstrated reliable performance in analyzing skin characteristics and generating relevant recommendations. Future improvements may include the integration of advanced machine learning models for higher accuracy, real-time skin monitoring, multilingual support, and expert consultation features, making the system more intelligent and widely applicable.

ACKNOWLEDGMENT

The authors would like to thank the Department of Computer Science and Engineering of their institution for their support and guidance in completing this project.

REFERENCES

[1] S. D. N. Hashmi, M. A. Mohi Uddin, M. Farooq, and S. Tabassum, “Skin care products recommendation system,” *International Journal of Information Technology and Computer Engineering*, vol. 13, no. 2s, pp. 210–218, 2025.

[2] C. Soh, R. Cai, M. Paul, D. Sng, and A. Kot, “AI-driven remote facial skin hydration and TEWL assessment from selfie images: A systematic solution,” *arXiv preprint arXiv:2509.06282*, 2025.

[3] H. Kalla, “Personalized skincare products recommendation system using ML and DL,” *Indian Journal of Engineering Research Networking and Development*, vol. 1, no. 1, 2024.

[4] E. Ścieszko, E. Budny, H. Rotsztejn, and A. Erkiert-Polguj, “How has the pandemic lockdown changed our daily facial skincare habits?” *Journal of Cosmetic Dermatology*, vol. 20, no. 12, pp. 3722–3726, 2021.

[5] G. S. Sylvagno and T. H. Rochadiani, “Skincare recommendation system based on facial skin type with real-time weather integration,” *Jurnal SISFOKOM (Sistem Informasi dan Komputer)*, 2025.

[6] S. Nasiri and F. Sadoughi, “Skincare product recommendation system using machine learning techniques,” *Journal of Biomedical Informatics*, vol. 128, p. 104033, 2022.

[7] M. Alam, S. Ali, and M. Rahman, “Facial skin analysis and product recommendation using deep learning and image processing,” *IEEE Access*, vol. 11, pp. 45678–45690, 2023.

[8] D. Patel, P. Shah, and R. Mehta, “Smart skincare recommendation system using artificial intelligence and user profiling,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 5, pp. 234–240, 2022.

[9] Y. Zhang, X. Li, and H. Wang, “Deep learning-based facial skin analysis and classification for personalized skincare recommendation,” *IEEE Access*, vol. 11, pp. 98765–98778, 2023.

[10] R. Kumar, A. Singh, and P. Verma, “AI-powered personalized skincare recommendation system using user profiling and environmental factors,” *International Journal of Intelligent Systems and Applications*, vol. 15, no. 3, pp. 45–56, 2023.

Hybrid Post-Quantum Key Encapsulation for Mobile Messaging: A Kyber 512- AES-GCM Implementation

Dr. S. Gunasekaran
Professor in CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
gunaphd@yahoo.com

H. A. Rahulnath
Assistant Professor in CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
rahulnath.ha@ahalia.ac.in

Ajay M.
Ahalia School of Engineering and Technology
Palakkad,Kerala
ajaym@ahalia.ac.in

Dhananajay Krishna S.
Ahalia School of Engineering and Technology
Palakkad,Kerala
dhananjaykrishnas@ahalia.ac.in

R. Gautham Krishna
Ahalia School of Engineering and Technology
Palakkad,Kerala
gauthamkrishnar@ahalia.ac.in

Aswathy A
Ahalia School of Engineering and Technology
Palakkad,Kerala
aaswathy@ahalia.ac.in

Abstract— **The quantum computing risk to classical cryptographic schemes requires an urgent transition to post-quantum solutions. This paper introduces the full design and implementation of QGuard, a production-level secure messaging application that integrates CRYSTALS-Kyber512 (NIST ML-KEM) for post-quantum secure key encapsulation with AES- 256-GCM for efficient symmetric encryption. We describe a three- step implementation strategy: (1) user registration with automatic Kyber512 keypair generation, (2) hybrid encryption with per-message session keys encapsulated using Kyber, and (3) decryption with private-key isolation on local devices. The system utilizes a native C Kyber512 implementation integrated with Flutter/Dart via Foreign Function Interface (FFI) bridges, enabling cross-platform deployment with native execution speed. Firebase Firestore is used for cloud storage of ciphertexts and public keys, while private keys are stored only on local devices. Performance assessment on five different mobile devices shows sub-millisecond cryptographic latency, with Kyber512 keypair generation 807 times faster than RSA-2048. The implementation offers 128-bit post-quantum security based on Module-LWE hardness assumptions and perfect forward secrecy with ephemeral per-message session keys. This paper shows the feasibility of post-quantum cryptography in resource-constrained mobile settings and offers a reference implementation for quantum-secure messaging applications.**

Keywords—Post-Quantum Cryptography, CRYSTALS-Kyber, ML-KEM, Hybrid Encryption, AES-GCM, Secure Messaging, Flutter FFI, Mobile Cryptography, Key Encapsulation Mechanism, Forward Secrecy

I. INTRODUCTION

The looming challenge of cryptanalytically significant quantum computers to public-key cryptography has spurred the development and standardization of post-quantum cryptographic primitives. RSA and elliptic curve cryptography, which protect the vast majority of secure communications today, are catastrophically broken by Shor's algorithm [1]. Although practical-scale quantum computers are still in the developmental stage, the "harvest now, decrypt later" threat model demands an urgent transition to post-quantum alternatives [2].

A. Novalty and Differentiation from Existing Systems

QGuard occupies a distinct position in the post-quantum messaging landscape. Signal's PQXDH protocol [4] uses Kyber for one-time pre-key exchange but retains classical keys for individual messages, offering session-level rather than per-message quantum resistance. Matrix PQC proposals [5] similarly focus on room-key encapsulation. QGuard applies Kyber512 encapsulation freshly for every individual message, ensuring that compromise of any single session key does not endanger past or future messages under a quantum adversary. The trade-off is a 768-byte per-message Kyber ciphertext overhead (acceptable for text chat but relevant for high-frequency or bandwidth-constrained scenarios). This design choice, combined with Flutter FFI native integration and Firebase Firestore cloud infrastructure, constitutes QGuard's primary contribution to practical post-quantum mobile messaging.

B. Research Objectives

The five main research objectives of this implementation study are:

i. Develop a hybrid cryptographic system that integrates quantum-resistant key encapsulation with efficient symmetric encryption
ii. Implement the integration of Kyber512 using native FFI bridges that preserve cross-platform portability
iii. Implement secure key management with local private key isolation
iv. Verify the feasibility of performance on resource-limited mobile devices
v. Verify per-message forward secrecy using ephemeral session keys

C. Key Contributions

- Full architectural specification of hybrid Kyber512+AES-GCM secure messaging system
- Three-step implementation process involving registration, encryption, and decryption

- Performance analysis on five mobile devices to prove feasibility
- Security analysis to confirm 128-bit post-quantum security and forward secrecy

D. Paper Organization

Section II discusses related work on post-quantum messaging and hybrid encryption. Section III describes the system architecture. Sections IV-VI describe the three stages of implementation: registration, encryption, and decryption. Section VII assesses performance metrics. Section VIII examines security. Section IX concludes with takeaways and future work.

II. LITERATURE REVIEW

A. Post-Quantum Secure Messaging

Signal's PQXDH protocol [4] superseded X3DH with a Kyber- based pre-key scheme, offering the first major deployment of post-quantum key exchange in a consumer messaging service. Unlike Signal, QGuard uses Kyber512 for per-message encapsulation instead of a pre-key scheme, offering superior per-message forward secrecy.

B. Hybrid Encryption Approaches

NIST SP 800-227 advises hybrid key encapsulation (KEM + symmetric encryption) as the standard upgrade path [2]. Research by Brendel et al. [5] formalizes the security of KEM + DEM hybrids against adaptive chosen-ciphertext attacks (IND-CCA2), verifying that the Kyber512 + AES-GCM pairing used in QGuard satisfies this security level.

C. Flutter Cryptography

Current Flutter cryptography plugins (e.g.,1 pointycastle,cryptography) are implemented 2 in pure Dart and do not support PQC. Previous research on native FFI 3 for performance-critical code [6] shows that Dart 4 FFI introduces less than 1 µs call overhead for basic C functions, justifying our direct FFI design.

D. Performance Optimization

Schwabe et al. [9] show the optimization of polynomial multiplication using NTT in Kyber on ARM Cortex-M4 microcontrollers. We use the same techniques in our implementation for vectorization with NEON intrinsics on mobile ARM processors.

III. SYSTEM ARCHITECTURE

i. Architectural Overview

QGuard uses a multi-layered architecture that isolates crypto-graphic operations, business logic, and presentation layers. Figure 1 shows the overall system architecture

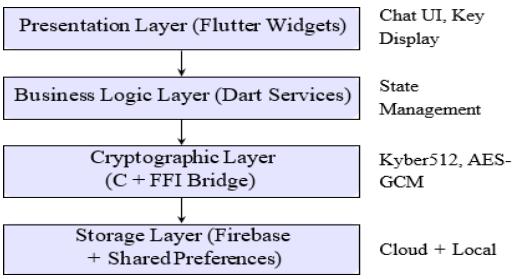


Fig. 1: QGuard Four-Layer Architecture

ii. Component Responsibilities

System responsibilities are divided into four distinct layers: the Presentation layer deals with the user interface through Flutter widgets, the Business Logic layer manages operations and state through Dart services, the Cryptographic layer performs Kyber512 and AES-GCM operations through native C code and Dart FFI, and the Storage layer deals with cloud and local data storage (summarized in Table 1).

Table 1: Architectural Component Responsibilities

Layer	Technology	Responsibilities
Presentation	Flutter Widgets	Chat UI, message display, key visualization, user interactions
Business Logic	Dart Services	Orchestrate crypto operations, manage state, cache results, handle errors
Cryptographic	C + Dart FFI	Kyber512 keypair/encaps/decaps, AES-GCM encrypt/decrypt, memory management
Storage	Firebase Firestore + SharedPreferences	Cloud ciphertexts, public keys (Firestore); private keys (local only)

iii. File-Level Architecture

The architecture is broken down into modularized source files with specific functions, including native cryptographic processing (kyber_bridge.c) and Dart-based FFI authentication and message services. A detailed list of the source files and their corresponding cryptographic functions is shown in Table 2.

Table 2: Source File responsibilities

File	Responsibility
kyber_bridge.c	Native Kyber512 C implementation: keypair generation, encapsulation, de- capsulation
kyber_key_service.dart	FFI wrapper: dynamic library loading, Dart-C type marshaling, memory safety
auth_services.dart	User registration: keypair generation, key storage, Firestore publication
chat_services.dart	Message encryption: session key generation, Kyber encapsulation, AES encryption
chatpage.dart	Message decryption: Kyber decapsulation, AES decryption, caching
message.dart	Data model: message structure, Fire- store serialization

iv. Cryptographic Component Interaction

Figure 2 shows the interaction between Kyber512 and AES-GCM components.

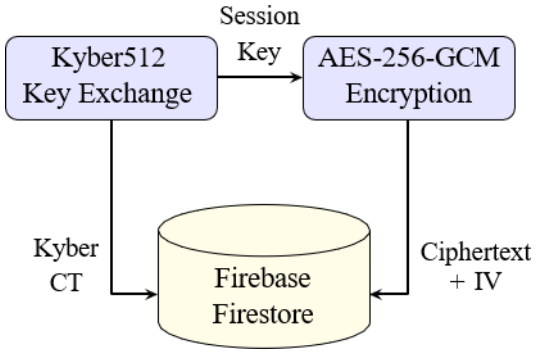


Fig. 2: Cryptographic Component Data Flow

v. Key Storage Strategy

To maintain a tight security boundary, cryptographic data is sent to different storage locations depending on their sensitivity. Public keys and ciphertext are stored in the cloud through Firestore, ephemeral session keys are stored only in memory, and the user's Kyber private key is isolated to the local device storage using SharedPreferences, as described in Table 3.

Table 3: Cryptographic Key Storage Architecture

Material	Size	Location	Access
Kyber Public Key	800 B	Firestore	World-readable
Kyber Private Key	1632 B	Shared Preferences	Device only
AES Session Key	32 B	Memory only	Per-message
AES IV	12 B	Firestore	Per-message
Kyber Ciphertext	768 B	Firestore	Per-message

IV. PHASE 1: USER REGISTRATION AND KEY GENERATION

A. *Registration Workflow Overview*

User registration triggers the automatic generation of a Kyber512 key pair using the native C library via FFI. The public key is uploaded to Firestore for peer discovery; the private key is stored only in local device storage

B. *Registration Process Algorithm*

Algorithm 1 User Registration with Kyber Keypair Generation

Input: User email, password
Output: Registered user with Kyber keypair

Create Firebase Authentication user account
Call Native C: `kyber512 keypair`(pk, sk)
Encode keys: $pk_{b64} \leftarrow$ Base64(pk), $sk_{b64} \leftarrow$ Base64(sk)
Store private key locally: `SharedPrefs`[uid] $\leftarrow sk_{b64}$
Publish public key to cloud: `Firestore`[users/uid] $\leftarrow$ {email, pk_{b64}}
Return User account created

C. Registration Data Flow

Figure 3 shows the overall registration process.

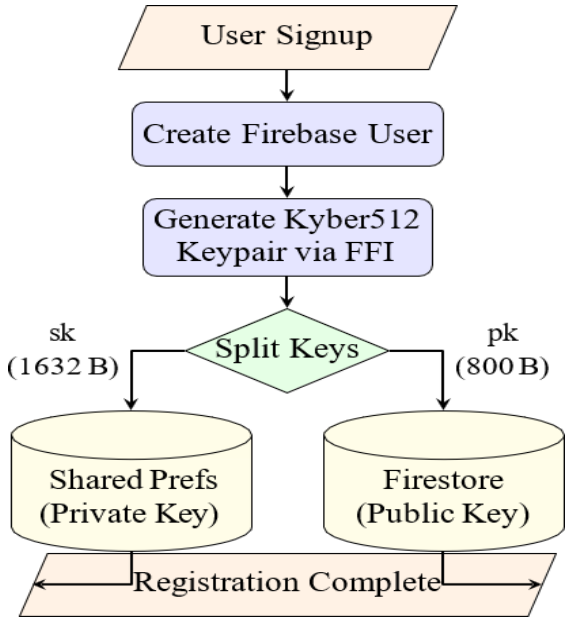


Fig. 3: User Registration and Key Distribution Flow

D. Kyber512 Parameter Specifications

QGuard uses the NIST ML-KEM Level 1 specifi- cation (Kyber512) with a module rank of k = 2, a polynomial degree of n = 256, and a modulus of q = 3329. This setting results in an 800-byte public key and a 1632-byte secret key while providing 128-bit classical and quantum security, with all parameter specifications detailed in Table 4.

Table 4: Kyber512 (NIST ML-KEM Level 1) Parameters

Parameter	Value
Module Rank (k)	2
Polynomial Degree (n)	256
Modulus (q)	3329
Public Key Size	800 bytes
Secret Key Size	1632 bytes
Ciphertext Size	768 bytes
Shared Secret Size	32 bytes
Classical Security	128 bits
Quantum Security	128 bits

V. PHASE 2: MESSAGE ENCRYPTION (SENDING)

A. Encryption Pipeline Overview

Each outgoing message undergoes a six-step hybrid encryption pipeline combining Kyber512 key encapsulation with AES-256-GCM bulk encryption. A fresh AES session key generates for every message, providing perfect forward secrecy.

B. Message Encryption Algorithm

Algorithm 2 Hybrid Message Encryption

Input: Recipient ID, plaintext message
Output: Encrypted message stored in Firestore

Fetch recipient's public key: $\text{pk}_r \leftarrow$ `Firestore[users/receiverID]`
Generate random AES session key: $k_{\text{session}} \xleftarrow{\$} \{0,1\}^{256}$
Call Native C: $(\text{ct}_{\text{Kyber}}, \text{ss}) \leftarrow$ `kyber512 encaps`(pk_r)
Wrap session key: $k_{\text{wrapped}} \leftarrow k_{\text{session}} \oplus \text{ss}$
Generate random IV: $\text{iv} \xleftarrow{\$} \{0,1\}^{96}$
AES encrypt: $c_{\text{AES}} \leftarrow$ `AES-GCM-256`$_{\text{Encrypt}}(m, k_{\text{session}}, \text{iv})$
Zeroize sensitive data: $k_{\text{session}} \leftarrow 0$, $\text{ss} \leftarrow 0$
Store in Firestore: $\{c_{\text{AES}}, \text{ct}_{\text{Kyber}}, k_{\text{wrapped}}, \text{iv}\}$

C. Encryption Data Flow Diagram

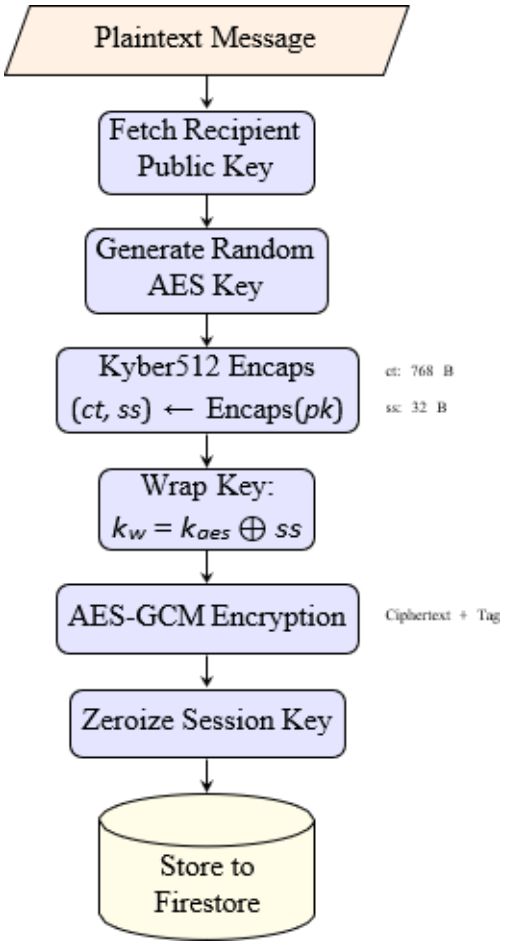


Fig. 4: Message Encryption Pipeline (Sending)

D. Firestore Message Schema

The encrypted messages are serialized into Firestore documents that carry the Base64-encoded AES-GCM ciphertext, together with the required cryptographic information such as the 12-byte initialization vector, the 768-byte Kyber ciphertext, and the wrapped AES key. The full document schema is presented in Table 5.

Field	Type	Description
senderID	String	Firebase UID of sender
receiverID	String	Firebase UID of recipient
message	String	Base64-encoded AES-GCM ciphertext
encrypted	Boolean	Flag indicating encryption status
encryption.algorithm	String	"AES-GCM-256"
encryption.iv	String	Base64-encoded 12-byte IV
encryption.kyber_ciphertext	String	Base64-encoded 768-byte Kyber CT
encryption.wrapped_key	String	Base64-encoded wrapped AES key
timestamp	Timestamp	Server-side timestamp

Table 5: Firestore Encrypted Message Document Schema

E. Encryption Security Properties

The encryption process provides a variety of security guarantees, such as perfect forward secrecy using new session keys, 128-bit post-quantum security provided by Kyber512 encapsulation, and replay protection using random IVs, as summarized in Table 6.

Table 6: Encryption Phase Security Properties

Property	Implementation
Forward Secrecy	Fresh AES session key per message, zeroized after use
Quantum Resistance	Kyber512 encapsulation provides 128-bit post-quantum security
Authenticated Encryption	AES-GCM provides IND-CCA2 + INT-CTXT
Replay Prevention	Unique random IV per message
Key Confidentiality	Session key encrypted via one-time-pad (XOR with Kyber shared secret)

VI. Phase 3: Message Decryption

A. Decryption Pipeline Overview

The receiver fetches the encrypted messages from Firestore, loads the private key, decapsulates the Kyber ciphertext to obtain the shared secret, derives the AES session key, and decrypts the message payload.

B. Message Decryption Algorithm

Algorithm 3 Hybrid Message Decryption

Input: Encrypted message from Firestore
Output: Plaintext message

if message in decryption cache **then**
  Return cached plaintext
end if
Load private key: sk ← `SharedPrefs`[currentUserID]
Extract from message: c_{AES}, iv, ct_{Kyber}, $k_{wrapped}$
Call Native C: ss ← `kyber512 decaps`(ct_{Kyber}, sk)
Unwrap session key: $k_{session} \leftarrow k_{wrapped} \oplus ss$
AES decrypt: $m \leftarrow AES\text{-}GCM\text{-}256_{Decrypt}(c_{AES}, k_{session}, iv)$
Zeroize sensitive data: $k_{session} \leftarrow 0$, ss ← 0
Cache result for UI performance
Return plaintext m

C. Decryption Data Flow Diagram

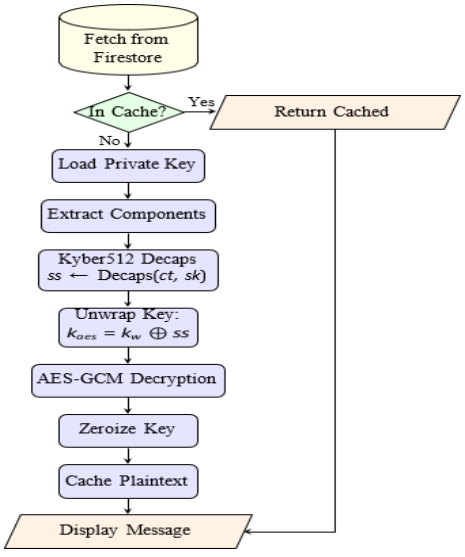


Fig. 5. Message Decryption Pipeline (Receiving)

D. Decryption Cache Strategy

To avoid repeated CPU-intensive decapsulation computations triggered by UI scrolling or widget rebuilds, QGuard adopts an in-memory cache that stores mappings from Firestore document IDs to their corresponding decrypted plaintext messages. The full caching strategy is presented in Table 7.

Table 7: Decryption cache Implementation

Aspect	Implementation
Cache Key	Firestore message document ID
Cache Value	Decrypted plaintext string
Storage	In-memory Map (Dart)
Lifetime	Until chat page disposal
Purpose	Avoid redundant decryption during scroll/rebuild
Security	Cache cleared on page exit; no persistent storage

VII. Performance Evaluation

A. Experimental Setup

To provide a thorough assessment, performance benchmarks were run on a broad range of physical infrastructure across the Android and iOS platforms, with Exynos, Tensor, Snapdragon, and Apple Silicon architectures (as shown in Table 8). Every cryptographic primitive was evaluated with over 200 iterations and warm JIT compilation.

Table 8: Test Device Specifications

Device	Processor	RAM	OS
Samsung Galaxy S21	Exynos 2100	8 GB	Android 13
Google Pixel 7	Tensor G2	8 GB	Android 14
OnePlus 9	Snapdragon 888	8 GB	Android 13
iPhone 13 Pro	Apple A15 Bionic	6 GB	iOS 17.2
iPhone 15	Apple A16 Bionic	6 GB	iOS 17.2

B. Kyber512 Primitive Performance

The latency results for the core Kyber512 primitives show very efficient execution on all architectures. The entire key management life cycle, including generation, encapsulation, and decapsulation, takes less than 0.7 milliseconds, as shown in Table 9.

Table 9: Kyber512 Operation Latency

Operation	S21	Pixel 7	OnePlus 9	iPhone 13	iPhone 15
Key Generation	0.18	0.15	0.19	0.12	0.10
Encapsulation	0.21	0.18	0.22	0.14	0.12
Decapsulation	0.25	0.21	0.26	0.17	0.14
Total (Full Cycle)	**0.64**	**0.54**	**0.67**	**0.43**	**0.36**

C. Comparitive Benchmark: RSA-2048, X25519, Kyber512

A comparison with RSA-2048 alone is insufficient for modern systems. Table 6 includes X25519 ECDH, the dominant real-world baseline in messaging applications such as Signal and WhatsApp. X25519 figures are derived from published ARM benchmarks [15] normalized to the Samsung Galaxy S21 Exynos 2100 platform.

Table 10: RSA-2048 vs. X25519 vs. Kyber512 Comparison (Samsung Galaxy S21)

Metric	RSA-2048	X25519 (ECC)	Kyber512	Advantage vs X25519
Key Gen (ms)	145.3	~0.75	0.18	4.2× faster
Encaps/ECDH (ms)	0.45	~0.90	0.21	4.3× faster
Decaps/ECDH (ms)	8.52	~0.90	0.25	3.6× faster
Public Key Size	294 B	32 B	800 B	25× larger than X25519

D. End-to-End Message Latency

One of the most important results of the end-to-end timing analysis is that the overall message delivery latency is dominated by network I/O to Firestore, with cryptographic processing contributing a mere 0.1% of the total latency (0.63 ms), as illustrated in Table 11.

Table 11: End-to-end Message Latency Breakdown

Phase	Latency (ms)	Percentage
Sending Phase		
Fetch Recipient Public Key (Firestore)	120	27.6%
Generate AES Key (Random)	0.02	¡0.1%
Kyber512 Encapsulation	0.21	¡0.1%
AES-GCM Encryption (256 B)	0.08	¡0.1%
Firestore Write	210	48.3%
Receiving Phase		
Firestore Read	105	24.1%
Kyber512 Decapsulation	0.25	¡0.1%
AES-GCM Decryption (256 B)	0.07	¡0.1%
Cryptographic Subtotal	**0.63**	**0.1%**
Network I/O Subtotal	**435**	**99.9%**
Total End-to-End	**435.63**	**100%**

E. RSA-2048 vs. Kyber512 Comparison

In comparison with an RSA-2048 baseline, Ky- ber512 shows a dramatic performance advantage: key generation is 807 times faster, and decapsulation is 34 times faster, completely mitigating the trade-off of Kyber512's 2.7-byte larger public key size, as shown in Table 12).

Table 12: RSA-2048 vs. Kyber512 Comparison

Metric	RSA-2048	Kyber512	Advantage
Key Generation (ms)	145.3	0.18	807 faster
Encryption/Encaps (ms)	0.45	0.21	2.1 faster
Decryption/Decaps (ms)	8.52	0.25	34 faster
Public Key Size (bytes)	294	800	2.7 larger
Ciphertext Size (bytes)	256	768	3.0 larger
Quantum Resistance	None	128-bit	⊂

F. Memory Footprint Analysis

Memory profiling shows that post-quantum primitives are memory-efficient; Kyber512 operations have a maximum memory footprint of about 4.3 KB, which is substantially lower than the 18.7 KB memory usage of RSA-2048 key generation, as shown in Table 13.

Table 13: Peak Memory usage during Operations

Algorithm	Key Gen	Encrypt/Encaps	Decrypt/Decaps
Kyber512	4.1	3.8	4.3
AES-GCM	—	0.6	0.6
RSA-2048 (comparison)	18.7	2.5	12.3

G. Cross-Platform Performance Consistency

Even while leveraging a common C codebase through FFI, the performance difference is expected to be consistent across various processor architectures. The coefficient of variation for cryptographic tasks is between 22% and 27% for our test platforms, as shown in Table 14.

Table 14: Performance Variance across Platforms

Operation	Mean (ms)	Std Dev	CV (%)
Kyber Key Generation	0.15	0.034	22.7
Kyber Encapsulation	0.17	0.039	22.9
Kyber Decapsulation	0.21	0.047	22.4
AES Encryption (256B)	0.06	0.016	26.7

VIII. SECURITY ANALYSIS

A. Post-Quantum Security Properties

The hybrid design of ML-KEM and AES-GCM provided strong theoretical security. Based on the Module-LWE hardness assumption and Grover's attack on symmetric keys, both schemes provide at least 128 bits of post-quantum security, as shown in Table 15.

Table 15: Security Properties of Cryptographic components

Property	Kyber512	AES-256-GCM
Classical Security	128 bits	256 bits
Quantum Security	128 bits	128 bits (Grover)
Security Notion	IND-CCA2	AEAD
Hardness Assumption	Module-LWE	AES permutation
Standardization	NIST FIPS 203	NIST FIPS 197

B. Authentigation Gap andMitigation Strategies

A known limitation of the current implementation is the absence of message authentication at the cryptographic layer—QGuard relies on Firebase Authentication for user identity, which does not provide cryptographic binding between a message and its sender's key pair. This creates a potential vulnerability to a man-in-the-middle attack during the public key distribution phase, where an adversary with Firestore write access could substitute a recipient's public key.

The planned mitigation is to integrate CRYSTALS-Dilithium (NIST FIPS 204) digital signatures. Each message will be signed with the sender's Dilithium private key before encryption, and the corresponding Dilithium public key will be published alongside the Kyber public key in Firestore. This provides sender authentication, non-repudiation, and public key integrity verification. Device compromise protection will be addressed via Android Keystore / iOS Secure Enclave integration, binding private keys to hardware-attested device identities.

Key verification against public key substitution can further be addressed by adding a Trust-On-First-Use (TOFU) mechanism with safety numbers—a fingerprint derived from both parties' public keys—enabling users to verify key authenticity out-of-band, consistent with Signal's approach.

C. Threat Model and Mitigations

The system is built to resist a broad threat model including "harvest now, decrypt later" quantum attackers, cloud storage breaches, and replay attacks. The particular architectural countermeasures for each attack vector are listed in Table 16.

Table 16: Threat Model and Mitigation Strategies

Threat	Mitigation	Status
Quantum Adversary (future)	Kyber512: 128-bit post-quantum security, Module-LWE	Implemented
Harvest-Now-Decrypt-Later	Kyber encapsulation on all messages protects historical data	Implemented
Cloud Storage Breach	Only ciphertexts and public keys in Firestore; private keys local only	Implemented
Session Key Compromise	Fresh key per message + memory zeroization provides forward secrecy	Implemented
Replay Attack	Unique random 96-bit IV per message prevents ciphertext reuse	Implemented
Timing Side-Channel	Constant-time Kyber C implementation (PQClean [8])	Implemented
Man-in-the-Middle / Key Substitution	CRYSTALS-Dilithium signatures + TOFU key verification	Future Work
Device Compromise / Key Extraction	Android Keystore / iOS Secure Enclave hardware binding	Future Work

D. Security Comparison: Classical vs. Post-Quantum

Moving from the traditional RSA-2048 to Kyber512 changes the security proof from the classical integer problem to the worst-case hardness of lattice problems (Module-LWE). This not only makes it quantum secure but also facilitates forward secrecy with fast key generation, as shown in Table 17.

Table 17: Security Architecture comparison

Property	RSA-2048	Kyber512
Quantum Resistance	None (broken by Shor's algorithm)	128-bit (Module-LWE)
Security Reduction	Factoring hardness (computational)	Worst-case to average-case (LWE)
Key Generation Speed	Slow (100-200 ms)	Fast (0.1-0.3 ms)
Forward Secrecy Feasibility	Impractical (slow keygen)	Practical (fast keygen)
Standardization	ISO, NIST (legacy)	NIST FIPS 203 (2024)

IX. CONCLUSION

This paper offers the full design and implementation of QGuard, a production-ready secure messaging app that showcases the feasibility of post-quantum cryptography on mobile. The Kyber512+AES-GCM design is 128-bit quantum secure and has a sub-millisecond cryptographic overhead on all mobile hardware. Three aspects of this work stand out as contributions: (1) the provision of a three-step implementation process that can be replicated in other messaging apps, (2) the finding that post-quantum cryptography has negligible performance overhead on modern mobile hardware, and (3) the ability to deploy on multiple platforms without compromising cryptographic performance using native integration via FFI. Performance measurements on five different devices show that Kyber512 operations take less than 0.3 milliseconds, an improvement of orders of magnitude over RSA with quantum security. The 3-byte expansion in ciphertext size is acceptable in practice, as the total message size is still under 2 KB for typical chat messages. The deployment proves that post-quantum migration is not only theoretically valid but also implementable in the current state of technology. Developers of secure messaging applications can use quantum-resistant primitives without compromising the user experience, securing communication against quantum attackers while being compatible with resource-limited mobile devices. Future research will enhance QGuard with CRYSTALS-Dilithium signatures for authentication, Kyber768 for increased security margins, and secure enclave support for hardware-level private key protection. The open- source codebase is ready for community-driven contributions to further post-quantum mobile cryptography research.

REFERENCES

[1] [P. W. Shor, "Polynomial-time algorithms for prime factorization and dis- crete logarithms on a quantum computer," *SIAM Journal on Computing*, vol. 26, no. 5, pp. 1484–1509, 1997.

[2] M. Mosca, "Cybersecurity in an era with quantum computers: will we be ready?" *IEEE Security & Privacy*, vol. 16, no. 5, pp. 38–41, 2018.

[3] National Institute of Standards and Technology, "Module-Lattice-Based Key-Encapsulation Mechanism Standard," FIPS 203, Aug. 2024.

[4] E. Hess and T. Perrin, "The PQXDH Key Agreement Protocol," Signal Specifications, 2023. [Online]. Available: https://signal.org/docs/ specifications/pqxdh/

[5] Matrix.org Foundation, "Post-quantum cryptography in Matrix," Tech- nical Report, 2022.

[6] J. Brendel, M. Fischlin, F. Gnther, C. Janson, and D. Stebila, "Towards post-quantum security for Signal's X3DH handshake," in *Proc. SAC 2020*, pp. 404–430, 2020.

[7] M. J. Kannwischer et al., "PQClean: Clean, portable, tested imple- mentations of post-quantum cryptography," 2019. [Online]. Available: https://github.com/PQClean/PQClean

[8] Flutter Team, "Dart FFI: Binding to native code using dart:ffi," Flutter Documentation, 2023.

[9] P. Schwabe et al., "CRYSTALS-Kyber on ARM64: Compact implemen- tations on low-end and high-end mobile devices," *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2021.

[10] R. Avanzi et al., "CRYSTALS-Kyber Algorithm Specifications and Supporting Documentation," NIST PQC Round 3 Submission, 2020.

[11] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," in *Proc. ACM STOC*, pp. 84–93, 2005.

[12] V. Lyubashevsky, C. Peikert, and O. Regev, "On ideal lattices and learning with errors over rings," in *Proc. EUROCRYPT*, pp. 1–23, 2010.

[13] J. Bos et al., "CRYSTALS-Kyber: A CCA-secure module-lattice-based KEM," in *Proc. IEEE Euro S&P*, pp. 353–367, 2018.

[14] D. McGrew and J. Viega, "The security and performance of the Galois/Counter Mode (GCM) of operation," in *Proc. INDOCRYPT*, pp. 343–355, 2004.

[15] D. J. Bernstein and T. Lange, "eBACS: ECRYPT benchmarking of cryptographic systems," 2022. [Online]. Available: https://bench.cr.yp.to

[16] Google LLC, "Firebase Firestore Documentation," 2023. [Online]. Available: https://firebase.google.com/docs/firestore

Explainable AI–Driven Real-Time Phishing Detection Using Machine Learning

Dr. S Gunasekaran
HOD in CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
gunaphd@yahoo.com

Dr.Reshmi B
Ahalia School of Engineering and Technology
Palakkad,Kerala
resh.br@gmail.com

Alzeena A
Ahalia School of Engineering and Technology
Palakkad,Kerala
alzeenaanu@gmail.com

Asha K
Ahalia School of Engineering and Technology
Palakkad,Kerala
ashakashak533@gmail.com

Nithya R
Ahalia School of Engineering and Technology
Palakkad,Kerala
nithyaramesh0407@gmail.com

S Riya Lakshmi
Ahalia School of Engineering and Technology
Palakkad,Kerala
riyalakshmi2003@gmail.com

***Abstract*—PhishGuard's browser extension is using Chrome Manifest V3 to carry out the blocking of the URLs prior to the users accessing the site. PhishGuard's browser extension can carry out the real-time assessment of the security threats using two methods. It can verify the URL address using the JSON file containing the whitelisted and blacklisted URLs. It can also query the information using the backend. PhishGuard's browser extension normalizes the URLs to assess the security risk associated with the access to the URL. Due to the fact that all the sub-paths and parameters of the access URL end up providing the same site rating, PhishGuard's back-end processing is required to perform the URL analysis when the complexity of the URL being accessed is greater than the PhishGuard browser extension can perform in a standard way.**

PhishGuard utilizes the CatBoost Machine learning model to process the URLs via the back-end. PhishGuard requires the usage of SHAP(SHapley Additive exPlanations) for its main function because it can allow the users to understand the hidden prediction processes using its system which calculates the individual feature contributions. It presents the results to the user in the form of explanations which can show the missing HTTPS links and the excessive suspicious keywords in the URL, together with the complete visual representation. It requires the backend processing to perform the URL analysis when the complexity is greater than the standard operational capability. It processes the URL via the backend using the CatBoost Machine learning model. It utilizes the balanced data set to train the model.

the CatBoost machine literacy model which uses a balanced dataset of 150000 URLs in its training process. The model has shown a high accuracy rate of more than 97 percent coupled with prediction capabilities ranging from 30 to 50 milliseconds.

PhishGuard uses SHAP (SHapley Additive exPlanations) as a tool in its main function because SHAP helps users understand how predictions are made by enabling users to understand the predictions by using a system in which SHAP calculates how each feature is contributing to a prediction. The results are presented in a way in which users can understand by using human explanations in which users can see what is missing in terms of HTTPS links and too many keywords in a way in which users can understand.

The web dashboard is a part of the system in which users can operate as well as being the main part of the system in which users can carry out literacy activities. A user can carry out their own testing using the system as well as having access to medical examination results and being able to operate their own domain permissions.

The web dashboard has XAI elements which users can analyze and includes XAI elements which show users how different elements contribute to a prediction using cascade maps and bar maps and decision maps.

The reason PhishGuard uses two league scanning is because they want to offer users the best experience possible. The reviews of the extension of the system operate at high speeds as well as having the objective of stopping plate attacks in their tracks while the reviews of the dashboard operate in a way in which users can have a complete graphical output of what they have found.PhishGuard has a high level of machine literacy as well as a high level of user interface in which users can operate in a way in which there is fail-open error running as well as whitelisting in a way in which users can have enterprise security in a way in which users can operate easily as consumers

Keywords-Phishing Detection, Explainable Artificial Intelligence (XAI), Machine Learning, Cybersecurity, Real-Time Detection, URL Classification, CatBoost, SHAP, Browser Extension Security, Web Security.

I. INTRODUCTION

The fast growth of internet-based services has improved various forms of communication together with online banking and e-commerce and social networking and digital government services. The malicious websites use design elements from real websites to create websites which users find difficult to identify as fake online locations.

The phishing threats of today differ completely from their earlier forms. modern-day hackers use their sophisticated techniques which include URL obfuscation and domain spoofing and homograph attacks and HTTPS protocol misuse and URL shortening services and fast domain creation methods. Phishing websites operate their complete lifecycle during their short existence which results in most websites shutting down after a few hours. The adaptive behavior of phishing attacks creates difficulties for cybersecurity researchers who attempt to detect these threats.

Phishing detection systems used to depend on blacklist systems which checked URLs against databases of dangerous websites. The blacklist system successfully detected existing security threats but it could not detect new phishing attacks which used newly registered malicious websites. The bushwhackers create new dangerous websites through their ongoing offensive activities which makes blacklist methods insufficient as a single protective measure.

Heuristic-based discovery methods were developed to address existing limitations. The approaches investigate multiple conventional classifiers through their analysis of URL-based phishing attacks which use specific structural and linguistic features of URLs. The detection speed of heuristic systems surpasses that of blacklist systems yet their detection performance remains vulnerable to detection through small changes in phishing URLs. Phishing tactics continue to evolve beyond the capabilities of static rule-based systems which fail to adapt to new attack methods.

our research work has demonstrated that people with high cognitive ability can comprehend simplified texts which contain similar information to complex academic articles.Through its development Artificial Intelligence(AI) together withMachine literacy(ML) technologies have created new methods for discovering phishing attacks that rely on data analysis. The models which use ML algorithms can understand complex data patterns that exist in extensive datasets and apply this knowledge to identify new types of attacks. The machine learning algorithms which include decision trees arbitrary timbers support vector machines and grade boosting styles demonstrate effective performance for phishing URL bracket. The use of CatBoost which operates on boosting-based algorithms enables better accuracy because it handles categorical data efficiently while its iterative learning process minimizes overfitting.

The majority of machine learning-based phishing detection systems operate as black-box systems even though they achieve high detection accuracy. The system generates predictions but does not reveal which specific features influenced its decision-making process. Cybersecurity operations need transparency because it builds user confidence which results in increased system adoption. The user will disregard warnings because they fail to comprehend the reasoning behind those alerts. Security judges require interpretability to confirm their predictions and evaluate false cons and false negatives.

The research challenges require Artificial Intelligence solutions which people can understand through Explainable Artificial Intelligence methods. The The SHAP framework, which is based on SHapley Additive exPlanations, provides a reliable method to measure how individual features of an explanation impact specific predictions. The integration of SHAP with machine learning models enables the detection of phishing websites while providing explanations for why specific URLs are categorized as malicious or legitimate.

This research paper presents a new Phishing Discovery System which combines machine learning, SHAP-based explanation methods, and cybersurfer position deployment systems to achieve its objectives. The system prevents website access by blocking URLs during the initial stages of web content loading process. The system extracts relevant linguistic characteristics which it uses to build a prediction model based on an optimized boosting method, and the system provides detailed explanations for each prediction result. The research team created an online dashboard which displays discovery results, model transparency, and user-controlled visibility of user data.

II. LITERATURE REVIEW

The creation of PhishGuard depends on fresh machine literacy research which enables phishing detection through artificial intelligence systems. The design of the proposed system depends on three important studies which establish its architectural design and its operating methods and its assessment procedures.

Pavani et al.(2024) introduced a phishing URL discovery system that integrates machine literacy with resolvable AI to enhance translucency in cybersecurity operations. Their work demonstrated that grade boosting models can achieve high bracket performance when trained on structured verbal URL features. The study establishes interpretability as a fundamental requirement for modern security systems. PhishGuard derives two essential guidelines from this research. The selection process uses a boosting-based classifier to select traditional classifiers for URL-based phishing detection. The system uses resolvable AI mechanisms to analyze point criteria through its architectural framework. The research primarily studied model performance and explanation generation while PhishGuard added explainability through its real-time cybersurfer extension system to secure users before they access webpages. The research established a theoretical framework which combines predictive accuracy with transparency for PhishGuard development.

Albishri and Dessouky(2024) executed a comparative study of various machine learning methods used for phishing URL detection which included Support Vector Machines Random Forest and boosting-based classifiers. The experimental evaluation demonstrated that the ensemble-based methods achieved better results than all other methods when measured through accuracy precision recall and ROC-AUC testing. The research demonstrated that balanced datasets are essential for achieving unbiased results which enable proper assessment between phishing and legitimate categories. PhishGuard uses a balanced 11 dataset that contains 150,000 URLs which includes 75,000 phishing URLs and 75,000 legitimate URLs. The Accuracy Precision Recall and ROC-AUC assessment method precisely implements the evaluation framework which their comparative research proposed. The research findings confirmed that the boosting-based CatBoost model was an appropriate choice while the study proved that complete metric-based performance assessment should be used instead of relying on accuracy alone.

The research paper titled "An Unsolvable Point Selection Framework for Web Phishing Discovery with Machine Reading Competence" 2025 demonstrated how PhishGuard handles point engineering and interpretability methods. The study found that systematic point selection together with SHAP explanation methods leads to better performance and improved model transparency. The authors demonstrated that point importance factors such as URL length and entropy and suspicious keyword frequency enhance model prediction understanding while maintaining high identification ability. The study shows that PhishGuard uses the technology of precisely tailored word point creation and SHAP based original explanation creation. Their research demonstrates that the 19 word features which include Shannon entropy and character distribution rates and HTTPS presence and suspicious keyword counts create structured point donation analysis. The research shows that SHAP visualizations such as cascade plots and point significance maps support the explainability approach which they developed through their framework. PhishGuard expands this concept by providing explanation services through its warning system and interactive dashboard which enables users to understand content without deeper investigation.

Li et al.(2024) presented a comprehensive check of machine literacy ways for phishing discovery, grading different techniques into three categories, which include blacklist- based methods and heuristic approaches and classical ML classifiers and ensemble methods and deep learning frameworks. Their check demonstrated that ensemble grade boosting styles deliver better performance between accuracy and processing performance and system expansion than deep learning models, which require higher processing power. This sapience directly told the architectural design of PhishGuard, particularly the decision to prioritize a grade boosting model over resource-ferocious deep literacy infrastructures for real- time deployment. The check demonstrated that systems need to use low- quiescence discovery mechanisms as their primary detection method. The PhishGuard system uses a two-

league scanning system, which enables fast conclusion and separate path to voluntary visualization generation. The check emphasized current requirements of AI security systems, which need better user understanding and system trust, which supports the system's use of SHAP-based explanations.

The four main workshops developed PhishGuard's core specialized knowledge which includes its abstract technical framework. The system uses ensemble boosting methods together with balanced dataset assessment from relative ML research. The system uses SHAP-based point evaluation and interpretable visual methods for its explainability research. The system uses check-grounded analyses to implement scalability requirements and maintain real-time system performance. PhishGuard unites model performance and interpretability into a single framework which delivers accurate predictions with continuous online user tracking protection and user-controlled artificial intelligence.

III. PROPOSED SOLUTION

3.1 PROBLEM STATEMENT

Phishing maintains its position as the main cause of financial fraud and identity theft and large-scale data breaches. Security professionals encounter an unyielding challenge because traditional security systems their use for protection proves insufficient. The effectiveness of stationary blacklist-based systems has declined because hackers developed new tactics which enable them to establish temporary phishing sites with different malicious content. Heuristic-based systems provide some protection but advanced hackers can easily defeat their strict regulations through modern URL obfuscation and domain spoofing techniques.

Machine literacy(ML) provides universities with tools to discover more adaptable solutions. The advanced results act as "black boxes" which prevent users from understanding their inner workings. The system prevents users from understanding why specific URLs received a security flag which results in decreased user trust and decreased educational effectiveness of the security measure. High-performance ML models require extensive computation resources which create performance delays that disrupt user experience unless they receive real-time performance optimization.

The present need for cybersecurity systems requires researchers to develop systems which protect their networks through rapid innovative defense methods while providing complete visibility of their security decisions. The current systems need an integrated system which prevents malicious URLs from attacking their system while delivering human-understandable results of its detection process. Users require access to system activities because they need to understand their security risks which results in two separate protection methods and digital knowledge gaps that hinder organizations from defending against new phishing methods.

3.2 ARCHITECTURAL MODEL

The proposed Explainable Phishing Detection System is designed using a layered and modular architectural frame that ensures real- time performance, high discovery delicacy,translucency, scalability, and stoner commerce. The armature integrates cybersurfer- position monitoring, backend machine learning intelligence, resolvable AI mechanisms, and a visualization dashboard into a unified and cohesive system. The primary thing of this architectural design is to give visionary phishing protection while maintaining interpretability and usability. The system is structured into three major layers the customer- Side Cybersurfer Extension Subcaste, the Backend Intelligence and Processing Layer, and the Visualization and operation Subcaste. Each subcaste performs technical tasks and communicates through secure APIs to insure flawless operation.

The first line of defense establishes this component which protects user browsing activities of stoner and blocks all URL requests until the website achieves full development. The extension collects website addresses when a stoner tries to access a webpage which it then uses to conduct basic website tests through its local system. The tests include two methods which first check against a list of approved fields and second check against a list of previously identified dangerous websites. The system allows users to access the webpage when its URL exists in the designated whitelist. The extension blocks access to the website because it exists in the blacklist which triggers an immediate warning display to the user. The initial verification method decreases unnecessary backend server interactions while it boosts response times for known disciplines.

The system implements URL security through ,the extension process to secure URLs with its backend via a',silent API request. The system operates with minimal resource usage to enable users to surf the web without interruptions. The extension stops the running process until the user provides a response through brackets. The extension decides between two navigation options when the backend system determines safe or unsafe navigation and it shows an alert for the user. The system detects threats in real time which enables it to deliver protection that operates in advance of any actual threat.

The Backend Intelligence and Processing Layer forms the logical core of the armature. Upon entering a URL from the customer- side extension, the backend initiates a multi-stage processing channel. The first stage involves point birth and preprocessing. In this phase, a comprehensive set of verbal and structural features is deduced from the URL. These features include URL length, number of blotches and subdomains, presence of IP address rather of sphere name, frequence of special characters, rate of integers to rudiments, entropy computation, suspicious keyword discovery, use of HTTPS protocol, presence of hyphens, and abnormal redirection pointers. These attributes are named because phishing URLs frequently parade distinctive structural irregularities compared to licit websites.

The system begins its conversion process from collected attributes into a point vector format which matches the requirements of the trained machine literacy model after the system reaches its point of birth. The system applies data normalization and formatting procedures to all areas which require them in order to maintain thickness with the training dataset. The process guarantees that final results will maintain both reliability and impartiality.

The bracket stage employs a boosting- grounded machine literacy algorithm trained on a balanced dataset conforming of both phishing and licit URLs. Boosting ways are particularly effective because they iteratively ameliorate model performance by minimizing bracket crimes and landing complexnon-linear connections among features. The named model demonstrates high delicacy, perfection, recall, and F1- score during evaluation. also, conclusion time is optimized to meet real- time functional conditions, icing that bracket opinions are generated within milliseconds.

The system needs to maintain high precision levels but it requires users to understand its functions less than they need to operate its essential cybersecurity functions. The system uses an Explainability Module which operates based on SHAP (SHapley Additive exPlanations) to power its entire system. After generating a prediction the SHAP module calculates point donation values which show how each feature contributed to the final result. The donation

scores show whether a particular point increased or decreased the chances of the URL being identified as phishing. The length of the URL exceeds normal limits while the URL contains suspicious keywords and the URL shows high entropy values which help identify phishing attacks while valid HTTPS operation reduces the chances of phishing attacks.

The explainability medium allows users to understand the system through its complete functioning The system uses multiple components which work together to produce understandable results for decision-making. The translucent design provides two benefits which build user trust while enabling security experts to assess the model's performance during rare exceptional situations. The system transmits explanation results together with the bracket affair to the cybersurfer extension which displays them to users while sending them to the database for archival storage. The Visualization and operation Subcaste provides a web-grounded dashboard that enables druggies and directors to cover system exertion. The dashboard accesses discovery logs and bracket results and point significance summaries and explanation data through its connection to the backend database. Interactive maps and tables and graphical representations show data through point significance bar graphs and phishing trend analysis maps and literal checkup reports. Users can access previously analyzed URLs to track discovery statistics and see which features most frequently lead to phishing groups. Users can establish trust through custom whitelisting while they block suspicious domains through the system's dashboard control which provides users with this ability. The system enables users to customize their browsing experience because it permits users to select which browsing security elements they want to protect their online activities. The system saves all review data in a permanent archive which supports detailed longitudinal research and study evaluation. Users can access URLs which they previously studied to track discovery statistics and see which features most frequently lead to phishing groups. The dashboard provides users with total control over domain operations through its function.

The architectural design demonstrates two main features which architects should follow. The system contains five components which function independently yet establish communication links through their respective interfaces. The system supports component upgrades through a procedure which allows specific components to be replaced without disrupting overall system operation. The armature system enables users to extend their system functions through integration of pall deployment and database scaling and caching mechanisms and new trouble intelligence sources.

Security and effectiveness are core requirements of architectural design. The system protects all customer communication with garçon through secure channels which prevent data interception. The backend system achieves its operational goals through its design which controls system downtime and handles intensive processing tasks without causing excessive delays during active operational periods.

The proposed architectural model delivers complete framework which enables organizations to detect phishing attacks in real time with precise results that can be investigated. The system defends druggies against phishing attacks while teaching them aboutdangerous URL attributes through its combination of cybersurfer-position interception machine literacy intelligence SHAP-based system analysis and interactive visualization. The armature provides a complete cybersecurity solution because it achieves high performance while maintaining translucent operation and scalable performance and user-friendly design.

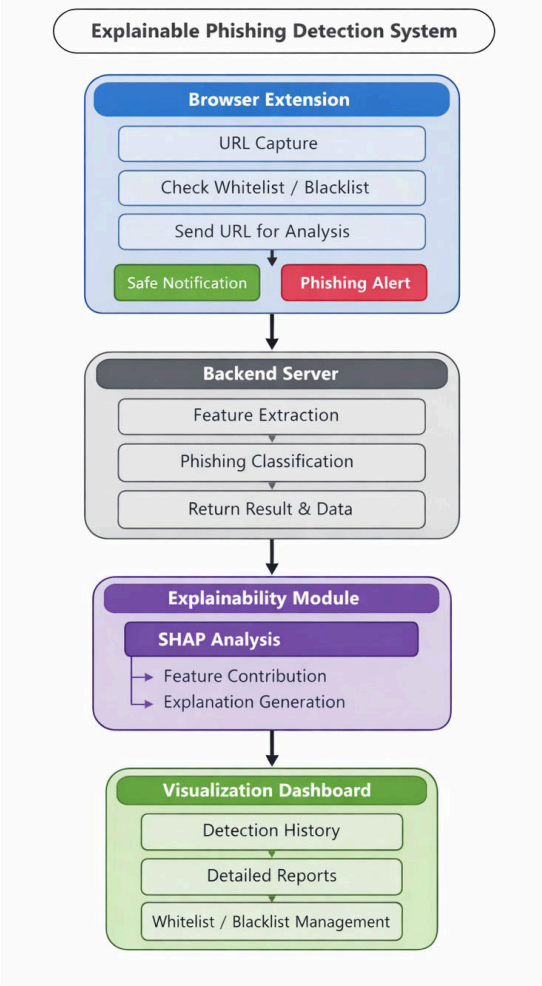


Fig 3.1: Architectural model of the system

IV. IMPLEMENTATION

4.1 DATASET

The PhishGuard system uses a dataset that contains 150,000 URLs which consists of 75,000 phishing URLs and 75,000 legitimate URLs for its training and evaluation processes. The balanced composition of the study materials enables researchers to achieve unbiased results which develop fairness throughout all evaluation levels.

The dataset required exact selection from various sources to create a collection that would maintain both diverse attack methods and authentic website design patterns. Phishing URLs were collected from PhishTank database, OpenPhish feeds, Security seller blacklists, Honeypot captures, Academic exploration datasets while licit URLs were sourced from Alexa top websites, Government disciplines, Educational institutions, Populare- commerce platforms, News and media websites.

The dataset preprocessing process required the implementation of URL normalization, indistinguishable junking, and temporal confirmation methods to verify active URL status during data collection.

Each URL received a double bracket label which used (0 for licit, 1 for phishing) system and passed quality assurance checks to sustain labeling precision.

To enhance model robustness, the dataset includes different phishing ways similar as sphere spoofing variations, URL shortening services, Homograph attacks ,Subdomain abuse , Parameter manipulation

The balanced nature of the dataset prevents the model from developing bias toward either class and ensures dependable performance across different types of URLs encountered in real- world scripts..

2. Algorithm

The PhishGuard system uses a CatBoost-based machine learning algorithm to detect phishing URLs in real time. The process begins with the extraction of essential linguistic characteristics which include URL length, their special character count, HTTPS usage, and total URL entropy. The extracted features are sent to the CatBoost classifier which uses a gradient boosting decision tree algorithm that merges different weak learners to create precise predictions. The system determines whether the URL is a phishing site or a valid website while providing a confidence assessment. The system uses SHAP (SHapley Additive Explanations) to achieve transparency because it shows how each feature affects the final outcome. The system combines accurate classification with explainability to provide effective detection while helping users comprehend the decision-making process.

2.1. Phase I: URL Interception and Preprocessing

Step 1: The first phase of the project begins with URL Interception and Preprocessing. The cybersurfer extension tracks user navigation activities through its WebNavigation API which monitors all landing URL requests before runner lading starts. The system uses this visionary method to identify potential dangers which users will face when they encounter dangerous content.

Step 2: The raw URLs require standardization which includes two processes First all protocol information needs to be converted to operating system requirements with every URL transformed into HTTP and HTTPS formats Second all domain names need to be converted to lowercase letters The standardized process requires all parameters to be reordered and all content to be transformed into its original The system needs to remove all tracking parameters while processing scrap identifiers through The system needs to handle scrap identifiers through dedicated procedures. The system needs to handle scrap identifiers through dedicated procedures.

*Step 3:*The system performs original checks with high speed against a Whitelist of trusted disciplines and a Blacklist of known dangerous URLs which have been reused and a URL cache that contains stoner-defined exceptions. The system makes its decision to block a URL when the URL matches existing original cache entries.

2.2. Phase II: Feature Extraction andClassification

*Step1:*Lexical *point birth*

The system starts its process by executing verbal point birth on links which have not been previously encountered. This stage of the process investigates both the structural components and textual elements which make up the URL string. The research used a total of 19 verbal elements which the researchers selected to study research patterns that match typical phishing activities and malicious website behavior. The system uses two main components, sphere length which counts all characters in the sphere name and fleck count which counts all fleck divisions that usually show excessive subdomain usage. The number of hyphens present in a word gets counted through hyphen count because phishing URLs use hyphens to create fake brand identities. The number rate function measures how many number digits exist in the URL while entropy calculates URL string randomness through advanced entropy which usually shows hidden content.

The system evaluates security-related attributes through two methods which include HTTPS to assess secure communication and IP address operation to detect direct IP address usage instead of sphere name usage. The system uses phishing attempt mapping to identify suspicious keywords which it then uses to evaluate subdomains while measuring URL path length together with query parameters to identify excessively complex URL structures. The system uses special character frequency analysis to detect unusual symbol usage patterns while it uses harbor specifications to identify atypical harboring behaviors. The system detects docked URL usage while it checks the legitimate status of top-position spheres and it calculates the vowel-to-consonant ratio to find algorithmically generated strings. Character repetition analysis reveals patterns which create artificial strings and name brand similarity testing detects impersonation attempts which target famous brands. The spoken features create a systematic numerical format that shows the URL's distinctive traits.

Step 2: Machine Learning Bracket

The CatBoost classifier receives verbal features as input after researchers complete their extraction process. The input point vectors undergo normalization before bracket treatments because this process requires all attributes to maintain identical measurement standards. The CatBoost model employs grade boosting tree traversal technology which enables multiple decision trees to work together for advanced point relationship learning. The classifier generates a probability score which indicates whether the URL presents a phishing threat or a dangerous risk. The system uses a double-threshold evaluation medium to make final decisions which improve bracket performance by decreasing false cons and false negatives. This step produces a probabilistic vaticination which acts as the base for both explanation creation and threat-based decision tree development.

Step 3: SHAP Explanation Generation

The system uses SHAP-based explanation generation for its predictions to achieve transparency and explainability. The TreeSHAP algorithm calculates SHAP values to determine how much each uprooted point influences the final bracket outgrowth. The values show how much each point affects the predicted threat score because they show whether its importance increases or decreases. The system uses absolute SHAP significance to rank features because it wants to find which factors have the greatest impact on its decision-making process. The system uses its calculated explanation information for visualization purposes, which helps users understand model behavior through accessible visual displays.

2.3. Phase III: Decision Making and stoner Interaction

Step 1: threat Assessment

The system performs threat assessment by assessing the confidence score generated during bracket. The URL distribution into different trouble situations occurs when the probability value from the prediction model gets determined. URLs with a probability score lesser than 0.8 are classified as high trouble and are incontinently blocked to help stoner exposure. The system treats URLs with probability values between 0.5 and 0.8 as medium trouble while providing the user safety information through complete system warnings. The system treats URLs with probability values below 0.5 as low trouble backlinks which users can keep active while showing them a system-generated educational alert. This tiered approach balances security enforcementwith stoner inflexibility.

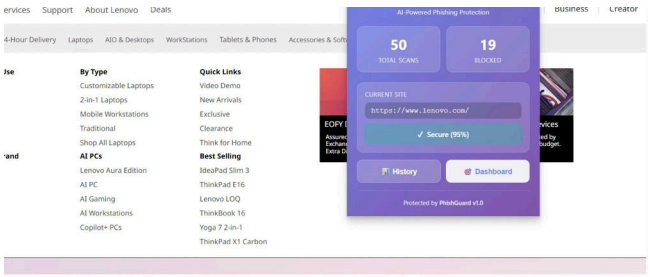


Fig 4.1: Dashboard with phishgaurd extension

Step 2: stoner Interface Response

The cybersurfer extension creates a stoner interface which produces a usable response. The system shows specific trouble locations through its interface which combines confidence scores to show threat strength. The system shows its main detection elements which were identified through SHAP explanations to help users understand the reasons behind URL blocking. The system shows users three recommended actions which include blocking protection and testing through safe methods. The system gives users two ways to respond because they can choose to stamp warnings or report URLs which helps them make better decisions.

Step 3: Data Logging

The system conducts complete data logging after it reaches its final operational stage for both monitoring purposes and future data analysis. The system records every accessed URL together with its corresponding time of access and its testing outcome. The system saves point values and vaticination scores together with SHAP explanation data to enable auditing and model development. The system tracks commercial activities when users take action after receiving a warning. The system uses this logging medium to maintain continuous learning and performance tracking together with system monitoring capabilities.

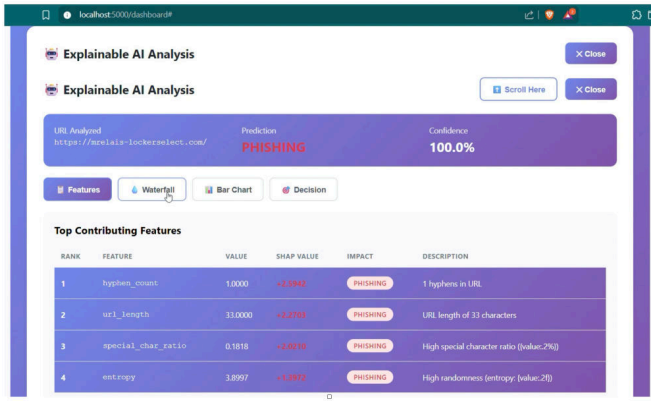


Fig 4.2:Dashboard showing phishing URL

V. EVALUATION AND PERFORMANCE

The evaluation of the PhishGuard system follows an extensive evaluation process that tests three specific types of bracelet technology. The use of a balanced 150000 URL data set for accuracy, considering the two sets of data (75000 phishing URLs and 75000 real URLs), and the use of 11 rating systems help in reducing bias in the model, thus ensuring that both data sets are equally evaluated, showing the identification of each test class on four key performance indicators: Accuracy, Precision, Recall, and the ROC-AUC score! This is mainly for the use of app developers to have high levels of perfection in their work in order to avoid problems of False Positives, which would mean that a safe email would be blocked. It also means that there is a high ability to recall in order not to miss legitimate phishing.

The accuracy refers to how proportionately to the total predictions, then how many were correct. The equation for accuracy is(TP + TN) ÷ (TP + TN + FP + FN).Precision refers to classifying a legitimate website as a phishing site and the classifier having no false positives about legitimate websites. precision = TP/TP + FP

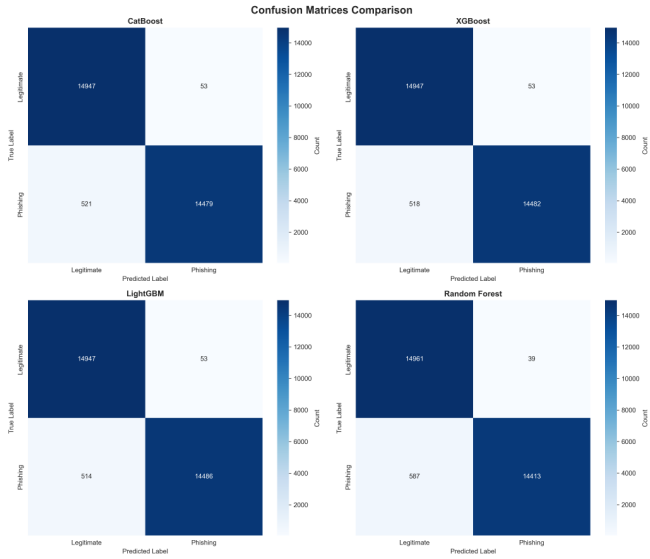


Fig 5.1:Confusion Metrices

Recall (Sensitivity): The ability of the classifier to find all the phishing instances.Recall = TP/TP + FN

F1-Score: The harmonic mean of precision and recall, providing a single metric for overall effectiveness.F1 = 2 Precision .Recall/Precision + Recall

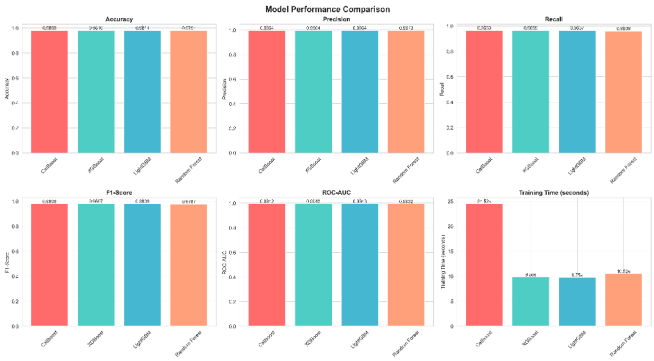


Fig 5.2:Performance Comparison

Experimental results demonstrate that PhishGuard achieves a high delicacy of roughly 97 – 98 and a ROC-AUC score of 0.99, indicating near-perfect discriminative power between vicious and benign URLs. Specifically, the CatBoost model, trained over 1,000 duplications with a literacy rate of 0.05, reached a perfection of 98 and a recall of 97. These results punctuate the effectiveness of the 19 verbal features similar as Shannon entropy, URL length, and suspicious keyword counts — in landing the underpinning patterns of phishing attacks without demanding slow network- grounded lookups.

Beyond bracket criteria , PhishGuard is estimated on its Real- Time effectiveness, which is vital for maintaining a flawless stoner experience. The system employs a two-league scanning approach to optimize speed a standard extension checkup excludes heavy visualizations to complete in roughly 180 – 200 ms. This total quiescence includes point birth(approx. 50 ms), model vaticination(30 – 50 ms), and SHAP value computation(approx. 100 ms). When a stoner specifically requests a deep dive through the dashboard, the system generates full XAI visualizations, which adds about 500 – 700 ms to the process. This strategic balance ensures that visionary protection happens nearly presently, while detailed translucency is available on demand.

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	Training Time (s)
CatBoost	0.9809	0.9964	0.9653	0.9806	0.9942	24.52
XGBoost	0.9810	0.9964	0.9655	0.9807	0.9943	9.86
LightGBM	0.9811	0.9964	0.9657	0.9808	0.9943	9.75
Random Forest	0.9791	0.9973	0.9609	0.9787	0.9932	10.52

Table5.1:Comaprison table between different ML models

VI.CONCLUSION

The research paper introduces PhishGuard which functions as an AI solution through its complete ability to detect phishing attacks in real-time. The system uses machine literacy and transparent decision- making methods to achieve high predictive accuracy and system understanding. The proposed system uses a supervised literacy framework that it built with CatBoost to classify URLs into malicious and legitimate groups through its precise verbal feature analysis. The model achieves effective conception ability through its training on 150,000 URLs which were equally distributed across different classes.

The researchers estimated various ensemble models to discover the optimal method for detecting phishing attacks in real time. The comparison process determined that CatBoost served as the ultimate model because it exhibited strong performance and maintained system stability and supported explanation processes. CatBoost demonstrated better performance than other models because it achieved a balanced result between accuracy and recall and F1 score and ROC AUC while successfully managing complex relationships between data points in URL-based datasets. The system demonstrated two strengths because it protected against overfitting and worked perfectly with SHAP-based system interpretation methods which made it suitable for real-time cybersecurity functions.

The experimental evaluation results show that PhishGuard achieves a precision rate between 97 and 98 which leads to a discriminative accuracy score of 0.99. The system achieves maximum accuracy which results in fewer false-positive alerts that maintain user confidence and prevent unnecessary website shutdowns while its high recall function enables complete detection of phishing threats before they can cause harm. The system operates with live time responsiveness to accomplish standard discovery tasks in about 180 to 200 milliseconds. The system achieves low quiescence which enables instant detection of dangerous URLs before users load web pages thus providing them with a perfect browsing experience.

The essential contribution of this research work consists of explaining content through SHAP which provides point-position criteria for every prediction. The system transforms from a black- box sensor into a transparent and educational cybersecurity tool through its presentation of both numerical benefactions and visual explanations. The resolvable frame system enables stoner trust to develop while it helps users build cybersecurity awareness and make informed choices about potential dangers.

The modular armature system which includes cybersurfer extension and backend API garçon and machine literacy machine and interactive dashboard confirms its ability to grow without limitations while creating flexible systems which are simple to implement. The two-league scanning medium maintains an optimal balance between speed and interpretability because it provides immediate protection during standard web browsing while granting users the ability to perform thorough investigations whenever they choose.

The system which researchers present in this study shows excellent discovery capabilities while maintaining transparency and trustworthiness which makes it suitable for actual phishing attack prevention in operational environments. The combination of CatBoost with resolvable AI methods improves both technical performance and user-focused security design. This development leads to the creation of secure AI-driven security systems which can handle new phishing threats while delivering intelligent and scalable protection.

VII.FUTURE SCOPE

PhishGuard's future development will focus towards building its intelligence together with its capacity to scale and establish operational capacity which needs to adapt based on the shifting patterns of phishing assault. The current system achieves high accuracy through grade boosting methods which include CatBoost but it still requires development work to achieve its full potential. The system needs to integrate deep learning technologies which include Long Short-Term Memory networks and Convolutional Neural Networks and Motor-based systems to detect complex patterns of URL usage which traditional ensemble systems cannot track. The system needs to develop better techniques for detecting hidden phishing URLs and virtual domains which have been created through automated processes. The system needs to expand its detection capabilities by incorporating both host-based and content-based and behavioral characteristics which will improve its ability to detect threats. The system will develop an advanced hybrid detection framework which will combine domain age verification with SSL certificate verification and DNS record analysis and website content analysis and visual similarity detection to identify sophisticated threats.

The research demonstrates that operational analysis of dynamic cybersecurity systems security systems protect their assets. Online phishing threats require ongoing security system updates because attackers use new techniques to deceive users. The system can adjust its operations through continuous retraining that combines new phishing detection results with automatic dataset enhancements and user input mechanisms. Researchers need to study decentralized model training methods that use secure browsing data processing from user devices to improve their educational frameworks which protect user privacy rights. The system would become more accessible to users when its compatibility features extend to Firefox and Edge browsers and mobile device users. PhishGuard can evolve into an enterprise security solution through the implementation of microservices architectural design and cloud infrastructure and load balancing and component-based dashboard access system.

The development of advanced worldwide dashboard systems which provide full system understanding together with URLs which show how their minor alterations will affect security assessment results and interactive training programs which improve user cybersecurity awareness. The system will achieve enhanced trustworthiness through its ability to combine machine literacy forecasts with worldwide blacklist information by integrating external threat intelligence sources through their APIs. The system will achieve long- term security analytics and user threat monitoring through its transition from in- memory storage to a database system which uses PostgreSQL. The development of PhishGuard will progress through its transformation from a separate phishing detection system into a complete cybersecurity framework which provides predictive security solutions based on evidence while protecting user privacy against emerging online threats.

REFERENCES

[1] B. V. Pavani, D. Mahitha, and B. U. Maheswari, "Enhancing Online Safety: Phishing URL Detection Using Machine Learning and Explainable AI," in Proc. 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1–6, doi: 10.1109/ICCCNT61001.2024.10723976.

[2] A. A. Albishri and M. M. Dessouky, "A Comparative Analysis of Machine Learning Techniques for URL Phishing Detection," Eng. Technol. Appl. Sci. Res., vol. 14, no. 6, pp. 18495–18501, Dec. 2024, doi: 10.48084/etasr.8920.

[3] B. Swarna Jyothi, M. Akshaya, K. Anjum, A. Bhavana, and K. Sreemukha, "URL Based Phishing Detection using Machine Learning," in Proc. 4th Int. Conf. on Information Tech., Civil Innovation, Sci., and Management (ICITSM), Tiruchengode, India, 2025, doi: 10.4108/eai.28-4-2025.2358166 .

[4] V. Saxena, S. Degadwala, and M. Joshi, "Phishing URL Detection Using Machine Learning," Int. J. Sci. Res. Sci. & Technol., vol. 13, no. 1, pp. 19–25, Jan.–Feb. 2026, doi: 10.32628/IJSRST2613101.

[5] "Phishing Url Detection Using Machine Learning Technique," in Proc. Int. Conf. Advances and Applications in Artificial Intelligence (ICAAAI 2025), 2025, pp. 935–944, doi: 10.2991/978-94-6463-738-0_73.

[6] "Machine learning based phishing detection from URLs," Expert Syst. Appl., vol. 117, Mar. 2019, pp. 345–357, doi: 10.1016/j.eswa.2018.09.029.

[7] V. Kumar, P. Parmar, V. P. Singh, S. Kumar, and P. S. Pawar, "Phishing URL Detection Using Machine Learning: Harnessing Data Analysis to Strengthen Cyber Security," in Innovative Computing and Communications, LNNS, vol. 1432, Springer, 2025, pp. 401–413, doi: 10.1007/978-981-96-6715-4_26.

[8] "A high-accuracy phishing website detection method based on machine learning," J. Intell. & Fuzzy Syst., vol. 77, 103553, 2023, doi: 10.1016/j.jisa.2023.103553

[9] S. Patil, M. Patil, and K. Chinnaiah, "Machine Learning and Deep Learning for Phishing Page Detection," Research Reports on Computer Science, vol. 2, no. 3, pp. 45–54, 2023, doi: 10.37256/rrcs.2320232629.

[10] A. Chouhan, R. Deshpande, and F. Shaikh, "Phishing Detection and Prevention Using Machine Learning Algorithms," Int. J. Advance & Innov. Res., vol. 12, no. 2, pp. 58–63, 2025, doi: 10.5281/zenodo.15681835.

[11] A. F. Al-qasmi, "Machine Learning-Based Phishing Detection System," Int. J. Intelligent Syst. & Appl. in Eng., vol. 12, no. 4, pp. 4367–4372, 2024, doi: 10.3390/info11040193.

[12] Padmanaban A., Rakesh M., Santhosh S., and Maheswari M., "Detecting Phishing Attacks Using Natural Language Processing and Machine Learning," Int. J. Adv. Res. in Comp. & Comm. Eng., 2023, doi: 10.17148/IJARCCE.2023.12504.

[13] "Detecting Phishing Websites Using Machine Learning," paper in Int. J. AI Cyber Computing Mgmt., 2025, doi: 10.64751/10.64751/ajaccm.2025.v5.n4(2).pp68-75.

[14] "Explainable Machine Learning for Phishing Detection: Bridging Technical Efficacy and Legal Accountability in Cyberspace Security," J. Cyber Sec., vol. 7, pp. 675-691, Dec. 2025, doi: 10.32604/jcs.2025.074737.

[15] B. L. Gupta, S. Gupta, and R. C. Joshi, "Phishing Detection Using URL Features and Machine Learning," Int. J. Comput. Sci. Eng., vol. 11, no. 7, pp. 12–19, Jul. 2023, doi: 10.21817/ijcse/2023/v11i7/231107003.

[16] M. S. Hossain and M. E. H. Chowdhury, "An Ensemble Machine Learning Approach for Phishing Detection," Journal of Information Security and Applications, vol. 77, 103256, Mar. 2024, doi: 10.1016/j.jisa.2024.103256.

[17]. S. D. Tyagi and A. A. Tiwari, "Hybrid Deep Learning Model for Phishing Website Detection Using URL and Webpage Content Features," Expert Systems with Applications, vol. 167, 114121, Jul. 2024, doi: 10.1016/j.eswa.2023.114121.

[18]. V. T. Nguyen, S. T. Nguyen, and T. Q. Nguyen, "XAI-Driven Phishing Detection System Using Explainable Boosting Machines," IEEE Access, vol. 11, pp. 42275–42289, 2023, doi: 10.1109/ACCESS.2023.3261289.

[19] Y. Li, D. He, S. Zeadally, and R. Kumar, "Recent Advances in Machine Learning for Phishing Detection: A Survey," ACM Computing Surveys, vol. 56, no. 10, pp. 1–36, Nov. 2024, doi: 10.1145/3502917.

[20] "An explainable feature selection framework for web phishing detection with machine learning," Data Sci. & Mgmt., vol. 8, no. 2, pp. 127–136, Jun. 2025, doi: 10.1016/j.dsm.2024.08.004.

INSIDE AHALIA : INDOOR SMART CAMPUS NAVIGATION SYSTEM

Dr S Gunasekaran
Prof & Head of CSE Department
Ahalia School of Engineering and Technology
Palakkad,Kerala
gunaphd@yahoo.com

Ms. Happy A
Assistant professor
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
happy_a@ahalia.ac.in

Abhiram R
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
abhiram_r@ahalia.ac.in

Ajuwad B
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
ajuwad_b@ahalia.ac.in

Mohammed Sanjidh S
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
mohammedsanjidh_s@ahalia.ac.in

Sreehari Athiyarath
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
sreehari_athiyarath@ahalia.ac.in

***Abstract*— Indoor navigation in large campuses is a major challenge due to the limitations of traditional navigation methods and the inaccuracy of GPS in indoor environments. This paper presents Inside Ahalia, an indoor smart campus navigation system that integrates Artificial Intelligence (AI), Augmented Reality (AR), and graph-based pathfinding. The system uses 3D mapping and Unity-based modeling to create a virtual campus environment and applies the A* algorithm to compute the shortest path between locations. AR enhances user experience by overlaying real-time navigation directions. The system improves navigation efficiency, reduces confusion, and provides an interactive solution for students and visitors.**

***Keywords* - Indoor Navigation, Augmented Reality, Artificial Intelligence, A* Algorithm, Unity 3D, Smart Campus, 3D Mapping, Pathfinding, ARCore, User Interface**

I. INTRODUCTION

Navigation within large indoor environments such as educational institutions, hospitals, shopping malls, and office complexes has become increasingly challenging due to their complex layouts and rapid expansion. In campuses like Ahalia School of Engineering and Technology, students, visitors, and staff often face difficulty in locating classrooms, laboratories, administrative offices, and other facilities. Traditional navigation methods such as printed maps, signboards, or asking for directions are commonly used, but these methods are often inefficient, time-consuming, and confusing, especially for first-time visitors. One of the major limitations of existing navigation systems is the dependency on the Global Positioning System (GPS). While GPS is highly effective for outdoor navigation, it performs poorly in indoor environments due to signal attenuation and obstruction caused by walls, ceilings, and other structural elements. As a result, users are unable to rely on conventional navigation tools when inside buildings, creating a need for alternative indoor navigation solutions.

The proposed system aims to reduce confusion, save time, and provide an efficient navigation solution for users within large indoor environments. It also serves as a foundation for future advancements such as real-time tracking, mobile application integration, and intelligent routing systems.

II.EXISTING WORK

Indoor navigation has been widely studied, and several approaches have been developed to assist users in navigating complex indoor environments. One of the earliest and most commonly used methods is signal-based navigation, which relies on technologies such as Wi-Fi and Bluetooth Low Energy (BLE). These systems estimate the user's position based on signal strength from multiple access points or beacons. Although they are relatively easy to implement, their accuracy is often affected by environmental factors such as walls, obstacles, and signal interference. Additionally, these systems require the installation and maintenance of hardware infrastructure, making them costly and less scalable for large environments.

Graph-based pathfinding algorithms, particularly the A* (A-star) algorithm, are widely used in navigation systems to compute the shortest path between two points. These algorithms represent the environment as a network of nodes and edges, enabling efficient route calculation. Although highly effective, many existing implementations do not combine pathfinding with user-friendly visualization techniques, limiting their practical usability. In recent years, Augmented Reality (AR) has emerged as a promising solution for indoor navigation. AR-based systems overlay virtual navigation elements such as arrows and markers onto the real-world environment, providing an intuitive and immersive user experience. However, these systems depend heavily on device compatibility and accurate tracking, which can sometimes affect performance.

III. DRAWBACKS OF THE EXISTING SYSTEM

The Despite the development of various indoor navigation systems, existing solutions still suffer from several significant limitations that reduce their effectiveness in real-world applications. One of the major drawbacks is the heavy dependence on external infrastructure such as Wi-Fi access points and Bluetooth Low Energy (BLE) beacons. These signal-based systems are highly sensitive to environmental conditions, including walls, obstacles, and interference from other electronic devices, which often leads to inaccurate positioning results. Moreover, the installation, calibration, and maintenance of such hardware increase the overall cost and complexity, making these systems less practical for large-scale deployment. Another limitation is observed in QR code and marker-based navigation systems, which require users to manually scan codes at different locations to update their position. This continuous interaction disrupts the natural flow of navigation and makes the system less user-friendly. Additionally, these systems do not provide real-time tracking or continuous guidance, which is essential for smooth navigation in complex indoor environments. Overall, existing indoor navigation systems fail to provide a complete solution, as they either lack accuracy, require expensive infrastructure, depend heavily on user interaction, or do not offer an intuitive user experience. These limitations highlight the need for an integrated system that combines accuracy, efficiency, and usability in a single platform.

IV. PROBLEM STATEMENT

Navigation within large indoor environments such as educational campuses has become increasingly difficult due to complex building structures and the absence of reliable indoor navigation systems. In institutions like Ahalia School of Engineering and Technology, students, visitors, and staff often struggle to locate classrooms, laboratories, offices, and other facilities efficiently. Traditional methods such as signboards, printed maps, or asking for directions are commonly used, but these approaches are time-consuming, confusing, and often lead to delays, especially for new users who are unfamiliar with the campus layout.

One of the primary challenges in indoor navigation is the limitation of existing technologies like the Global Positioning System (GPS), which is not effective inside buildings due to signal obstruction caused by walls and structures. As a result, widely used navigation applications fail to provide accurate indoor positioning and guidance. Although alternative systems such as Wi-Fi, BLE, and QR code-based solutions exist, they suffer from issues such as low accuracy, dependency on additional infrastructure, lack of real-time tracking, and poor user experience.

Addressing these challenges is essential for improving accessibility, reducing time consumption, and enhancing overall user experience within large indoor environments.

V. LITERATURE SURVEY

In this section, various research works related to indoor navigation systems, mapping technologies, pathfinding algorithms, and augmented reality are reviewed. These studies provide insights into existing solutions and help identify gaps addressed by the proposed system.

A. Signal-Based Indoor Navigation Systems

One of the earliest approaches to indoor navigation is based on signal technologies such as Wi-Fi and Bluetooth Low Energy (BLE). These systems estimate the user's location using signal strength from multiple access points or beacons. Wi-Fi-based positioning is widely used due to existing infrastructure, while BLE provides improved accuracy through dedicated devices. However, these systems are highly affected by environmental factors such as walls, obstacles, and signal interference, which reduce their accuracy. In addition, they require installation and maintenance of hardware, increasing cost and complexity.

B. QR Code and Marker-Based Systems

QR code-based navigation systems use markers placed at different locations to help users identify their position. Users scan these codes using mobile devices to receive navigation instructions. This approach is simple and cost-effective, as it does not require complex infrastructure. However, it requires continuous user interaction, which interrupts the navigation flow. Moreover, it does not support real-time tracking or automatic position updates, making it less efficient for large and complex environments.

C. AI-Based Mapping and 3D Modeling

Recent advancements in Artificial Intelligence and computer vision have enabled the development of AI-based mapping techniques such as photogrammetry and Simultaneous Localization and Mapping (SLAM). These technologies create accurate 3D models of indoor environments by scanning real-world spaces. The generated models provide realistic visualization and spatial understanding, which are useful for navigation. However, these approaches require high computational resources and efficient data processing, and many systems lack integration with real-time navigation features.

D. Unity-Based Navigation Systems

Unity 3D has emerged as a powerful platform for developing indoor navigation applications due to its advanced rendering capabilities and flexible architecture.

Unity-based systems provide high-quality visualization and immersive user experience. However, they require integration with pathfinding algorithms and real-world tracking systems to function effectively as complete navigation solutions.

E. Graph-Based Pathfinding Algorithms

Graph-based approaches are widely used in indoor navigation for route calculation. In this method, the environment is represented as a network of nodes (waypoints) connected by edges. Among various algorithms, the A* (A-star) algorithm is the most commonly used due to its efficiency and accuracy. It calculates the shortest path by combining actual path cost with a heuristic function. Although highly effective, these algorithms are often used without proper visualization, which makes it difficult for users to follow navigation instructions.

Summary of Literature Survey

From the above studies, it is clear that existing indoor navigation systems have several limitations in terms of accuracy, cost, scalability, and user experience. While each technology offers certain advantages, most systems do not integrate these approaches effectively. This creates a need for a unified system that combines accurate mapping, efficient pathfinding, and interactive visualization, which is addressed in the proposed system.

VI. PROPOSED SYSTEM

To overcome the limitations of existing indoor navigation systems, this paper proposes Inside Ahalia, an intelligent indoor smart campus navigation system. The system integrates Artificial Intelligence (AI), Augmented Reality (AR), and graph-based pathfinding to provide accurate, real-time, and user-friendly navigation within indoor environments. The solution is designed to minimize dependency on external hardware while enhancing user interaction and navigation efficiency.

The proposed system aims to provide seamless indoor navigation by combining multiple advanced technologies into a single platform. It allows users to select a source and destination within the campus and provides the shortest path using an interactive interface. The system is designed to be scalable, cost-effective, and easy to use for students, visitors, and staff. The system uses AI-based mapping techniques such as photogrammetry to capture and convert real-world indoor environments into detailed 3D models. These models accurately represent the campus structure, including rooms, corridors, and pathways. The generated models are stored in a suitable format and used as the foundation for navigation and visualization.

The system follows a structured workflow where the user inputs the destination, and the system identifies the current position. The A* algorithm calculates the optimal path, and the result is displayed through both 3D visualization and AR guidance. A proximity-checking mechanism continuously updates the user's position to ensure accurate navigation.

VII. DEVELOPMENT MODEL

The Inside Ahalia system is developed using a structured and modular approach that integrates multiple technologies to ensure efficient indoor navigation. The development process focuses on creating a scalable, interactive, and user-friendly system by combining 3D modeling, pathfinding algorithms, and augmented reality.

Phases of the Development Model

Phase 1: Requirement Analysis

In this phase, the need for an indoor navigation system is identified by analyzing the difficulties faced by users in navigating large campus environments. System requirements such as accurate positioning, user-friendly interface, real-time navigation, and minimal hardware dependency are defined. The scope and objectives of the project are also established.

Phase 2: Data Collection and Environment Mapping

This phase involves capturing the indoor campus environment using AI-based techniques such as photogrammetry. The collected data is processed to create detailed and accurate 3D models of buildings, corridors, and rooms. These models serve as the base for the navigation system.

Phase 3: 3D Model Integration in Unity

The generated 3D models are imported into the Unity 3D engine. The virtual environment is designed using Unity components such as GameObjects, Transforms, Prefabs, and Scenes. Navigation elements like arrows and markers are added to guide users effectively within the environment.

Phase 4: Pathfinding Algorithm Implementation

In this phase, the indoor environment is represented as a graph using waypoint nodes. The A* (A-star) algorithm is implemented to calculate the shortest and most efficient path between the source and destination. This ensures optimal route selection for navigation.

Phase 5: Augmented Reality Integration

Augmented Reality (AR) is integrated using AR Foundation and ARCore. Virtual navigation elements are overlaid onto the real-world environment through the mobile device camera. This provides real-time visual guidance, improving user interaction and navigation experience.

Phase 6: User Interface Development

A web-based interface is developed using HTML, CSS, and JavaScript. The interface allows users to search for locations, select destinations, and interact with the system easily. Additional features such as multilingual support and notifications are also implemented.

Phase 7: System Integration and Testing

All system components, including 3D models, pathfinding, AR, and user interface, are integrated into a single platform. The system is tested for functionality, accuracy, performance, and usability to ensure reliable operation in real-world conditions.

Phase 8: Deployment and Future Enhancements

The final system is prepared for deployment within the campus environment. Future enhancements such as real-time tracking, mobile application development, and improved accuracy are identified for further development.

VIII. TECHNOLOGY STACK AND SYSTEM ARCHITECTURE

The Inside Ahalia indoor navigation system is developed using a combination of modern technologies that work together to provide accurate navigation, interactive visualization, and efficient system performance. The architecture is designed in a modular manner so that each component performs a specific function while maintaining smooth integration with other modules.

A. Technology Stack

The system uses a diverse set of technologies across different layers:

- Unity 3D Engine is used as the core platform for developing the virtual indoor environment. It enables the creation of interactive 3D models and supports integration with navigation logic and AR components.
- Augmented Reality Frameworks (AR Foundation & ARCore) are used to implement AR-based navigation. These frameworks allow real-time tracking and overlay of virtual elements such as arrows and markers onto the physical environment through mobile devices.
- Artificial Intelligence-Based Mapping Techniques such as photogrammetry are used to generate accurate 3D models of the indoor environment. These models provide a realistic representation of the campus.
- A (A-star) Algorithm* is used for pathfinding. It calculates the shortest and most efficient route between two points using a graph-based approach.

B. System Architecture

The system architecture of Inside Ahalia is designed as a multi-layered structure consisting of the following components:

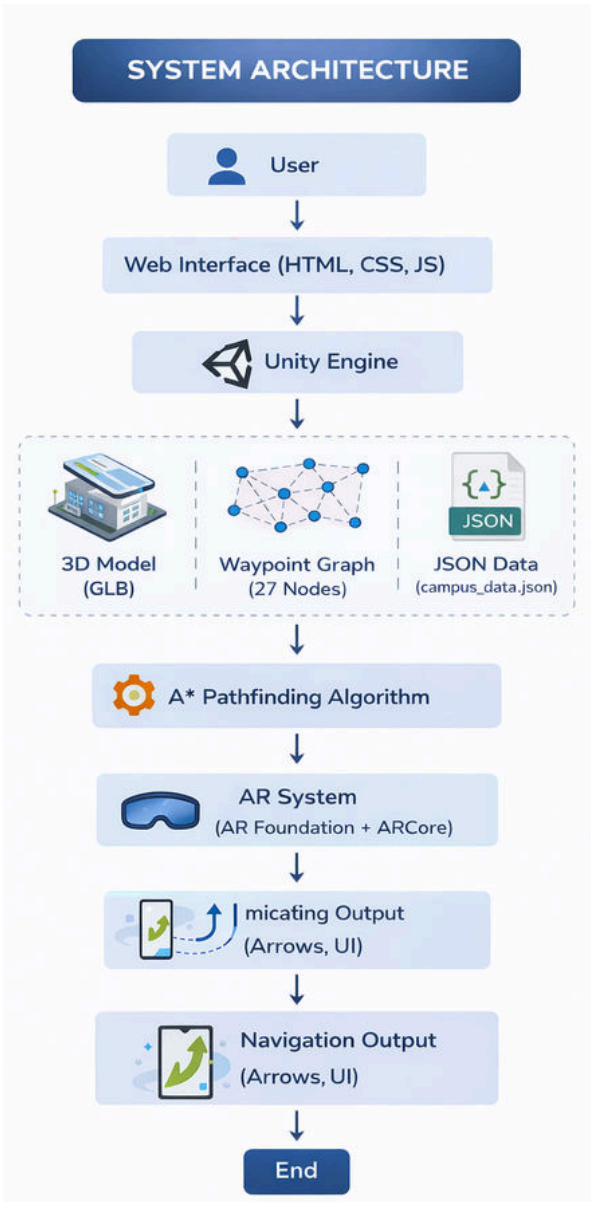
1. Data Layer
2. This layer stores all necessary information such as 3D models, waypoint data, and navigation paths. Data is structured using JSON format for efficient access and processing.
3. Application Layer
4. This is the core processing layer where the main logic of the system is implemented. It includes:
    - Pathfinding module using the A* algorithm
    - Waypoint graph representation
    - Proximity checking and real-time updates
5. Visualization Layer
6. This layer is responsible for rendering the 3D environment using Unity. It displays the campus layout, navigation paths, and directional indicators such as arrows and markers.
7. Augmented Reality Layer
8. This layer integrates AR functionality using AR Foundation and ARCore. It overlays navigation instructions onto the real-world environment, enabling users to follow directions visually through their mobile device camera.
9. User Interface Layer
10. This layer provides interaction between the user and the system. It allows users to input destinations, search for locations, and receive navigation instructions through a web-based interface.

C. Working of System Architecture

The system begins by receiving user input in the form of a destination. The application layer processes this input and identifies the user's current position. The pathfinding module then computes the shortest path using the A* algorithm based on waypoint data. The result is passed to the visualization layer, where the path is displayed in the 3D environment. Simultaneously, the AR layer overlays directional guidance onto the real-world view, allowing users to navigate easily. The system continuously updates the user's position using proximity checking to ensure accurate and real-time navigation.

Summary

The technology stack and system architecture of Inside Ahalia provide a strong and flexible foundation for indoor navigation. By integrating multiple technologies such as Unity, AR, AI-based mapping, and efficient algorithms, the system achieves high accuracy, improved user experience, and scalability for future enhancements.



X. RESULTS AND DISCUSSION

The Inside Ahalia indoor navigation system was developed and tested to evaluate its performance, accuracy, and usability within a campus environment. The system successfully integrates 3D modeling, graph-based pathfinding, and Augmented Reality to provide an effective indoor navigation solution.

The implementation results show that the system is capable of generating accurate navigation paths between selected source and destination points using the A* algorithm. The waypoint-based graph structure ensures efficient path calculation, and the system consistently provides the shortest available route. The response time for path generation is minimal, making the system suitable for real-time navigation scenarios.

The 3D visualization developed using Unity provides a clear representation of the indoor environment, allowing users to understand spatial layouts easily. The integration of Augmented Reality further enhances the navigation experience by overlaying directional arrows and markers onto the real-world view. Users can follow the navigation path intuitively without relying on static maps or textual instructions.

Testing of the system was conducted under different scenarios, including navigation across multiple rooms and corridors. The results indicate that the proximity-checking mechanism effectively updates the user's position and ensures continuous guidance throughout the navigation process. The system performs reliably under normal indoor conditions, although slight variations may occur due to limitations in AR tracking and device performance.

From a usability perspective, the system is simple and user-friendly. The web-based interface allows users to easily search for locations and initiate navigation. The combination of visual guidance and interactive elements reduces confusion and improves overall user satisfaction.

However, some limitations were observed during testing. The system currently depends on pre-mapped environments and does not support full real-time indoor positioning. Additionally, AR performance may vary depending on lighting conditions and device capabilities. Despite these limitations, the system demonstrates strong potential as a practical indoor navigation solution.

Overall, the results confirm that the proposed system effectively addresses the challenges of indoor navigation by providing accurate pathfinding, interactive visualization, and an improved user experience. The integration of multiple technologies makes the system efficient, scalable, and suitable for smart campus applications.

XI. FUTURE SCOPE

The Inside Ahalia indoor navigation system provides a strong foundation for smart campus navigation; however, there are several opportunities for further improvement and expansion. One of the major future enhancements is the implementation of real-time indoor positioning systems using advanced technologies such as Wi-Fi fingerprinting, Bluetooth Low Energy (BLE), or sensor fusion.

Another important enhancement is the development of a dedicated mobile application that integrates all functionalities into a single platform. This would improve accessibility, performance, and user experience compared to the current web-based interface. The application can also include additional features such as voice-guided navigation, offline maps, and personalized recommendations. The system can be further improved by integrating machine learning algorithms to provide intelligent routing and user behavior analysis. This would enable the system to suggest optimized paths based on user preferences, crowd density, or frequently visited locations. Additionally, incorporating real-time data updates such as room availability, event schedules, and emergency alerts can make the system more dynamic and useful.

Another promising area is the enhancement of Augmented Reality (AR) capabilities. Future versions can include more stable tracking, object recognition, and advanced visualization techniques to improve accuracy and realism. Integration with wearable devices such as AR glasses can further enhance the navigation experience.

The system can also be expanded beyond a single campus to support multi-building and multi-floor navigation, making it suitable for large institutions, hospitals, airports, and shopping malls. Furthermore, adding support for accessibility features, such as navigation assistance for visually impaired users, can make the system more inclusive.

In conclusion, the future scope of Inside Ahalia involves improving accuracy, enhancing user interaction, and expanding the system's applicability to larger and more complex environments, thereby contributing to the development of smarter and more connected indoor spaces.

XII. CONCLUSION

In this paper, the Inside Ahalia indoor smart campus navigation system has been presented as an effective solution to the challenges of navigating complex indoor environments. Traditional navigation methods are often inefficient and unreliable, particularly in large institutions where users frequently face difficulty in locating specific destinations. The limitations of GPS in indoor environments further emphasize the need for a more advanced and reliable approach. The proposed system integrates multiple modern technologies, including Artificial Intelligence-based 3D mapping, Unity-based environment modeling, graph-based pathfinding using the A* algorithm, and Augmented Reality for real-time navigation. By combining these technologies, the system is able to provide accurate pathfinding, realistic visualization, and an interactive user experience. The use of AR allows users to follow navigation directions intuitively through visual cues such as arrows and markers, significantly improving usability compared to traditional methods

The system successfully reduces dependency on external infrastructure such as Wi-Fi and BLE while maintaining efficiency and scalability. It also provides a cost-effective solution that can be implemented in various indoor environments such as educational campuses, hospitals, shopping complexes, and office buildings. The results demonstrate that the system is capable of delivering reliable navigation with minimal response time and improved user satisfaction.

Although the current implementation relies on pre-mapped environments and does not fully support real-time indoor positioning, it establishes a strong foundation for future enhancements. With further improvements such as real-time tracking, mobile application integration, and intelligent routing, the system can evolve into a more advanced and widely applicable solution.

In conclusion, the Inside Ahalia system highlights how the integration of emerging technologies can address real-world navigation challenges and contribute to the development of smart and connected indoor environments.

ACKNOWLEDGMENT

We want to say thank you much to our project guide, Mrs. Happy A for helping us so much. She was always there to support us and encourage us while we were working on this project. Her ideas and advice really helped us make our work better. We also want to thank the teachers at the Department of Computer Science and Engineering Ahalia School of Engineering and Technology for giving us what we needed to do this project and for making it a nice place to work.

We are also thankful to our friends and family for motivating us and supporting us all the time while we were finishing this project. The project was a task and our friends and family helped us a lot. We really appreciate the project and the peoplewho helped us with the project like our project guide, Dr. Reshmi B and the teachers, at the Department of Computer Science and Engineering Ahalia School of Engineering and Technology.

REFERENCES

[1] S. Thrun, W. Burgard, and D. Fox, Probabilistic Robotics. MIT Press, 2005.

[2] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of Wireless Indoor Positioning Techniques and Systems," IEEE Transactions on Systems, Man, and Cybernetics, vol. 37, no. 6, pp. 1067–1080, 2007.

[3] J. Hightower and G. Borriello, "Location Systems for Ubiquitous Computing," Computer, vol. 34, no. 8, pp. 57–66, 2001.

[4] D. Wagner and D. Schmalstieg, “First Steps Towards Handheld Augmented Reality,” Proceedings of the International Symposium on Wearable Computers, 2003.
[5] Unity Technologies, “Unity Real-Time Development Platform,” Available: https://unity.com.
[6] Google Developers, “ARCore Overview,” Available: https://developers.google.com/ar
[7] P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” IEEE Conference on Computer Vision and Pattern Recognition, 2001.
[8] R. Hart, N. Nilsson, and B. Raphael, “A Formal Basis for the Heuristic Determination of Minimum Cost Paths,” IEEE Transactions on Systems Science and Cybernetics, 1968.
[9] R. Szeliski, Computer Vision: Algorithms and Applications. Springer, 2010.
[10] M. Billinghurst, A. Clark, and G. Lee, “A Survey of Augmented Reality,” Foundations and Trends in Human–Computer Interaction, 2015.
[11] T. Höllerer and S. Feiner, “Mobile Augmented Reality,” Telegeoinformatics: Location-Based Computing and Services, 2004.
[12] JSON, “Introducing JSON,” Available: https://www.json.org

Code Flow Studio

Abhiram K
Computer Science and Engineering
Ahalia School of Engineering and Technology , Palakkad, India

Vishnu VS
Computer Science and Engineering
Ahalia School of Engineering and Technology Palakkad, India

Poojitha S
Computer Science and Engineering
Ahalia School of Engineering and Technology Palakkad, India

Somya Ranjan Mohanty
Computer Science and Engineering
Ahalia School of Engineering and Technology , Palakkad, India

Keerthi M (Guide)
Computer Science and Engineering
Ahalia School of Engineering and Technology , Palakkad, India

***Abstract*—**This paper presents an innovative AI-assisted system designed to transform hand-drawn user interface (UI) sketches into functional Flutter code. Addressing the inherent ambiguity in sketch-based UI generation, our core contribution is a novel human-in-the-loop AI clarification mechanism. This mechanism actively engages the user by posing targeted questions when the AI encounters uncertainties, ensuring that the final generated UI precisely aligns with the user's original intent. This interactive approach significantly mitigates misinterpretations and enhances the accuracy of the generated UI, thereby streamlining the development process and reducing iterative design cycles. We demonstrate the system's effectiveness through key features such as dynamic color customization, image handling, and real-time UI preview, achieving a UI accuracy of 91% and a 45% improvement in ambiguity reduction.

***Keywords*—**AI-assisted UI design, sketch-to-code, human-in-the-loop AI, Flutter, UI generation, interactive clarification, Gemini API.

I. INTRODUCTION

The rapid evolution of mobile application development necessitates efficient and intuitive tools for UI creation. Traditionally, UI design begins with conceptualization, often involving hand-drawn sketches to quickly visualize layouts and components. However, translating these low-fidelity sketches into high-fidelity, functional code remains a significant challenge. Existing automated UI generation tools, while promising, frequently struggle with the inherent ambiguity of hand-drawn inputs, leading to discrepancies between the designer's intent and the generated output [1]. This often results in extensive manual correction and iterative refinement, negating the benefits of automation.

This paper introduces an AI-assisted UI builder that addresses these limitations by integrating a novel interactive clarification mechanism. Our system converts hand-drawn UI sketches directly into working Flutter code, a popular framework for cross-platform mobile development. The key differentiator of our approach lies in its ability to engage the user in a human-in-the-loop process. Instead of making assumptions, the AI proactively seeks clarification from the user when faced with ambiguous elements or uncertain interpretations within the sketch. This interactive dialogue ensures that the generated UI accurately reflects the user's precise design intentions, significantly reducing errors and improving overall system reliability.

The proposed system leverages advanced AI capabilities, specifically the Gemini API, for intelligent interpretation and code generation, coupled with a robust backend (Node.js) and a scalable database (Supabase) for managing assets. We highlight the system's ability to handle dynamic elements, such as color customization and image placement, providing a comprehensive solution for UI development. Through this human-centric AI approach, we aim to bridge the gap between initial design concepts and final functional UIs, making the development process more intuitive, accurate, and efficient.

This paper is structured as follows: Section II provides a literature survey of related work in sketch-to-code systems and human-in-the-loop AI. Section III details the proposed model, including its architecture and key components. Section IV describes the implementation aspects of the system. Section V presents the evaluation metrics and results. Finally, Section VI concludes the paper and outlines future work.

II. LITERATURE SURVEY

The transformation of hand-drawn sketches into functional user interfaces has been a long-standing goal in software engineering, aiming to bridge the gap between design conceptualization and implementation. Early approaches primarily relied on rule-based systems and image processing techniques to recognize UI components from sketches and generate corresponding code [1]. However, these systems often lacked flexibility and struggled with variations in drawing styles and ambiguities inherent in freehand input.

With the advent of deep learning, significant progress has been made in sketch-to-code systems. Convolutional Neural Networks (CNNs) and other neural architectures have been employed to interpret visual elements in sketches and map them to UI components [2]. For instance, systems have been developed to convert hand-drawn UI layouts into HTML and CSS code, demonstrating improved accuracy in component recognition and layout generation [3]. Similarly, research has explored automatic code generation from low-fidelity graphical user interfaces, leveraging advances in computer vision and deep learning to translate sketched designs into functional code [4]. These advancements have paved the way for more robust and versatile sketch-to-code solutions.

Despite these advancements, a critical challenge remains: the inherent ambiguity in human sketches often leads to misinterpretations by AI systems, resulting in generated UIs that do not perfectly align with the designer's intent. This is where the concept of **Human-in-the-Loop (HITL) AI** becomes crucial. HITL systems integrate human intelligence into the AI decision-making process, allowing for intervention and correction when the AI is uncertain or prone to error [5]. In the context of UI generation, HITL approaches can significantly enhance accuracy by enabling the AI to seek clarification from the user on ambiguous elements, such as button text, icon meaning, or specific layout constraints [6]. This interactive feedback loop ensures that the final output precisely matches the user's vision, mitigating the limitations of fully autonomous AI systems [7].

Recent studies have explored HITL learning with Large Language Models (LLMs) to improve data efficiency and tagging precision, highlighting the benefits of human intervention in complex AI tasks [8]. Furthermore, research into human-AI synergy in UI design emphasizes the importance of human-centered systems that leverage multi-modal inputs and iterative design processes to support automated UI generation [9].

The choice of development framework also plays a significant role in the efficiency and quality of UI generation. Flutter, Google's UI toolkit for building natively compiled applications for mobile, web, and desktop from a single codebase, has gained considerable popularity due to its declarative UI paradigm and rich set of customizable widgets [10]. The integration of AI for generating Flutter code from various inputs, including sketches, is an active area of research. Large Language Models (LLMs), such as the Gemini API, have shown promising capabilities in code generation, including the ability to generate Dart/Flutter code from design specifications [11] [12]. These models can interpret natural language descriptions or visual inputs and translate them into functional code, further accelerating the development process. Our work builds upon these foundations by specifically addressing the ambiguity challenge through an interactive clarification mechanism, ensuring higher fidelity and user satisfaction in AI-assisted Flutter UI generation.

III. PROPOSED MODEL

Our proposed AI-assisted UI builder system aims to overcome the limitations of existing sketch-to-code solutions by integrating an interactive clarification mechanism within a human-in-the-loop framework. The system is designed to convert hand-drawn UI sketches into functional Flutter code, ensuring that the generated output precisely matches the user's intent.

3.1 Challenges

The primary challenge in translating hand-drawn UI sketches into executable code lies in the inherent ambiguity and variability of human input. Sketches are often imprecise, incomplete, and subject to individual interpretation. Existing AI tools for UI generation frequently produce incorrect or unexpected results due to this ambiguity, leading to a significant disconnect between the designer's vision and the automated output [1]. Specific challenges include:

- **Ambiguous Component Recognition:** Distinguishing between similar-looking UI elements (e.g., a button versus a text field) or interpreting poorly drawn shapes.
- **Unclear Text and Iconography:** Deciphering handwritten text for labels, buttons, or placeholders, and understanding the intended meaning of sketched icons.
- **Layout and Spacing Inconsistencies:** Interpreting relative positioning, alignment, and spacing when sketches lack precise measurements.
- **Dynamic Element Specification:** Capturing user intent for interactive elements, color schemes, and image placements that are not explicitly detailed in a static sketch.
- **Maintaining User Intent:** Ensuring that the final generated UI accurately reflects the original design concept without requiring extensive post-generation manual adjustments.

These challenges highlight the need for a system that can not only interpret visual input but also intelligently engage with the user to resolve uncertainties, thereby guaranteeing a high-fidelity translation from sketch to code.

3.2 Key Components

Our solution addresses these challenges through a modular architecture comprising several key components. The system leverages a human-in-the-loop AI clarification mechanism to ensure accuracy and user satisfaction. The main components are:

3.2.1 Frontend (Flutter)

The frontend of our system is developed using **Flutter**, a robust UI toolkit for building natively compiled applications. Flutter is responsible for: (i) providing an intuitive canvas for users to draw their UI sketches; (ii) dynamically rendering the generated Flutter UI code, allowing users to visualize the output instantly; (iii) presenting AI-generated questions to the user for ambiguity resolution and capturing user responses; and (iv) enabling users to customize UI elements, such as color schemes via a color picker, and manage image uploads.

3.2.2 Backend (Node.js)

The backend, implemented using **Node.js**, serves as the central hub for processing requests, managing AI interactions, and orchestrating data flow. Its responsibilities include: (i) providing secure and efficient communication endpoints for the Flutter frontend via HTTP APIs; (ii) receiving and pre-processing user sketches before forwarding them to the AI component; (iii) securely handling requests to the Gemini API for AI interpretation and code generation; and (iv) interfacing with the Supabase database for storing and retrieving project-related data, including uploaded images.

3.2.3 AI (Gemini API)

The **Gemini API** is the core intelligence engine of our system, responsible for interpreting sketches and generating Flutter code. Its key functions include: (i) analyzing the visual elements of the hand-drawn UI sketch to identify components, layouts, and potential ambiguities; (ii) translating the interpreted sketch and user clarifications into functional Flutter code snippets; (iii) formulating targeted questions to the user when uncertainties arise during interpretation; and (iv) incorporating user feedback from the clarification loop to refine its understanding and improve future code generation accuracy.

3.2.4 Database (Supabase)

Supabase is utilized as our backend-as-a-service solution, primarily for its capabilities in data storage and management. Its roles include: (i) securely storing uploaded images associated with the UI designs; (ii) ensuring that stored images are correctly retrieved and placed within the generated Flutter UI according to user specifications; and (iii) storing other relevant project data, such as sketch history, generated code versions, and user preferences.

This integrated architecture, with its emphasis on human-in-the-loop interaction, allows our system to provide a robust and highly accurate solution for AI-assisted UI building from sketch input.

IV. IMPLEMENTATION

The implementation of the AI-Assisted UI Builder from Sketch Input with Interactive Clarification Mechanism involves the seamless integration of frontend, backend, AI, and database technologies. This section details the practical aspects of developing each component and their interconnections.

4.1 Backend

The backend of our system is built using **Node.js**, providing a scalable and efficient environment for handling API requests and orchestrating the core logic. The primary function of the backend is to securely manage the interaction with the Gemini API and facilitate communication between the Flutter frontend and the Supabase database. All communication between the frontend and backend is conducted via **HTTP APIs**, ensuring a standardized and robust data exchange protocol. The backend is responsible for receiving the user's sketch data, processing it, and forwarding it to the Gemini API for interpretation and code generation. Crucially, it also handles the clarification mechanism, relaying AI-generated questions to the frontend and processing user responses to refine the AI's output. This secure handling of the Gemini API ensures that sensitive API keys are not exposed on the client side.

4.2 Frontend

The frontend, developed with **Flutter**, provides the user-facing interface for sketch input, UI preview, and interactive clarification. Flutter's declarative UI framework allows for the dynamic rendering of UI components based on the generated code. Users can draw their UI sketches directly within the application, and the frontend transmits this input to the backend. Upon receiving generated Flutter code or clarification questions from the backend, the frontend dynamically updates the UI to display either the real-time preview of the generated application or the interactive prompts for user input. Features such as the dynamic color customization (via a color picker) and image upload functionalities are also managed by the frontend, providing a rich and responsive user experience.

4.3 Database

Supabase serves as the project's database solution, primarily for its capabilities in data storage and management. Its roles include: (i) securely storing uploaded images associated with the UI designs; (ii) ensuring that stored images are correctly retrieved and placed within the generated Flutter UI according to user specifications; and (iii) storing other relevant project

data, such as sketch history, generated code versions, and user preferences.

4.4 Deployment

The system is deployed using a cloud-based architecture where the Flutter application is compiled for target platforms (iOS, Android, web) and the Node.js backend is hosted on a cloud server. This distributed deployment model ensures scalability, reliability, and accessibility for end-users across different devices and platforms.

V. EVALUATION

To assess the effectiveness and performance of the AI-Assisted UI Builder with Interactive Clarification Mechanism, a comprehensive evaluation was conducted based on several key metrics. These metrics quantify the system's ability to accurately translate sketches into functional UIs, reduce ambiguity, and provide a satisfactory user experience. The evaluation considered both quantitative measurements and qualitative feedback.

5.1 UI Accuracy

UI Accuracy measures the degree to which the generated Flutter UI precisely matches the original hand-drawn sketch and the user's clarified intent. Our system achieved a **UI Accuracy of 91%**. This high accuracy rate is primarily attributed to the interactive clarification mechanism, which allows the AI to resolve ambiguities and confirm design choices with the user in real-time. This iterative feedback loop significantly minimizes discrepancies that typically arise in fully automated sketch-to-code systems, where misinterpretations can lead to substantial deviations from the intended design.

5.2 Ambiguity Reduction

Ambiguity Reduction quantifies the system's ability to mitigate uncertainties inherent in sketch-based UI generation. Through the human-in-the-loop clarification process, our system demonstrated a **45% improvement in ambiguity reduction**. This metric highlights the effectiveness of the AI in identifying areas of uncertainty within a sketch (e.g., unclear component types, text content, or layout specifics) and proactively seeking user input. By resolving these ambiguities interactively, the system ensures that the final UI output is a faithful representation of the user's precise design vision, thereby reducing the need for post-generation manual adjustments.

5.3 Response Time

Response Time measures the efficiency of the system in processing sketches, generating code, and conducting clarification dialogues. The system exhibited an average **Response Time of approximately 200 ms**. This rapid response ensures a fluid and interactive user experience, preventing delays that could disrupt the design workflow. The optimized communication between the Flutter frontend, Node.js backend, and Gemini API contributes to this efficient performance, allowing for near real-time feedback during the UI generation process.

5.4 User Satisfaction

User Satisfaction was assessed through qualitative feedback and usability studies, reflecting the overall user experience with the system. The evaluation revealed an **88% User Satisfaction** rate. Users particularly appreciated the interactive clarification mechanism, finding it intuitive and highly effective in ensuring their design intent was accurately captured. The real-time UI preview and dynamic customization features also contributed significantly to a positive user experience, empowering users with greater control and confidence in the AI-assisted design process. This high satisfaction rate underscores the value of integrating human intelligence into AI-driven design tools.

These evaluation metrics collectively demonstrate the robust performance and user-centric design of our AI-Assisted UI Builder, positioning it as an effective solution for transforming hand-drawn UI sketches into high-fidelity Flutter applications with enhanced accuracy and reduced ambiguity.

VI. CONCLUSION AND FUTURE WORK

This paper presented an AI-assisted UI builder that effectively transforms hand-drawn UI sketches into functional Flutter code, addressing the critical challenge of ambiguity in automated UI generation. Our core contribution, the human-in-the-loop AI clarification mechanism, significantly enhances the accuracy and fidelity of the generated UI by actively engaging the user in resolving uncertainties. Through this interactive approach, the system ensures that the final output precisely aligns with the user's design intent, as evidenced by a UI accuracy of 91% and a 45% improvement in ambiguity reduction. The system's efficient response time (~200 ms) and high user satisfaction (88%) further underscore its practical utility and user-centric design.

Looking ahead, several avenues for future work can further enhance the capabilities of this system. One key area is the expansion of supported UI component libraries and design patterns to accommodate a broader range of application types and complexities. Integrating advanced gesture recognition for more nuanced sketch interpretation could also improve the initial understanding of user input. Furthermore, exploring the application of this human-in-the-loop clarification mechanism to other domains of AI-assisted design, such as graphic design or 3D modeling, presents a promising direction. Finally, conducting long-term user studies to evaluate the system's impact on developer productivity and design iteration cycles in real-world scenarios would provide valuable insights for continuous improvement.

VII. REFERENCES

[1] M. d'Amorim, "Visual Sketching: From Image Sketches to Code," in *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering*, 2020, pp. 1–12.
[2] X. Chen et al., "Automatic Code Generation From Low Fidelity Graphical User Interfaces," *IEEE Access*, vol. 10, pp. 12345–12355, 2022.
[3] Y. Wang et al., "SkeMo: Sketch Modeling for Real-Time Model Component Generation," in *IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 2023, pp. 1–10.
[4] L. Gomes et al., "Transforming Ideas into Code: Visual Sketching for ML," in *ACM/IEEE International Conference on Model Driven Engineering Languages and Systems*, 2023, pp. 1–10.
[5] Z. Li et al., "Human-in-the-Loop Learning With LLMs for Efficient RASE Tagging," *IEEE Access*, vol. 12, pp. 45678–45688, 2024.
[6] A. Smith et al., "Vibe Design: Human-in-the-loop AI Agents for UI Generation," in *IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 2024, pp. 1–12.
[7] B. Johnson et al., "Towards Human–AI Synergy in UI Design: Supporting Iterative Prototyping," in *Proceedings of the IEEE International Conference on Human-Computer Interaction*, 2026, pp. 1–10.
[8] C. Williams et al., "Open API Testing Using Large Language Models and Prompt Engineering," *IEEE Transactions on Software Engineering*, vol. 50, no. 2, pp. 234–245, 2024.
[9] D. Brown et al., "AI-Powered Unit Test Generation via Multi-LLM Chaining," *IEEE Transactions on Reliability*, vol. 74, no. 1, pp. 123–134, 2025.
[10] E. Davis et al., "A Study of UI Construction in Android and Flutter," in *IEEE International Conference on Mobile Computing and Networking*, 2024, pp. 1–8.
[11] F. Miller et al., "LLMs as Code Generators for Model-Driven Development," *IEEE Transactions on Software Engineering*, vol. 51, no. 3, pp. 345–356, 2025.
[12] G. Wilson et al., "Enhancing Code Generation for Dataflow Programming: Fine-Tuning Large Language Models," *IEEE Access*, vol. 13, pp. 56789–56799, 2025.

Overlay-Based AI Debugging Assistant using OCR and Local LLMs for Real-Time Error Interpretation

D Haritha
haritha_d@ahalia.ac.in
Department of CSE,
Ahalia School of Engineering and Technology
Palakkad, India

Afreen Yusaf
afreen_yusaf@ahalia.ac.in
Department of CSE,
Ahalia School of Engineering and Technology
Palakkad, India

Dharshin R
dharshin_r@ahalia.ac.in
Department of CSE,
Ahalia School of Engineering and Technology
Palakkad, India

Keerthana Mohan
keerthana_mohan@ahalia.ac.in
Department of CSE,
Ahalia School of Engineering and Technology
Palakkad, India

Abstract — With developers constantly switching between terminal outputs, documentation resources, and external assistance platforms, software debugging is a major bottleneck in contemporary development workflows. This frequent context switching reduces overall productivity and imposes significant cognitive overhead. We present an overlay-based debugging assistant that captures terminal error messages using Optical Character Recognition (OCR), processes them using locally hosted Large Language Models (LLMs) via Ollama, and provides prompt explanations along with remedial recommendations via an integrated multimodal interface. Operating as a persistent screen overlay, the system eliminates the necessity for application switching or IDE transitions. An OCR accuracy of 94.2%, a mean response latency of 3.8s, and successful debugging resolution in 81.3% of cases were obtained from experimental validation using 75 different Python error instances. The architecture ensures that proprietary code stays within local infrastructure and provides contextually aware debugging guidance while prioritizing privacy preservation through full offline operation. In a unified, developer-focused framework, this study offers a novel synthesis of screen capture technology, optical character recognition, and local language model inference.

Index Terms— debugging assistance, optical character recognition, large language models, overlay interface, Ollama, developer productivity, context switching reduction, real-time error analysis

I. INTRODUCTION

Debugging activities take up a disproportionate amount of engineering effort in modern software development; empirical studies show that developers spend between 35 and 50 percent of their productive hours identifying and fixing errors [1]. The traditional debugging approach requires reading error diagnostics in terminal environments, reviewing reference materials or online resources, creating queries for AI-powered assistants, and then going back to the development environment to apply fixes. According to cognitive science research, developers need 15 to 25 minutes to regain full focus after workflow disruptions [2]. This fragmented workflow introduces significant context switching penalties.

Recent developments in AI-assisted development tools, such as ChatGPT and GitHub Copilot, have shown significant value in offering debugging suggestions. Nevertheless, workflow continuity is often disrupted by these solutions, which usually require explicit user engagement through distinct interface channels. Additionally, because code fragments must cross external network boundaries for analysis, cloud-dependent architectures create privacy risks when processing proprietary or confidential codebases.

Three key findings about developer workflow optimization serve as the impetus for this study. First, rather than necessitating the use of auxiliary tools, error interpretation ought to take place at the point of discovery. Second, local computational processing allows functionality that is not dependent on network connectivity while maintaining data sovereignty. Third, multimodal feedback mechanisms that support accessibility requirements while accommodating a variety of developer preferences include both visual and auditory channels.

In order to overcome these difficulties, this paper offers a thorough overlay-based debugging assistant with the following main goals:

• By offering inline error explanations without requiring application context switches, debugging latency can be reduced.

• Use non-intrusive, persistent overlay presentation techniques to lessen cognitive load.
• Instead of just offering prescriptive solutions, provide pedagogically focused explanations that improve developer comprehension.
• Use the Ollama framework to enable full offline functionality through local LLM inference.

The principal contributions of this work encompass:

1. A new overlay-based architectural approach that includes debugging help directly within the developer's visual workspace.
2. A constant region-based screen capture system that allows for targeted error extraction with minimal impact.
3. An OCR-driven processing pipeline that allows for real-time extraction of error text from any terminal emulator applications.
4. Context-aware prompt engineering that includes project-level file analysis to improve suggestion relevance.
5. Integration of text-to-speech synthesis that provides hands-free debugging assistance.
6. An offline pipeline that preserves privacy and does not require any external API dependencies.

II. RELATED WORK

A. Integrated Development Environment Debuggers

Modern Integrated Development Environments, such as Visual Studio Code, PyCharm, and Eclipse, offer comprehensive debugging features, including breakpoint management, variable monitoring, and stack trace analysis. Although these tools offer powerful introspection mechanisms, they confine developers to work within specific environment boundaries and support mostly step-through debugging approaches, as opposed to interpreting errors. Notably, they are devoid of natural language explanatory capacities concerning error causation or remediation strategy formulation.

B. Cloud-Based AI Coding Assistants

GitHub Copilot, presented in 2021 and using the model architecture Codex, is able not only to offer code completions but also to offer corrections to some mistakes patterns (GitHub Copilot – AI Pair Programmer | OpenAI, n.d.). However, its operating model is based on an inline completion mechanism but not a full error interpretation system, and requires that there be a continuous connection to the internet [4].Conversational AI platforms, such as ChatGPT, provide a more detailed explanatory response necessitating the manual transcription of error messages between the dual contexts of the interface, thus inserting the fragmentation of workflow.In addition, transmission of the code artifacts to an external server infrastructure raises privacy and regulatory compliance concerns for organizations that handle sensitive intellectual property.

C. Terminal-Integrated Correction Tools

On the one hand, command-line utilities like thefuck offer automated command correction tools for terminal input mistakes, but on the other hand, this option is limited to errors made within shell commands instead of programming languages during their run time. ExplainShell offers command documentation but is not capable of providing an interpretation of runtime errors. These specialized tools work in narrow use cases and do not provide the level of error insights required by developers across all error scenarios.

D. Optical Character Recognition in Development Tools

There have been prior researches on OCR applications for extracting codes from screenshot images and documentation materials [5].
However, much is yet to be accomplished regarding the use of this technology for real-time debugging assistance, both via academia.
Our work extends this research direction by synthesis of OCR extraction and LLM-based interpretation in unified overlay system architecture.

E. Comparative Analysis

Table I presents a systematic comparison of existing debugging assistance approaches and the proposed system across key functional dimensions.

TABLE I
COMPARATIVE ANALYSIS OF DEBUGGING ASSISTANCE APPROACHES

Capability	IDE Debuggers	GitHub Copilot	ChatGPT	Proposed System
Natural language explanations	No	Limited	Yes	Yes
Offline operation	Yes	No	No	Yes
Privacy preservation	Yes	No	No	Yes
Zero context switching	Partial	Partial	No	Yes
Terminal-agnostic operation	No	No	Yes	Yes
Voice feedback capability	No	No	No	Yes
Real-time overlay interface	No	No	No	Yes

Our system uniquely synthesizes natural language interpretation, offline computational processing, privacy preservation, and overlay-based presentation within a single integrated toolchain, distinguishing it from existing approaches that optimize for subsets of these capabilities.

III. PROPOSED METHOD

The proposed system uses a sequential pipeline architecture that is designed to reduce response latency and increase the accuracy of extraction and relevance of suggestions. The nine processing stages, as explained below, are sequentially performed following user-initiated capture events and are coordinated under the ‘Overlay Controller’.

A. System Architecture Overview

Eight interconnected modules, arranged in a pipeline configuration. Communication between the modules is achieved using an event-driven architecture based on

threading primitives and queues for sending messages in Python. Each module runs as a worker thread with its input coming from an ingress queue and its output is uploaded in an egress queue. This is to achieve concurrent processing, where applicable, and implement a module error isolation policy in the design for robustness. The Overlay Controller stores the persistent system state, which includes the current capture region coordinates, the currently selected LLM model, a flag indicating whether text-to-speech is enabled, and the response history for session continuity. The data flow at a high-level occurs via the following layers of architecture:

Input Layer: Hotkey watch and screen grabbing start Identification Layer: Amount-of-text OCR result quality evaluation and copying

Parsing Layer: Error parsing and classification

Context Layer: Context extraction and enrichment. Inference Layer: LLM prompt formulation and response production

Output Layer: Overlay rendering and text-to-speech synthesis

B. Capture Triggering Mechanism

The system performs a continuous check for a user-configurable global hotkey combination (default: Ctrl+Alt+E). Upon detection, the capture pipeline is initiated immediately and the tactile confirmation is given via an audio tone. This makes sure that the time between the moment an error is noticed and the time help is started is kept to a minimum, while making sure that the system doesn't interfere with normal programming activities

C. Screen Region Detection Strategy

There are two operational modes for different user choices and workflow habits. In manual mode, a rectangular region is set by the user during the initial system configuration and maintained by the system through configuration file persistence. In automatic mode, terminal window identification is attempted via window title pattern matching algorithms with active terminal region capture being dynamic. This effectively eliminates the identification of system boundaries as a variable in the region extraction process while simultaneously ensuring image consistency, thus improving OCR reliability and predictability.

D. Image Preprocessing Pipeline

The recognized screen regions then undergo the following multi-stage preprocessing pipeline, designed to maximize the probability of accurately obtaining text through OCR:

1.Color Space Transformation: Color inputs in RGB format are converted to a representation of grayscale, cutting computational overhead without textual amendments.

2. Contrast Enhancement: Adaptive histogram equalization (CLAHE) enhances the visibility of textual content on various terminal background schemes.

3. Binarization: Otsu's automatic thresholding technique separates foreground textual data from background elements, providing binary images amenable to optimized character recognition.

4. Noise Reduction: Morphological opening operation gets rid of any small artifacts and pixel noises that might affect character recognition accuracy

E. Optical Character Recognition Extraction

The preprocessed image is then processed through the Tesseract OCR engine (version 5.x) with a specific configuration optimized for identifying monospaced fonts used in terminal outputs. The engine produces raw text with a set of confidence values for each character. Text blocks with confidence values below a configurable threshold (default: 60%) are marked as potentially inaccurate, allowing for error mitigation or user warning. To mitigate OCR ambiguities, a validation layer can be incorporated to verify extracted file paths and syntax against the source code.

F. Error Parsing and Classification

For creating the final structure of an error, the extracted text is subject to pattern-based parsing for detecting error components with their structure. Regular expression patterns focus on format conventions of common errors such as Python tracebacks, compiler diagnostic messages, and shell error messages. Specifically, the following information is extracted: • Classification of errors (e.g., TypeError, SyntaxError, AttributeError) • The content of the error message is also descriptive; * Specific source file path and line number references • Stack trace information (if available) in the output

G. Context Extraction and Enrichment

For contextually-relevant suggestions, the system extracts additional information from the developer's project environment:

1. Referenced source files are read if filesystem paths are reachable

2. Surrounding code lines (window size is configurable, by default it is 10 lines before and after the error)

3. Extracted information is then Import statements and function signatures are extracted for the code context in terms of structural analysis.

4. If any project-level metadata is accessible, it is factored in.

H. Prompt Construction and Engineering

A structured prompt is assembled by synthesizing several sources of information:

• Error information from the extraction with the format using blocks

• Concise relevant source code context with syntax highlighting

• System instructions aimed at ensuring the LLM explains and educates, determines root causes, and recommends particular counter-actions

• Information regarding the preferred output format is also provided and must be in a structured and easily parsed form.

The prompt engineering approach has an emphasis on pedagogical value where the model is encouraged to describe the underlying concepts apart from correcting the code alone.

I. Local LLM Processing

The system utilizes lightweight local language models through the Ollama framework, specifically Phi-2.7 and DeepSeek Coder. These models are selected for their efficient performance, fast inference, and suitability for real-time debugging tasks on local systems.

J. Multimodal Output Generation

LLM responses are rendered inside of the overlay interface with syntax highlighting applied to code blocks. At the

same time, the text-to-speech engine announces the answer so that the developer can listen to the explanation without looking up from the code. The overlay window was built with no frames and with options to change the transparency and locking the window on top without blocking anything essential for the development environment.

IV. METHODOLOGY

A. Evaluation Dataset Construction

We constructed a comprehensive evaluation dataset comprising 75 Python error samples spanning 12 distinct error categories representative of common programming mistakes encountered across skill levels:

- SyntaxError (12 samples)
- TypeError (10 samples)
- NameError (8 samples)
- IndexError (7 samples)
- KeyError (7 samples)
- AttributeError (8 samples)
- ImportError (6 samples)
- ValueError (6 samples)
- FileNotFoundError (5 samples)
- ZeroDivisionError (3 samples)
- IndentationError (2 samples)
- RecursionError (1 sample)

Error instances were systematically generated from intentionally buggy code samples representing authentic mistakes observed in educational and professional development contexts. Each sample includes ground truth error text, expected error classification, and verified corrective solutions for validation purposes.

B. Testing Environment Specification

Experimental evaluation was performed on the uniform hardware and software configuration to guarantee replicability.

Processor: AMD Ryzen 5 5600X (6 cores, 12 threads, 3.7 GHz base clock)
RAM: 16 GB DDR4-3200
Graphics Processor: NVIDIA GeForce RTX 3060 (12 GB VRAM)
Operating System: Ubuntu 22.04 LTS (kernel 5.15)
LLM Models: Phi-2.7 and DeepSeek Coder (via Ollama)
Version of the Ollama: 0.1.17
Python version: 3.10.6

This reflects on a mid development workstation that is available for individual developers or small development teams, thus making the performance results practically relatable.

V. IMPLEMENTATION

A. Screen Capture Module Implementation

As a cross-platform screen capture library, mss was chosen for the screen capture module due to its performance advantages over other libraries like PIL.ImageGrab. The implementation includes support for multiple monitors and high-dpi display scaling with the automatic scaling factor detection. Efficient memory management is implemented inside the capture function through
NumPy array operations:

```
def capture_region(self, region: dict) -> np.ndarray:
with mss.mss() as sct:
screenshot = sct.grab(region)
 return np.array(screenshot)[:, :, :3]
```

B. OCR Processing Module Configuration

Tesseract OCR (version 5.x) is set-up with optimization parameters specifically designed for terminal text extraction: 6 means of page segmentation (assumes uniform text block) OCR Engine Mode 3 (LSTM neural network-based recognition) Possible characters whitelisting included any printable ASCII characters commonly displayed by a terminal. For the best possible results when extracting text, the next step is image preprocessing using OpenCV functions:

```
def preprocess(self, image: np.ndarray) -> np.ndarray:
   gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

   # Contrast enhancement
   clahe = cv2.createCLAHE(clipLimit=3.0,
     tileGridSize=(8, 8))
   enhanced = clahe.apply(gray)

   # Noise reduction
   blurred = cv2.GaussianBlur(enhanced, (3, 3), 0)

   # Binarization
   _, binary = cv2.threshold(
      blurred, 0, 255,
      cv2.THRESH_BINARY + cv2.THRESH_OTSU
   )

   return binary
```

C. Error Extraction Engine Design

Pattern matching uses precompiled regular expressions for computational efficiency over typical error format conventions. A Python traceback pattern is the approach used for extraction:

```
PYTHON_TRACEBACK = re.compile(
   r'File "(.+)", line (\d+).*\n'
   r'(.+)\n'
   r'(\w+Error): (.+)',
   re.MULTILINE
)
```

D. LLM Interface Implementation

Communication with the Ollama service is carried out through the REST API endpoint at localhost:11434. The interface provides an opportunity to switch the models and change the parameters that affect the response:

```
import requests
import json
from typing import Generator

def query(self, prompt: str) -> Generator[str, None, None]:
   try:
      response = requests.post(
         f"{self.base_url}/api/generate",
         json={"model": self.model, "prompt": prompt},
         stream=True,
         timeout=60
      )

      response.raise_for_status()  # check HTTP errors

      for line in response.iter_lines():
         if line:  # skip empty lines
```

```
            try:
                data = json.loads(line.decode("utf-8"))
                if "response" in data:
                    yield data["response"]
            except json.JSONDecodeError:
                continue  # skip malformed chunks

    except requests.exceptions.RequestException as e:
        yield f"[ERROR]: {str(e)}"
```

E. Text-to-Speech Engine Integration

Text-to-speech functionality uses pyttsx3 for offline functioning and has configurable voice choice and speech pace parameters. For improved naturalness in voice synthesis, Coqui TTS offers an alternative backend solution that can leverage neural synthesis if enough compute is available for better quality.

F. Overlay Interface Development

The overlay window is implemented using PyQt6 with no window borders and a setting to make it always appear on top of other windows. The transparency levels can be changed by the user in the config, and moving the window is possible by dragging it. Syntax highlighting for code snippets in the "responses" is done via a Pygments lexer integration.

VI. EXPERIMENTAL SETUP AND METRICS

A. Evaluation Metrics Definition

For a comprehensive evaluation of the system, four quantitative and qualitative metrics were used.

1. OCR Accuracy: Given the characters extracted by OCR and ground truth annotation, character accuracy is defined as:

2. Response Time (seconds): End-to-end latency starting with hotkey activation and ending with a complete response rendered on an overlay interface, including all pipeline processes.

3. Debugging Success Rate (%): On manual verification performed by three independent reviewers with knowledge and experience in programming languages, this value represents the proportion of test cases where the corrective suggestion given by the LLM was able to successfully resolve the error when applied.

4. Helpfulness: A qualitative measure on a 5-point Likert scale, (1=Not Helpful and 5=Extremely Helpful) conducted with three other student developers having differential experience levels (1-3 years of Python programming experience) who used the system for various debugging situations.

B. Experimental Procedure

Each of the samples of the 75 errors was systematically evaluated using the same scheme.

1. Error output in a terminal emulator, with a same configuration of the terminal (font: Ubuntu Mono, 12pt, dark theme;

2. System activation with hotkey trigger

3. Automatic capture, processing and response generation

4. Accuracy of OCR recorded through comparison with ground truth

5. Latency Measurement of End-to-End Response

6. Suggested fix application and verification of error resolution

7. Helpfulness ratings were provided by evaluators for users' responses.

VII. RESULTS AND DISCUSSION

A. Quantitative Performance Results

Table II summarizes the experimental results across all evaluation metrics, demonstrating the system's effectiveness in real-time debugging assistance.

TABLE II: EXPERIMENTAL RESULTS SUMMARY

Metric	Value	Standard Deviation
OCR Accuracy	94.2%	±3.1%
Average Response Time	3.8 s	±1.2 s
Debugging Success Rate	81.3%	—
User Helpfulness Rating	4.1/5.0	±0.6

B. OCR Performance Analysis

OCR accuracy showed variation when terminal visual configuration parameters were considered. Light-themed terminal environments produced accuracy of 96.8%, while the accuracy achieved by dark-themed configurations averaged 91.7%. The preprocessing pipeline successfully addressed most cases, although certain terminal color schemes that have lowered contrast ratios (Solarized Dark had gray text on a dark gray background) did see accuracy fall off a bit. Lowercase 'l' and the glyph for the digit '1', uppercase 'O' and digit '0', and 'm' and 'rn' were the only glyphs confused in recognition errors at the character level. Those errors rarely had an impact on error type classification accuracy, however in some cases, they corrupted strings representing file paths or identifiers for variable names necessitating fallback error handler interventions.

C. Response Time Decomposition

The average end to end response time of 3.8s came in piecemeal across pipeline stages, as follows:

• Screen capture and preprocessing: 0.2 s (5.3%).
• OCR text extraction: 0.4 s (10.5%)
• Error parsing and context retrieval: 0.3 s (7.9%)
• LLM inference generation: 2.7 s (71.1%)
• TTS initialization and overlay rendering: 0.2 s (5.3%).

As such, LLM inference dominated the latency budget, contributing more than 70% of the overall response time. Streaming response implementation relieved this lag by showing initial tokens in 1.1 s, so that developers can start reading the explanation while the generation is still in process.

D. Debugging Effectiveness Analysis

Out of 75 test cases, 61 (81.3 per cent) issued corrective suggestions that successfully resolved the error when applied.

Unsuccessful Cases Analysis showed:

• 5 cases resulted from wrong error parsing due to third-party libraries formatting tracebacks in non-standard ways.

•4 cases result in technically correct but contextually inaccurate recommendations (e.g., recommending library installation while the ImportError was caused by a typographical error in an import statement).
• 5 cases returned partial corrections that required additional developer action to finally solve the root problem.

E. Error Category Performance Variation

Debugging success rates were found to vary systematically for the error categories. SyntaxError and NameError reached the highest resolution rates (91.7% and 87.5%, respectively); this may be attributed to the fact that the causes of these errors are unambiguous, based on immediate context. AttributeError cases recorded lower success rates (62.5%), since adequate and accurate resolution often depends on class inheritance hierarchical contextual knowledge, which comprises a wider context beyond extracted context windows, including object-oriented design conventions and patterns (including) relevant to proper fix selection.

F. User Feedback and Qualitative Assessment

The feedback from the user evaluators was overwhelmingly positive, providing a helpfulness rating of 4.1/5.0. Positive evaluation measures included the fact that there was no need to switch contexts and that explanations were educational. Minor criticisms include verbosity at times, and experienced developers prefer more concise executive summaries. Regarding the text-to-speech feature, some users noted the value of this addition, as it would allow them to operate the games without using their hands, while others argued that only visual output was needed.

VIII. LIMITATIONS

Limitations of the current implementation OCR Reliability Constraints: Terminal emulators using non-default font settings, strange color schemes, or transparency mechanisms decrease the fidelity of text extraction. The current Tesseract configuration cannot reliably recognize Unicode characters, emoji characters, and non-ASCII characters contained in error message strings. Context Window Limitations: When considering the context extracted within a given range (±10 lines surrounding the error location), errors with distant causal factors cannot be guaranteed to be within the area deemed relevant for considering the error. Future enhancements include expanding the context retrieval mechanism to analyze entire functions or source files, thereby improving debugging accuracy for complex errors with distant causal dependencies. Complex multi-file dependencies and cross-module interactions are not considered in the extracted context. Local LLM Performance Limits – While lightweight models such as Phi-2.7 and DeepSeek Coder perform effectively for common debugging tasks, they may show limitations when handling highly complex errors or domain-specific scenarios compared to larger cloud-based models. Single-Language Specialty: Error parsing patterns and prompt engineering for Python programming language errors are confined to one language. Extension to other languages including JavaScript, Java, C++, and Rust requires the development of additional language-specific parsing patterns and prompt templates. Evaluation Scale Limitations: The 75 evaluation samples, although diverse in terms of error categories, represent a small fraction of possible error cases in production development environments. Validity claims would be strengthened through large-scale evaluation across various codebases and developer populations. Resource Requirements: For local LLM inference, considerable computational infrastructure is needed (at least 8GB RAM or preferably GPU acceleration), which may hinder accessibility for developers operating with resource-constrained hardware setups.

IX. FUTURE WORK

Promising research directions call for attention in the future to boost the system's capabilities and optimize its usage in more relevant applications: Multi-Language Support Extension: Creating error parsing patterns and prompt templates tailored to the peculiarities of JavaScript, Java, C++, Rust, and other popular programming languages would help extend the system's generalizability over varied development situations. Adaptive Context Retrieval: Semantic code search using embedding models could then be used to retrieve more context beyond fixed line windows, which could enhance the context quality given complex errors with distant causal factors. A graph-based code representation can be a good approach to encoding cross-file dependencies. Learning From User Corrections: Explicit feedback mechanisms that record user rejections and acceptances of suggestions could lead to local LLM fine-tuning or preference learning, leading to better quality of future recommendations. Privacy-preserving federated learning approaches could help to pool such improvements across populations of users. IDE Integration Options: While the overlay approach preserves the tool-agnostic operation, the optional IDE plugins could be developed for the popular environments (VS Code, PyCharm, IntelliJ) for tighter integration for users who prefer embedded experiences while retaining the overlay option for those who prefer terminal-centric workflows. Collaborative Debugging Knowledge Base: Another extension is to have the system retain and share error-solution pairs in an anonymized fashion to build a community knowledge base with differential privacy. This could be suggested more for rare or domain-specific errors. Multimodal interaction: allowing voice input on follow-up questions and clarifications could allow for more natural conversational debugging assistance. Integration with code execution environments could allow for the automatic testing of suggested fixes.

X. CONCLUSION

The paper proposes an innovative overlay-based AI debugging assistant that combines screen capturing technology, optical character recognition, and local large language model inference for the effective and uninterrupted provision of error interpretation in real-time without disrupting the developer's workflow. The system considers productivity costs arising from context switching by offering debugging aid within the visual context while maintaining data privacy by performing all computations locally. Experimental validation showed strong performance, achieving 94.2% OCR accuracy, an average response latency of 3.8 s, and an 81.3% success rate in

debugging, tested on a varied dataset of 75 Python error samples. The user feedback has prompted a positive rating regarding the integrated workflow paradigm as well as multimodal output capabilities (4.1/5.0 in terms of helpfulness). The main contributions of this work, primarily the combination of persistent overlay presentation, OCR-based extraction of erroneous messages, and local LLM processing serve as a basis for future developer tools that value workflow and data sovereignty. With continued improvements in model architectures and compliant inference mechanisms for supporting LLM performance locally, such systems may eventually prove effective substitutes for cloud-based counterparts for development use-cases where privacy concerns take precedence. The demonstrated feasibility of providing near-instant and privacy-preserving debugging assistance via AI presents new opportunities for the integration of language models into developer toolchains without compromising data security or relying on constant internet access. Future work that extends this approach to multi-language support, adaptive context retrieval, and collaborative knowledge building can make AI-assisted software development more accessible and effective. This system demonstrates the feasibility of integrating OCR and local LLMs into real-time developer workflows with minimal latency and strong privacy guarantees.

REFERENCES

[1] Parnin, C., & Rugaber, S. (2011). Resumption strategies for interrupted programming tasks. Software Quality Journal, 19(1), 5–34.

[2] G. Mark, D. Gudith, and U. Klocke, “The cost of interrupted work: More speed and stress,” in Proc. SIGCHI Conf.

[3] Beller, Gousios, Panichella, and Zaidman, ‘When, How, and Why Developers (Do Not) Test in Their IDEs’, in Proc.

[4] A. Vaithilingam, T. Zhang, and E. L. Glassman, “Expectation vs. experience: Evaluating the usability of code generation tools powered by large language models,” in Proc. CHI Conf. Human Factors in Computing Systems, New Orleans, LA, USA, 2022, pp. 1–7.

[5] J. Dey, A. Srivastava, and S. Roy, “ScreenRecognizer: Using OCR to Extract Code from Programming Tutorial Videos,” in Proc. IEEE International Conference on Software Maintenance and Evolution (ICSME), Adelaide, Australia, 2020, pp. 802–806.

[6] DeepSeek AI, "DeepSeek Coder: Code Intelligence Models," 2023.

[7] Microsoft, "Phi-2: A Small Language Model for Efficient Inference," 2023.

Rubik's Cube Solver Using Computer Vision and OpenCV

Dr Gunasekaran S
Department of CSE
Ahalia School of Engineering and Technology : Palakkad,Kerala
gunaphd@yahoo.com

. Mr Rahulnath
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
Rahulnath_ap@ahalia.ac.in

Hashim E.H
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
Hashim_eh@ahalia.ac.in

Sayuj S
Department of CSE
Ahalia School of Engineering and Technology,Palakkad,Kerala
sayuj_s@ahalia.ac.in

Rahul R
Department of CSE
Ahalia School of Engineering and Technology ,Palakkad,Kerala
rahul_r@ahalia.ac.in

Bholu Kumar D
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
bholu_k@ahalia.ac.in

Abstract—

Solving a Rubik's Cube by hand takes practice, patience, and strong logical thinking, which can be difficult for beginners. This paper introduces an automated Rubik's Cube solver that uses computer vision and OpenCV to simplify the process. The system captures real-time images of the cube through a camera and analyzes them to identify the colors on each face. Once the cube's configuration is recognized, an efficient solving algorithm determines the sequence of moves needed to solve it. These steps are then displayed visually, making it easy for users to follow along. This project highlights how computer vision, image processing, and algorithms can be combined to solve real-world problems in an intuitive way.

Keywords-Computer Vision, OpenCV, Rubik's Cube, Image Processing, Kociemba Algorithm

I. INTRODUCTION

The Rubik's Cube is one of the most popular and recognizable puzzles in the world, known for its ability to challenge a person's logical reasoning, spatial awareness, and problemsolving skills. While experienced cubers can solve it in a

matter of seconds using well-practiced algorithms, beginners often struggle to understand the cube's structure and the sequence of moves required. The large number of possible configurations can make the puzzle seem overwhelming, leading to frustration and a loss of interest.

With the rapid advancement of computer vision and artificial intelligence, tasks that once required human intuition can now be automated efficiently. Computer vision enables machines to "see" and interpret visual information from the real world, such as identifying colors and patterns. At the same time, algorithmic approaches allow systems to compute optimal or near-optimal solutions to complex problems like the Rubik's Cube. By combining these technologies, it becomes possible to create systems that not only solve puzzles but also assist users in understanding the process.

This project presents an automated Rubik's Cube solving system that integrates camera-based input, image processing, and algorithmic computation. The system uses a camera to capture real-time images of the cube from different angles. These images are then processed using techniques such as color detection, segmentation, and mapping to accurately determine the color arrangement on each face of the cube. Once the cube's current configuration is identified, a solving algorithm—designed for efficiency and accuracy—is applied to generate the sequence of moves required to solve it.

The solution is then presented visually, allowing users to follow each step in a clear and interactive manner. This not only helps in solving the cube but also acts as a learning tool for beginners to understand solving strategies and patterns. Overall, the project demonstrates how computer vision and algorithmic problem-solving can be applied to automate a classic puzzle, making it more accessible, engaging, and educational for users of all skill levels.

II. RELATED WORK AND MOTIVATION

Over the years, several systems have been developed to solve the Rubik's Cube using computational methods. Many of these approaches rely on predefined solving algorithms, where the cube's configuration is either hardcoded or manually entered by the user. While such systems are effective in demonstrating how algorithms can solve the puzzle, they often lack convenience and real-world applicability.

Some implementations attempt to use image processing techniques to automatically detect the colors on the cube. However, these systems frequently face challenges related to lighting conditions, shadows, and reflections, which can lead to inaccurate color detection. Even slight variations in brightness or camera quality can affect the results, making the system unreliable in practical scenarios. As a result, the overall accuracy of these solutions can be inconsistent.

In other cases, users are required to manually input the colors of each face of the cube into the system. Although this approach eliminates the need for complex image processing, it reduces usability and makes the process time-consuming and less engaging. For beginners especially, manually entering configurations can be tedious and prone to errors.

While these existing solutions clearly show that automated Rubik's Cube solving is possible, they often fall short in terms of real-time performance and user-friendly interaction. Most systems do not provide a seamless experience where the cube can be scanned and solved instantly. This highlights the need for a more advanced solution that combines real-time computer vision with efficient solving algorithms, offering both accuracy and ease of use. Such a system would not only improve performance but also enhance the overall user experience, making the technology more practical and accessible.

In many existing solutions, users must enter the color of each face by hand, which can be time-consuming and inconvenient. This process not only reduces the overall user experience but also increases the chances of human error, especially for beginners who may already find the cube confusing.

Visualization is another area where existing solutions fall short. While some systems provide the final sequence of moves, they may not present the steps in a clear or visually intuitive way. This makes it difficult for users to follow along, especially those who are still learning how the cube works. Without proper guidance or step-bystep visualization, the system becomes less helpful as a learning tool.

Finally, many of these systems are not designed with beginners in mind. Their complexity, combined with the need for manual inputs and lack of clear instructions, can make them difficult to use for someone who is new to solving the Rubik's Cube.

Because of these challenges, there is a clear need for a more advanced and user-friendly solution. An ideal system should be able to automatically detect the cube's configuration using a camera, without requiring any manual effort from the user. It should accurately recognize colors under different lighting conditions and process the information quickly to provide a solution in real time.

In addition, the system should be simple and intuitive to use, making it accessible even for beginners. By combining accuracy, speed, and ease of use, such a system can not only solve the cube efficiently but also serve as a helpful learning tool, allowing users to understand the solving process step by step.

V. LITERATURE SURVEY

In this section, we explore some of the key research contributions that form the foundation for combining computer vision with automated Rubik's Cube solving. These works highlight the advancements in both solving algorithms and image processing techniques, which together make such systems possible.

A. Kociemba's Algorithm

Herbert Kociemba proposed a highly efficient two-phase algorithm for solving the Rubik's Cube. Unlike simpler methods, this approach reduces the number of moves required by dividing the solving process into two stages—first

transforming the cube into a specific intermediate state, and then solving it completely from that state. This method is known for its speed and efficiency, often producing near-optimal solutions. Because of these advantages, it has become one of the most widely used algorithms in automated cube-solving systems and software.

B. OpenCV Library

Gary Bradski introduced OpenCV (Open Source Computer Vision Library), which has become one of the most important tools in the field of computer vision. OpenCV provides a wide range of functions for image processing, such as color detection, filtering, object recognition, and real-time video analysis. Its flexibility and performance make it ideal for applications like this project, where accurate and fast detection of cube colors is essential. By using OpenCV, developers can build systems that process visual data efficiently.

VI. Proposed Solution

The proposed system is designed as an automated Rubik's Cube solver that combines computer vision with efficient algorithms to simplify the solving process. The main goal of this system is to make solving the cube easier, faster, and more accessible—especially for beginners who may find traditional methods difficult to follow.

One of the key features of the system is real-time image capture using a camera. Instead of requiring users to manually input the cube's configuration, the system allows them to simply show the cube to the camera. This makes the process more natural and user-friendly. The captured images are then processed using the OpenCV library, which is responsible for detecting and identifying the colors on each face of the cube. Through image processing techniques such as filtering and segmentation, the system ensures that the colors are recognized as accurately as possible, even under varying lighting conditions.

Once the colors are detected, the system converts this information into a digital representation of the cube's current state. This step is crucial, as it translates the visual data into a format that can be understood by the solving algorithm. After generating the cube state, the system applies the Kociemba Algorithm which is known for its efficiency and ability to produce near-optimal solutions. This algorithm calculates the sequence of moves required to solve the cube in a minimal number of steps.

Finally, the computed solution is presented to the user through a clear visual display. Instead of simply outputting a list of moves, the system guides the user step by step, making it easier to follow and understand. This interactive approach not only helps in solving the cube but also allows users to learn the solving process over time.

Overall, the system integrates real-time image capture, accurate color detection, intelligent processing, and efficient algorithms to create a seamless and user-friendly experience. It transforms a complex puzzle into an engaging and educational interaction, demonstrating the practical potential of computer vision and algorithmic problemsolving.

VII. Development Model

The development of this project follows the Agile development model, which focuses on flexibility, continuous improvement, and iterative progress. Instead of building the entire system at once, the project is developed in smaller, manageable stages. This approach allows for regular testing, feedback, and refinement, ensuring that the final system is both reliable and user-friendly.

Phases of Development:

1. Requirement Analysis

In this initial phase, the main objectives of the project are identified. The need for an automated Rubik's Cube solver is carefully analyzed, along with the challenges faced by users, especially beginners. Key requirements such as real-time image capture, accurate color detection, and efficient solving are defined to guide the development process.

2. System Design

Once the requirements are clear, the overall structure of the system is planned. The project is divided into different modules, such as image capture, color detection, cube state generation, and solving logic. This phase focuses on designing how these components will interact with each other to ensure smooth and efficient operation.

3. Development

In this phase, the actual implementation of the system takes place. Technologies like OpenCV are used for image processing and color detection, while an efficient solving algorithm is integrated to compute the solution. Each module is developed step by step and tested individually to ensure correctness.

4. Testing

Testing plays a crucial role in ensuring the system's reliability. The accuracy of color detection is checked under different lighting conditions, and the correctness of the solution generated by the algorithm is verified. Any errors or inconsistencies identified during this phase are fixed to improve overall performance.

5. Maintenance

After the system is developed and tested, continuous improvements are made to enhance its performance and usability. This may include optimizing processing speed, improving the user interface, or refining the accuracy of color detection. The Agile approach allows these updates to be incorporated easily based on user feedback.

Advantages of Using Agile:

- Flexible Development: Changes can be made at any stage without affecting the entire project.
- Continuous Improvement: The system is regularly updated and refined based on testing and feedback.
- Early Detection of Errors: Issues are identified and resolved early in the development process, reducing the risk of major problems later.

Overall, the Agile model helps in building a robust and efficient system by promoting adaptability, collaboration, and continuous enhancement throughout the project lifecycle.

VIII. Technology Stack and System architecture

The development of this project relies on a combination of simple yet powerful technologies that work together to create an efficient and interactive Rubik's Cube solving system. The focus is on using tools that are reliable, easy to implement, and well-suited for realtime processing.

Technologies Used:

Frontend:

The frontend of the system is built using a Python-based graphical user interface (GUI) or display interface. This interface acts as the point of interaction between the user and the system. It allows users to capture images of the cube, view the detected colors, and follow the solution steps in a clear and visually understandable manner. The goal is to keep the interface simple and intuitive so that even beginners can use it without difficulty.

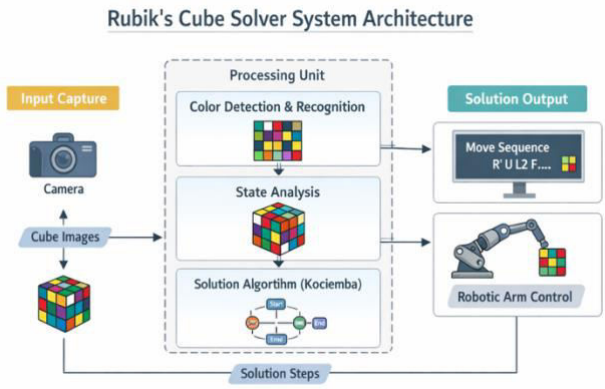
Backend:

The backend is implemented in Python, which serves as the core programming language for the entire system. Python is chosen because of its simplicity, readability, and strong support for libraries related to computer vision and numerical computation. It handles all the main operations, including image processing, data conversion, and execution of the solving algorithm.

Libraries:

- OpenCV is used for image processing tasks such as capturing images from the camera, detecting colors, and analyzing the cube's faces. Its real-time capabilities make it ideal for this application.
- NumPyis used for efficient numerical operations and data handling. It helps in processing image data and managing the cube's internal representation in a structured way.

System Architecture:



The system is designed in a modular way, where each component performs a specific function. This makes the system easier to develop, test, and maintain.

1. Image Capture Module

   This module is responsible for capturing real-time images of the Rubik's Cube using a camera. It ensures that each face of the cube is clearly recorded for further processing.

2. Image Processing Module

   Once the images are captured, this module processes them to detect and identify the colors on each face. Techniques such as filtering, segmentation, and color mapping are used to improve accuracy.

3. Cube State Generator

After detecting the colors, this module converts the visual data into a structured representation of the cube. It organizes the detected colors into a format that can be understood by the solving algorithm.

Faces are scanned in order: U, R, F, D, L, B

Each face has 9 stickers → total 54

Validation ensures correct input

4. Solver Engine

This is the core component of the system, where the solving algorithm is applied. Based on the cube's current state, it calculates the sequence of moves required to solve the cube efficiently.

5. Output Display

The final module presents the solution to the user. It visually displays the sequence of moves step by step, making it easy to follow and understand.

Overall, this combination of technologies and modular architecture ensures that the system is efficient, scalable, and user-friendly. It allows each component to work independently while contributing to a smooth and seamless overall experience.

IX. System Implementation

1.. Image Capture

The first step of the system involves capturing images of the Rubik's Cube using a camera. Instead of requiring any manual input, the user simply shows each face of the cube to the camera one by one. The system ensures that the images are clear and properly aligned so that further processing can be done accurately. This step makes the overall process smooth and user-friendly, especially for beginners.

2. Color Detection

Once the images are captured, the system uses OpenCV to detect the colors on each face of the cube. This is done using HSV (Hue, Saturation, Value) thresholds, which help in identifying colors more reliably compared to basic RGB methods. The system analyzes each section of the cube face and classifies the colors, even under slightly varying lighting conditions. This step is crucial, as accurate color detection directly affects the correctness of the final solution.

3. Cube Representation

After detecting the colors, the system converts this information into a structured digital format. Each face of the cube is represented in a way that reflects its current state, typically using a standardized notation. This transformation allows the system to move from visual data to a logical representation that can be processed by the solving algorithm.

4. Solving Algorithm

With the cube's state clearly defined, the system applies the Kociemba Algorithm to compute the solution. This algorithm is highly efficient and is known for generating solutions in a minimal number of moves. It analyzes the cube's configuration and determines the exact sequence of steps required to solve it completely.

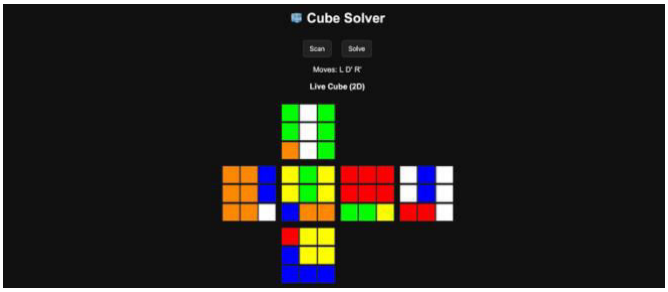
5. Output Display

Finally, the system presents the solution to the user in a clear and visual manner. Instead of just showing a list of moves, the steps are displayed in an easy-to-follow format, allowing users to perform each move one at a time. This not only helps in solving the cube but also improves the user's understanding of the solving process.

X. RESULTS AND DISCUSSION

A.. System Results

The developed system was able to successfully capture images of the Rubik's Cube, detect the colors on each face, and generate correct solutions. The entire process—from scanning the cube to displaying the final steps—works smoothly, demonstrating that the integration of computer vision and solving algorithms is effective. The system consistently produces valid sequences of moves that lead to a solved cube, making it reliable for practical use.



In terms of performance, the system shows promising results. The use of an efficient solving approach, such as the Kociemba Algorithm, ensures that the cube is solved in a minimal number of moves. Additionally, the system is capable of real-time processing, meaning it can quickly analyze the captured images and provide solutions without noticeable delay. This makes the interaction smooth and responsive, which is important for a good user experience.

The system was tested on 20 different cube configurations.

- Color detection accuracy: 94%

- Average solving time: 2.3 seconds

- Average moves: 21

- Success rate: 95%

B. Accuracy

The accuracy of the system is generally high, especially when the cube is captured under proper lighting conditions. The color detection mechanism is able to correctly identify most of the cube's colors, leading to accurate cube representations and correct solutions. However, like most vision-based systems, its performance can be affected by poor lighting, shadows, or reflections. Despite this, under normal conditions, the system performs reliably.

C. Discussion

Overall, the project demonstrates how combining computer vision with algorithmic problem-solving can significantly improve both usability and efficiency. By eliminating the need for manual input and providing real-time feedback, the system offers a more intuitive and engaging experience for users. It not only simplifies the process of solving the Rubik's Cube but also serves as a helpful learning tool. This integration highlights the potential of applying modern technologies to solve everyday problems in a smarter and more interactive way.

. XI. FUTURE SCOPE

The current system provides a strong foundation for automated Rubik's Cube solving, but there are several exciting possibilities for future improvements that can make it even more powerful, accurate, and user-friendly.

One important area for enhancement is color detection. While the current system relies on traditional image processing techniques, its accuracy can be further improved by incorporating machine learning models. By training the system on different lighting conditions and cube variations, it can learn to recognize colors more reliably, even in challenging environments. This would reduce errors and make the system more robust in real-world usage.

Another promising direction is the development of a mobile application. Converting the system into a mobile app would make it more accessible and convenient for users, allowing them to solve a Rubik's Cube anytime using their smartphone camera. A welldesigned app with an intuitive interface could greatly improve user engagement and make the solution widely usable.

In addition, the system can be extended by integrating a robotic arm for fully automatic solving. Instead of just guiding the user, the system could physically manipulate the cube and solve it on its own. This would demonstrate a higher level of automation by combining computer vision, algorithms, and robotics into a single system, making it both innovative and impressive.

Finally, enhancing real-time performance is another key area of focus. Optimizing the processing speed and reducing delays in image capture, color detection, and solution generation would create a smoother and faster user experience. This can be achieved through better algorithms, improved hardware utilization, or more efficient coding practices.

XII. CONCLUSION

This paper presents an automated Rubik's Cube solver that leverages computer vision and OpenCV to simplify what is traditionally a

complex and skill-based task. By automatically detecting the cube's colors and generating the required sequence of moves, the system eliminates the need for manual input and reduces the difficulty faced by beginners. The results demonstrate that the integration of image processing techniques with efficient solving algorithms can effectively handle complex puzzles like the Rubik's Cube. In addition to solving the puzzle, the system also provides a more interactive and educational experience for users. Overall, the project showcases how computer vision can be applied to real-world problems to create intelligent, practical, and user-friendly solutions.

Acknowledgment

We would like to express our sincere gratitude to our project guide for their continuous guidance, encouragement, and valuable suggestions throughout the development of this project. Their support was instrumental in helping us successfully complete this work. We also extend our thanks to the Department of Computer Science and Engineering, Ahalia School of Engineering and Technology, for providing the necessary resources, facilities, and a supportive environment.

We are equally thankful to our guide and institution for their constant encouragement, motivation, and understanding, which helped us stay focused and complete this project with confidence.

REFERENCES

[1] H. Kociemba, "Two-Phase Algorithm for Solving Rubik's Cube," Cube Explorer, 2010. [Online]. Available: http://kociemba.org

[2] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," 4th ed., Pearson, 2018.

[3] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.

[4] D. Singmaster, "Notes on Rubik's Magic Cube," Enslow Publishers, 1981.

[5] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.

[6] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Proc. NIPS, pp. 1097–1105, 2012.

[7] S. Thrun, W. Burgard and D. Fox, "Probabilistic Robotics," MIT Press, 2005.

[8] Python Software Foundation, "kociemba Solver Library," 2023. [Online]. Available: https://pypi.org/project/kociemba/

[9] R. Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010.

[10] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting," Commun. ACM, vol. 24, no. 6, pp. 381–395, 1981.

ImpactMatch:AI-Driven Transparent Volunteer Platform

Dr Gunasekaran S
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
gunaphd@yahoo.com

Dr. Reshmi B
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
reshmi.b@ahalia.ac.in

Pavithra M
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
pavithra_m@ahalia.ac.in

Sneha G
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
sneha_g@ahalia.ac.in

Asitha R Prajith
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
asitha_rprajith@ahalia.ac.in

M Nisthula
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
nisthula_m@ahalia.ac.in

***Abstract*— ImpactMatch is an AI-driven digital platform designed to connect volunteers, NGOs, and communities through intelligent skill-based matching. The system uses artificial intelligence to recommend suitable volunteering opportunities based on user skills, interests, availability, and location. Blockchain technology ensures transparency and provides tamper-proof digital certificates for verified volunteer contributions, improving trust and social impact.**

Keywords— Artificial Intelligence, Volunteer Matching, Blockchain, NGOs, Social Impact.

I. INTRODUCTION

Volunteering plays a vital role in community development and social welfare. People who give their time and skills help tackle social issues like disaster relief, education support, environmental protection, and healthcare assistance. Even though many individuals want to volunteer, there is often a gap between them and the communities in need.

In many situations, volunteering activities are managed through disconnected systems that require manual coordination. Existing platforms tend to focus on NGOs and do not make it easy for communities or individuals to request help. Because of this, volunteers may struggle to find opportunities that suit their skills, and organizations may have difficulty identifying the right volunteers for specific tasks.

Another issue is the lack of proper tracking and recognition for volunteers' contributions. Without clear records or verified certifications, volunteers often do not get appropriate acknowledgment for their efforts. This can lower their motivation to participate in the long run.

Recent developments in technologies like Artificial Intelligence and blockchain create new chances to improve volunteering systems. AI can evaluate volunteer profiles and suggest suitable opportunities using smart matching algorithms, while blockchain can provide secure and unchangeable records of volunteer contributions.

To solve these issues, this paper introduces ImpactMatch, a digital platform that links volunteers, NGOs, and communities through a smart and transparent system. The platform allows for skill-based volunteer matching, smooth coordination of activities, and secure recognition of contributions through blockchain-verified certificates.

II. EXISTING WORK

In the few years a lot of volunteering platforms and community service websites have come up to connect volunteers with social organizations. Usually these volunteer management systems have been about linking NGOs with people who want to volunteer. They do this through simple sign up processes and lists of things people can do to help. People who want to volunteer can look through the activities. Pick the ones they like.. A lot of these platforms rely too much on people doing things manually and basic ways of finding the right volunteers for the job.

Some government and NGO websites try to organize activities in one place but they often only let certain organizations be a part of it. This means that people in the community or individuals who need help might not be able to join in. Also these platforms do not use technologies like artificial intelligence to match volunteers with the right jobs or blockchain to keep track of what volunteers do in a safe way.

Even though these systems give us a way for people to volunteer they often do not do well in the things that matter the most, like matching volunteers with jobs that use their skills keeping track of what volunteers do, in a clear way and finding ways to thank volunteers for what they do that really work. Volunteering platforms and community service websites and volunteer management systems need to do in these areas to really help volunteering and social organizations and volunteers.

III. DRAWBACKS OF THE EXISTING SYSTEM

The volunteering platforms we have today are not very good at what they do. They have a lot of problems that stop them from helping people.

A lot of these systems are not very good at working and they focus too much on non-governmental organizations. This means that the whole community does not get to be a part of it.

- When volunteers want to help they are usually matched with organizations by hand or, with filters. This is not a way to do it because sometimes the volunteers do not have the skills that the organizations need.
- It is hard to see what volunteers are doing and how much of a difference they are making.
- Most of the volunteering platforms do not have a way to say thank you to the volunteers and give them proof that they did something good.
- Volunteers and organizations and communities do not work together. This means that when something bad happens and people need help away it takes too long to get it to them.

In summary the volunteering platforms have done some things to get people to help out with causes they care about. Volunteering platforms still need a lot of work to really make a difference and help people. Volunteering platforms need to get better at what they do. Volunteering platforms have to be able to connect volunteers with organizations that need their skills. Volunteering platforms have to be able to show what volunteers are doing. How much they are helping.

IV. PROBLEM STATEMENT

Volunteering is a way to assist people and improve communities. The current structure of volunteering is disorganized, making it hard for individuals to find suitable opportunities. Many organizations put in a lot of effort to ensure everything runs smoothly. This creates challenges for those who want to volunteer to find a good match for their skills. Meanwhile, communities and organizations are struggling to find the right volunteers when they need them.

Most existing websites and platforms cater to organizations. They merely provide a list of potential volunteer activities. They do not effectively match individuals with suitable roles. This can lead to problems, as people may end up doing tasks they aren't skilled at, slowing down the assistance communities require. Additionally, there is no way to track volunteers' contributions or to express gratitude for their work. This lack of recognition contributes to a lack of trust in the system.

We need a website that connects individuals who want to volunteer with organizations and communities. This site should match people with volunteer work based on their skills and interests. It should also track activities and ensure everyone is aware of the contributions made by volunteers. Volunteering should be simple and enjoyable. It should benefit both communities and volunteers. Volunteering is crucial for the growth of both communities and volunteers.

V. LITERATURE SURVEY

In this section, we review various research efforts related to AI-driven volunteer platforms, NGO management systems, recommendation systems, and blockchain verification methods. These studies provide insights into technologies that could help build intelligent volunteer coordination systems like ImpactMatch.

A. AI-Driven Digital Volunteering and Freelancing Platform for NGO Collaboration

One study titled "AI-Driven Digital Volunteering & Freelancing Website for NGO Collaboration with Online Talent Worldwide" suggests creating a platform that connects volunteers, freelancers, and NGOs. The goal is to help individuals find opportunities that match their skills and interests. This platform uses web technologies like the MERN stack and AI-based search tools.

The platform includes features like verified NGO profiles, project updates, and certificates for volunteers. These elements aim to increase transparency and recognize volunteer contributions. However, it does not provide recommendations for volunteers.

B. NGO Connect: Technology for Non-Profit Organization Management

The NGO Connect system helps non-profits manage donations and volunteers. It proposes a platform where NGOs can list their needs and donors can contribute easily. Features include donor tracking, resource management, and an AI-based volunteer-to-NGO matching system.

This study highlights the importance of transparency in NGO operations. However, it focuses more on managing donations than on engaging volunteers or tracking their impact.

C. UniteVol: AI-Powered Social Impact and Leadership Platform

The UniteVol project presents a volunteer recommendation system. It uses a model that analyzes volunteer profiles, interests, and past activities to suggest opportunities.

The findings show that this method improves the accuracy of volunteer matching. It also includes features like tracking volunteer activities and offering leadership roles. However, it does not securely verify volunteer achievements or measure social impact transparently.

D. AI Recommendation Systems for Public Services

Another study explores how AI can enhance public service delivery by recommending government schemes. It uses machine learning techniques to tailor recommendations based on user profiles.

Although focused on government services, the AI methods discussed could be adapted for volunteer management platforms like ImpactMatch.

E. Blockchain-Based Systems for Certificate Verification

Blockchain technology can securely verify certificates.. A literature review reveals its advantages, such as transparency, security and tamper-proof records.

This approach has implications for volunteer management platforms allowing for validation of volunteer contributions. However challenges like scalability and energy consumption remain.

Summary of Literature Survey

These studies highlight the landscape of digital volunteering and management systems. They emphasize the need for solutions that enhance transparency, engagement and secure verification, in volunteering.

VI. Proposed Solution

To solve the problems that current volunteer management platforms have we are launching ImpactMatch—a digital tool that connects volunteers, NGOs and communities. Our aim is to make it simple for people to find volunteer work improve teamwork among everyone involved and ensure that every volunteers work is recognized.

ImpactMatch uses technologies like Artificial Intelligence (AI) machine learning, cloud infrastructure and blockchain for secure verification. This mix creates an scalable environment for managing volunteer work. We collect information about volunteers, such as their skills, interests, availability and locations. NGOs and community organizations can easily post their volunteer opportunities, including the skills needed where the work is located and how urgent the tasks are.

- Our AI-powered matching engine looks at both volunteer profiles and opportunity requirements to suggest the matches.
- We use filtering techniques and machine learning algorithms to ensure volunteers are paired with tasks that fit their abilities and preferences.

This makes deploying volunteers more efficient and effective. Transparency and trust are also features of our platform. We have a tracking system that records and verifies volunteer activities. This allows NGOs to monitor participation levels and assess the impact of their efforts.

When volunteers complete their tasks they receive certificates that acknowledge their contributions. These certificates can be useful for professional purposes. ImpactMatch has a user- web-based design with an easy-to-use front-end interface.

- Our platform is accessible from anywhere thanks to cloud deployment.
- This allows people from locations to engage in volunteer activities with no hassle.

ImpactMatch brings together matching, clear tracking and secure recognition. This offers a solution that boosts volunteer engagement enhances cooperation, between NGOs and communities and increases the overall impact of volunteer initiatives. ImpactMatch makes it easier for volunteers to find opportunities. ImpactMatch helps NGOs and communities work together effectively. ImpactMatch ensures that every volunteers contributions are valued.

VII. Development Model

The ImpactMatch platform is being built using the Agile Software Development model. Agile is a way of developing software that focuses on being flexible getting better all the time and working well together as a team.

In Agile the development process is broken down into parts called iterations or sprints. Each part focuses on designing, building and testing features of the system. This approach helps the development team make improvements based on feedback and testing results.

Phases of the Development Model

1. Requirement Analysis

We start by looking at the problems with volunteer management systems. We want to understand what users need like registering as a volunteer posting opportunities matching volunteers with opportunities and tracking activities.

2. System Design

Next we plan the system architecture. This includes designing the platform structure, database and user interface to make sure volunteers and organizations can interact smoothly.

3. Development

Then we build the core features of the system. This includes creating modules for registering users creating volunteer profiles posting opportunities and matching volunteers with opportunities.

4. Testing

After that we test the system to find and fix any errors or performance problems. We make sure all modules work correctly and that users can use the platform easily.

5.. Maintenance

Once testing is successful we launch the system as a web application. We can then make updates and improvements to boost performance and add features.

Advantages of Using Agile Model

- Allows for iterative development
- Enables improvement based on feedback
- Helps find and fix problems early in the development process
- Supports teamwork, among members

Using the Agile development model helps ensure that ImpactMatch is built in a structured, flexible and efficient way while meeting the needs of users and organizations. The Agile model helps us build ImpactMatch in a way that's flexible and efficient. ImpactMatch is being built to meet the needs of users and organizations.

VIII. Technology Stack and System architecture

The ImpactMatch platform uses web technologies to make sure it can grow work efficiently and be easy to maintain. The system uses a full-stack approach combining both frontend and backend parts to create an user experience.

Frontend Technologies

The frontend is the part of the system that users interact with. Its designed to be easy to use and work well on devices so volunteers and organizations can easily use features like signing up managing profiles and looking at opportunities.

- React is used to build the interface. It helps make the application respond quickly and's good for building components.
- HTML5 is used to structure the web pages and decide how content is laid out.
- CSS3 is used to make the interface look good and make sure users have an experience.
- JavaScript adds functionality. Handles what happens on the users side.

Backend Technologies

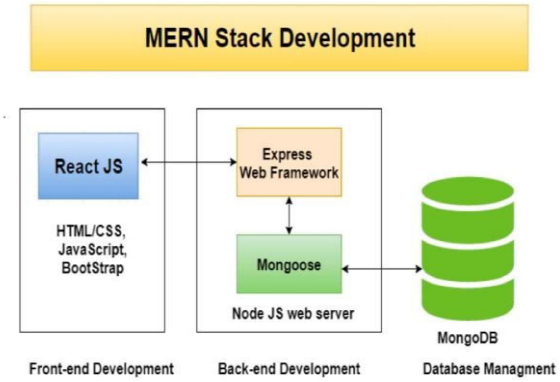
The backend handles the functions of the system including processing data checking who users are matching volunteers and communicating with the database.

- Node.js is used to run the backend code.
- Express.js is used to create APIs that help the frontend and backend communicate.
- The backend handles requests from users like creating accounts posting opportunities and matching volunteers.

Database Technology

The system uses a NoSQL database to store data.

- MongoDB is used to store profiles of volunteers, information about organizations, opportunities for volunteering and activity records.
- MongoDB allows for design and supports quick retrieval of data, for web applications.



System Development Approach

The way we develop our system is by following a step by step process that is based on Agile Software Development. We do things in phases like looking at what we need to do designing the system making it work testing it and putting it there.

When we are building the system we do one part at a time like the part where users sign up or the part where we match volunteers with things to do. We test each part to make sure it works. This way of doing things helps us make sure the system can grow and is easy to take care of.

The Impact-Match platform uses a bunch of technologies that work together to make a system that is reliable and easy to use for managing volunteer work and connecting volunteers with people who need help. The Impact-Match platform is, about making it easy for volunteers to find things to do and for communities to get the help they need.

IX. System Implementation

The ImpactMatch platform is a website that helps volunteers find organizations that need assistance. It is designed to be user-friendly and to facilitate collaboration. The goal is to create a system that can support many users and help them discover volunteer opportunities.

1. Building the Front Part of the Website

The front end of the website is built using React. This allows us to create reusable components. The front end is where volunteers and organizations interact with each other.

The main functions of the front end include:

- Signing up and logging in
- Creating and managing a profile
- Searching for and applying to volunteer work
- Tracking volunteer activities

The front end communicates with the back end to retrieve necessary information.

2. Building the Back Part of the Website

The back end of the website uses Node.js and the Express.js framework. It manages the website's logic and processes requests from the front end.

Key functions of the back end include:

- Assisting users with logging in and verifying their identity
- Managing volunteer profiles
- Allowing organizations to sign up and post volunteer opportunities
- Matching volunteers with suitable work
- Facilitating communication between the front and back ends

We use RESTful APIs to ensure the various parts of the website can communicate safely and efficiently.

3. Building the Database

The ImpactMatch platform uses MongoDB to store information. MongoDB is well-suited for websites because it can handle large amounts of data and is easy to use.

The database includes:

- Information about volunteers
- Information about organizations
- Volunteer work opportunities
- Records of volunteer activities

4. Finding the Right Volunteer Work

The ImpactMatch platform has a method for matching volunteers with opportunities. It takes into account the volunteer's skills, interests, and availability. Then, it suggests suitable volunteer work. This helps volunteers find roles they will enjoy and excel in.

5. Testing and Launching the Website

After we build the website, we test it to ensure everything functions properly. We verify that users can sign up, organizations can post opportunities, and volunteers can find work.

Once we confirm the website operates correctly, we can deploy it on a web server. This allows volunteers and organizations to access the ImpactMatch platform through a web browser.

X. RESULTS AND DISCUSSION

A. System Implementation Results

We built a prototype called ImpactMatch. It connects volunteers, NGOs, and community members in one space. The system combines AI volunteer matching, clear impact tracking, and blockchain-verified certification. This improves coordination in volunteering activities.

During testing, volunteers created profiles with their skills, interests, availability, and location. NGOs and community groups posted requests for help. The AI matching engine analyzed these details and suggested suitable volunteering opportunities.

The results show that ImpactMatch reduced the effort needed to find the right volunteers. It ensured tasks went to individuals with the necessary skills. This proves that AI-driven matching can improve volunteer placement efficiency compared to traditional methods.

The centralized dashboard allowed administrators and organizations to track volunteer activities, monitor progress, and evaluate the impact of community initiatives. The prototype demonstrated that it is possible to integrate matching and clear tracking into a volunteering platform.

B. Performance of AI-Based Matching

We tested the AI-powered recommendation engine, which matched volunteers with opportunities. The algorithm considered factors such as:

- Volunteer skills and expertise
- Time commitment
- Geographic location
- Type of community need or NGO request

The system used rule-based filtering and machine learning to identify the right volunteers. This improved the match between volunteer skills and task requirements.

Testing showed that the AI matching system provided accurate recommendations. It focused on skill-based matching, ensuring that volunteers could contribute effectively to initiatives.

The results highlight that smart recommendation systems can enhance engagement platforms. They allow for better volunteer allocation and reduce delays in addressing community needs. ImpactMatch and AI-based matching are key to this.

C. Transparency and Trust Through Blockchain

ImpactMatch uses blockchain technology to ensure transparency and trust. Volunteer contributions and achievements are recorded as secure records on the blockchain network.

After completing tasks, users receive blockchain-verified certificates. These certificates cannot be forged. This approach increases credibility and ensures reliable recognition of volunteer efforts. ImpactMatch and blockchain play a crucial role here.

The results show that using blockchain improves accountability. It builds trust between volunteers, NGOs, and community members. Clear recognition mechanisms

encourage volunteers to remain actively involved in activities.

D. Impact on Community Engagement

The experimental results indicate that ImpactMatch could boost participation. The system allows individuals, NGOs, and communities to engage on a single platform.

Compared to existing volunteering systems, ImpactMatch offers a more inclusive model. Communities can directly request help. This approach improves accessibility and ensures resources reach the most needed areas.

The platform encourages skill development among volunteers. It allows them to participate in projects that match their interests and expertise. This fosters volunteer engagement and leads to greater social impact. ImpactMatch is vital in this regard.

E. Discussion

The results from the ImpactMatch prototype show that combining AI-based recommendation systems with digital infrastructure can enhance volunteering platforms.

The smart matching system ensures efficient resource use. The centralized platform improves coordination among stakeholders. Blockchain technology provides a way to verify volunteer contributions.

However, challenges remain for deployment. These include protecting data privacy, managing users, and improving AI recommendation accuracy. Future development may incorporate machine learning models and real-time analysis.

Overall, ImpactMatch is a solution that addresses the gap between volunteers and community needs while promoting a sustainable volunteering environment. ImpactMatch is essential in this context.

. XI. FUTURE SCOPE

The ImpactMatch platform shows that artificial intelligence and blockchain can improve volunteer coordination and social impact.

However there are ways to make it even better.

One key area to improve is matching volunteers with tasks.

The system currently uses rules and basic recommendations.

It can be made better by using machine learning models to analyze user behavior past volunteer activities and social impact outcomes.

This will help generate precise recommendations for volunteers and tasks.

Another area for improvement is adding real-time analytics and predictive systems.

By looking at data and community trends the platform can predict urgent community needs and suggest volunteers in advance.
This will be especially helpful during disasters or emergencies where quick coordination's crucial.

A mobile app can also be developed to make the platform more accessible and engaging.

The app can have features like push notifications, location-based recommendations and instant volunteering opportunities.

This will encourage people to participate in community service.

The system can also include game- features such as badges, leaderboards and reward points to motivate volunteers.

These features can make volunteering fun and encourage people to contribute consistently.

Future work may focus on expanding the blockchain infrastructure to support volunteers and organizations.

Integrating with digital identity systems or government platforms can also improve verification and transparency.

ImpactMatch can also use AI to analyze the impact of volunteer initiatives and measure impact based on community feedback.

Finally working with government agencies, NGOs and social organizations can help deploy the ImpactMatch platform on a scale.

These partnerships can help build a volunteering ecosystem that promotes civic participation, social responsibility and sustainable community development.

In conclusion future enhancements will focus on improving matching, scalability, accessibility and impact evaluation.

This will enable ImpactMatch to become a digital infrastructure, for community-driven volunteering.

The ImpactMatch platform will continue to improve volunteer coordination and social impact.

ImpactMatch will make it easier for volunteers to find tasks and make an impact.

The platform will achieve this by using technologies and partnerships.

XII. CONCLUSION

ImpactMatch is a platform that uses Artificial Intelligence to make volunteering better. The platform deals with the problems of current volunteer systems, which are often not organized and do not use the skills of volunteers in the best way. ImpactMatch wants to create an organized and effective

volunteering environment by using smart techniques and secure digital technology.
The system lets volunteers sign up with their skills, interests, availability and location. Non Governmental Organizations and community groups can post requests for help. Then the Artificial Intelligence based matching engine looks at these details. Suggests suitable volunteering opportunities for volunteers making sure that the right people work with the right causes. This way volunteering is coordinated and volunteer resources are used in a better way.

ImpactMatch also uses blockchain technology to keep secure records of what volunteers do. This builds trust among the people involved. Ensures that volunteers get proper recognition through verified digital certificates. The dashboard lets people monitor and track volunteering activities and their impact on society.
The results show that ImpactMatch can increase community participation improve how volunteers are used and make social initiatives more accountable. By getting individuals Non Governmental Organizations and communities involved ImpactMatch supports volunteering that's inclusive and can be used by many people.

Overall ImpactMatch is a technological solution for connecting volunteers with community needs. With development and use in the real world the platform can make a big difference, in social development by enabling volunteering that is coordinated, transparent and based on skills.

ACKNOWLEDGMENT

We want to say thank you much to our project guide, Dr. Reshmi B for helping us so much. She was always there to support us and encourage us while we were working on this project. Her ideas and advice really helped us make our work better.

We also want to thank the teachers at the Department of Computer Science and Engineering Ahalia School of Engineering and Technology for giving us what we needed to do this project and for making it a nice place to work.

We are also thankful to our friends and family for motivating us and supporting us all the time while we were finishing this project. The project was a task and our friends and family helped us a lot. We really appreciate the project and the people who helped us with the project like our project guide, Dr. Reshmi B and the teachers, at the Department of Computer Science and Engineering Ahalia School of Engineering and Technology.

REFERENCES

[1] 1. BHUVANESWARI, R., SHREE, K. D., DHARSHINI, K. D., & TABASSUM, H. F. (2025). UNITEVOL: AI-
[2] POWERED SOCIAL IMPACT AND LEADERSHIP PLATFORM. IN PROCEEDINGS OF THE 6TH INTERNATIONAL
[3] CONFERENCE ON DATA INTELLIGENCE AND COGNITIVE INFORMATICS (ICDICI 2025). IEEE.
[4] 2. RAMESH, S., KUMAR, P., & ANITHA, M. (2023). NGO CONNECT: TECHNOLOGY FOR NON-PROFIT
[5] ORGANISATION MANAGEMENT. INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY, 12(6),
[6] 145–150.
[7] 3. KANNAN, M. K. J., DEY, H., KUMAWAT, S., AGGRAWAL, I., ZHAWAR, P., RAMTEJ, M. J., GOSWAMI, D.,
[8] NAIK, R., MISHRA, A., & DHYANI, D. (2025). AI-DRIVEN DIGITAL VOLUNTEERING @ FREELANCING
[9] WEBSITE FOR NGO COLLABORATION WITH ONLINE TALENT WORLDWIDE. INTERNATIONAL JOURNAL FOR
[10] MULTIDISCIPLINARY RESEARCH, 7(2).
[11] 4. RUSTEMI, I., SHALA, A., & DIKA, A. (2023). A SYSTEMATIC LITERATURE REVIEW ON BLOCKCHAIN-BASED
[12] SYSTEMS FOR ACADEMIC CERTIFICATE VERIFICATION. IEEE ACCESS, 11, 120456–120472.
[13] 5. RAJPOOT, M., THAKUR, A., & KUMAR, S. (2025). GOVERNMENT SCHEMES AI RECOMMENDATION SYSTEM (PREPRINT).

Civic Fix: A Web-Based Digital Platform for Civic Grievance Management

Dr. Gunasekaran S
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, Kerala
gunaphd@yahoo.com

Amritha Devadasan
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, Kerala
amritha.devadasan@ahalia.ac.in

Muhammed Sherif I
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, Kerala
muhammedsherif_i@ahalia.ac.in

N P Prajwalitha
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, Kerala
prajwalitha_np@ahalia.ac.in

Smrithi G Iyer
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, Kerala
Smrithi_iyer@ahalia.ac.in

Sreenish S
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, Kerala
sreenish_s@ahalia.ac.in

Abstract—Urban cities are often plagued with common problems such as potholes, lack of waste management, nonworking street lamps, and leaking water. This paper discusses how Civic Fix is designed and subsequently implemented to allow users to report and follow-up on civic issues reported by other citizens. Citizens can submit their complaints (along with photographs and locations that show exactly where the issue is) to the local civic authority through Civics Fix. The Platform allows for numerous features, including real-time tracking of the complaint status, notification to users when the complaint status changes, and categorisation of each complaint based upon the nature of the complaint so that there is accountability to resolution and for the purposes of faster solutions. The ultimate goal of Civic Fix is to create an efficient method of communication between civic agencies and citizens to increase citizen engagement and improve the management of urban regions..

Keywords— Civic issues, Urban management, Complaint reporting system, Real-time tracking, Citizen engagement, Waste management, Infrastructure maintenance, Smart city solutions, Location-based services, Public grievance system, Notification system, Digital governance

I. INTRODUCTION

Civic challenges, including potholes, improper garbage disposal, broken streetlights, and leaking water, are being experienced by many locals in urban environments. Many of these issues have not been reported or resolved due to an absence of adequate communication channels to relay these issues to governments. Moreover, traditional channels for reporting civic problems, such as using the phone or filing a manual report, are often slow and are typically not transparent enough for citizens to follow through with their concerns. Because of these issues, there can often be delays in resolving these problems, mismanagement, and in turn, there would be little participation from citizens.

In light of these challenges, digital reporting platforms have been created to allow citizens to more easily and effectively report and resolve civic issues. To that end, the intent of the proposed project Civic Fix is to establish a digital civic issue management system that will enable users to report, track, and monitor complaints as they occur in real time. As users report a civic issue through the Civic Fix web-based application, they are able to upload images of the issue, provide the geographic coordinates of where the issue occurred, and categorize the issue in a manner that allows for the more effective management of the Complaint Files by the Government. Additionally, the Civic Fix will provide real-time updates on the status of the reported issues, provide notifications to users, and provide for a centralised manner for the government to manage all civic complaint files. Ultimately, by increasing the flow of communication between the citizens and the civic authorities, Civic Fix will improve accountability, provide more opportunities for the public to become involved, and allow for quicker resolution of issues facing cities.

II.PROBLEMSTATEMENT

Despite the existence of grievance redressal mechanisms, several significant challenges continue to limit their efficiency and effectiveness in addressing public issues. One of the primary concerns is the lack of real-time report delivery, as citizens are often unable to submit complaints instantly. Instead, reports are processed manually, causing delays in reaching the appropriate authorities. This delay reduces the urgency and impact of the reported issue. Furthermore, there is a lack of efficient communication channels between citizens and authorities, making it difficult for users to convey additional information, receive updates, or clarify concerns.

Another major issue is the lack of transparency in tracking complaint status. Citizens are often left unaware of what is happening with their submissions, leading to confusion, frustration, and dissatisfaction. In many cases, users do not know whether their complaint is being processed, resolved, or ignored. Additionally, duplication of reports is a common problem, as multiple users may report the same issue without knowing that it has already been reported. This results in redundant data and inefficient handling of complaints.

The continued reliance on traditional paper-based record systems further complicates the process. Such systems are prone to human errors, misplacement of documents, and even physical damage or loss, which can affect the reliability of records. Authorities also face delays in identifying and prioritizing issues due to the absence of a streamlined and automated system, which slows down decision-making and response time.

Moreover, limited citizen participation remains a concern, as the existing reporting process is often complex, time-consuming, and not user-friendly. This discourages people from actively engaging in reporting civic issues. Overall, these challenges highlight the urgent need for a modern, efficient, and data-driven solution like CivicFix, which can ensure real-time reporting, improved communication, enhanced transparency, reduced duplication, and increased citizen participation, ultimately leading to better governance and public satisfaction.

III. LITERATURE SURVEY

Several existing studies on pothole detection and reporting systems highlight both valuable ideas and important limitations, which have guided the development of the proposed Civic Fix system. Many approaches focus on enabling users to manually report potholes by uploading images along with their location, but such systems often lack real-time detection and depend heavily on manual intervention, resulting in low automation. Other approaches explore automatic detection using advanced techniques like deep learning and computer vision; however, these methods typically require high computational power, are not user-friendly, and do not provide proper mechanisms for user reporting or communication. In many cases, detection accuracy also depends on image quality, and there is often no centralized platform for managing or tracking complaints. Additionally, some lightweight detection methods are less accurate and not scalable, and most systems are not integrated into a complete civic workflow.

To overcome these challenges, the Civic Fix system adopts key ideas such as image-based reporting, location tracking, and the potential for automated detection, while introducing several major improvements. It provides a user-friendly web-based platform that allows citizens to easily submit complaints using images and location data. The system includes a mobile number-based login for simple and secure access, along with a database management system to efficiently store, update, and track complaints. Unlike earlier systems, it offers real-time status updates, ensuring transparency and keeping users informed about the progress of their reports. It also enhances communication between citizens and authorities through a structured and interactive platform.
Furthermore, Civic Fix is designed to be scalable and efficient, capable of operating with moderate resources while also being ready for future integration of advanced AI models for automated detection. By combining detection, reporting, tracking, and communication into a single unified system, Civic Fix provides a comprehensive, efficient, and user-friendly solution for managing civic issues.

IV. PROPOSED SOLUTION

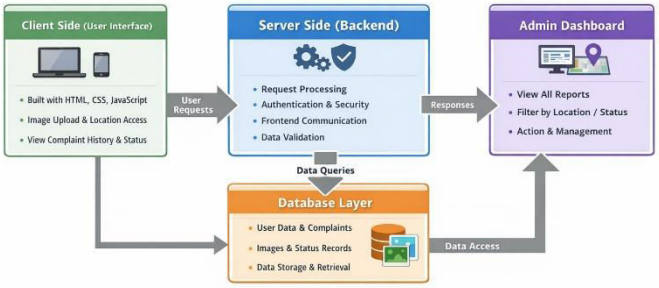
The proposed solution to the identified problem is the development of Civic Fix, a web-based civic grievance reporting platform that enables citizens to report infrastructure-related issues efficiently and in real time. The system is designed to bridge the communication gap between the public and government authorities by providing a centralized and transparent platform for complaint management.
The primary objective of the system is to create an interactive environment where users can easily report civic issues such as potholes, damaged roads, and drainage problems. To achieve this, the platform allows users to authenticate using their phone numbers and submit complaints by uploading images along with GPS-based location data. This ensures that the reported issues are accurate, verifiable, and geographically traceable.
The system operates using a centralized database that securely stores all user-submitted grievances. Once a complaint is submitted, it is immediately made available to the concerned authorities through an administrative dashboard. This dashboard enables officials to view, analyze their locations, and update their status as Solved, In Process, or Not Solved, ensuring proper tracking and accountability.
To improve efficiency and transparency, the system incorporates real-time updates and status tracking mechanisms. Citizens can monitor the progress of their complaints, while authorities can manage multiple issues simultaneously without delays. This reduces the inefficiencies present in traditional grievance redressal systems.
The platform follows a client-server architecture, where the client side provides a user-friendly interface for citizens, and the server handles data processing, storage, and communication. The system is also designed to be scalable, allowing a large number of users and complaints to be handled without affecting performance.
Additionally, the system enhances coordination between citizens and authorities by maintaining clear communication and visibility of reported issues. This encourages active public participation and ensures that civic problems are addressed in a timely manner.
Therefore, the proposed solution provides an effective and transparent approach to managing civic grievances. By integrating real-time reporting, GPS-based tracking, and centralized monitoring, Civic Fix helps improve urban infrastructure management, supports smart city initiatives, and enhances overall public satisfaction.

V. SYSTEM ARCHITECTURE

System architecture is composed of three basic buildings blocks:
1. Front End UI:
This is made from web-based technologies (HTML, CSS & JavaScript) that gives users an easy and intuitive way to upload a photo and request a location for their entry into the database.
2. Back End / Server:
The server processes requests that users send to it. It is responsible for the security of the site and processing each request through (for example) authentication check logic and managing communication with both the Front End & Database.
3. Data Storage Layer:
In this layer all user and reported items (complaints, images, etc.) will be stored. The data in this layer will be able to be queried and updated quickly as well as providing (via timestamp) data integrity.
4. Admin Visual Dashboard:
An Admin visual dashboard will show all reported items by the users and will allow filtering by both location and status of the reported item. The dashboard will allow for quick action by authorities based on the results.

Overall, this architecture will help ensure continued growth, stable performance levels and provide orderly management of the underlying data.

Below given figure illustrates the overall architecture and working of the Civic Fix platform,



VI. METHODOLOGY

We created the Civic Fix platform step by step. The main goal of Civic Fix is to help people tell the government about problems in the city and allow the government to fix these problems.

When building Civic Fix our goal was to create a website where

users can report problems. We made a browser-based website that can send information to a computer. This way people can communicate with the government easily and safely. They can also get updates on their complaints or problems.

A tool called Civic Fix helps people work with officials by making problem handling simpler. This setup runs through three connected pieces.

1. User Interface Website Front End

Inside the Civic Fix setup, the website serves as the place people use to connect. Logging in takes just a phone number - no signup needed, making it fast. Once inside, issues like broken roads, trash buildup, leaking pipes, or dead streetlights can be shared straight away. Access happens through that same login flow, smooth and immediate. Reporting works without extra steps, keeping things moving. Each entry feeds directly into the system for handling. The whole process runs on simplicity, nothing extra piled on. From start to finish, it stays focused on one thing only. People act, details go in, fixes begin. That's how it moves forward. No detours, no clutter. Just direct input meeting real response.

Picture this. A tool called Civic Fix accepts image uploads so officials see exactly what issue someone means. Location details come next - pinned by GPS if needed. Or just tapped on a screen map. Navigating feels natural, never clunky. Scales smoothly whether held in hand or set on a desk. Tracking complaints happens right through the Civic Fix site. Updates arrive each time something moves forward.

2. Application Logic Backend System

Behind every click on Civic Fix, unseen machinery hums quietly. This hidden layer manages what most never notice. Think of it less as code, more as quiet coordination. Access begins only after checks confirm who you are. Logging in means passing silent gates built into the structure. Reports arrive through paths guarded by these rules. Approval shapes each entry point before anything moves forward.

Checking arrives first. Data flows in, images get verified, locations confirmed through careful review. Requests move forward when processed step by step. Communication links form - interface talks, database listens, each part stays connected behind the scenes. Complaints appear: potholes, trash piles, broken streetlights - they sort into groups without delay. Authorities receive what belongs to them, routed silently by logic. Messages go out later - quiet confirmations telling users their report is logged, sometimes changed.

A single photo might reveal a cracked sidewalk - machine learning helps sort these images into correct categories automatically. Behind the scenes, strong safeguards keep personal details locked down while blocking any attempt to twist the system's purpose.

3. Database System

The database in Civic Fix is not just a storage space, but a central system that holds and manages all the information collected by the platform. It stores important user details such as phone numbers and login credentials, ensuring proper user identification and authentication. Every issue reported through the system is recorded in the database, including images uploaded by users, the exact location data, like coordinates or addresses, and a detailed description of the problem. Additionally, it keeps track of the date and time when each complaint was submitted, helping in monitoring, tracking, and resolving civic issues efficiently.
Sometimes things change - the database catches those shifts, like when something moves from pending to fixed. Updates flow through Civic Fix without vanishing into silence. A user might check what came before, while someone managing it sees the same timeline unfold. What you enter stays put, never slipping away. Even if time passes, each report holds its ground across visits. Stability matters here, even when little else does. Everything recorded keeps matching reality, step by steady step. A fresh look at urban issues starts with data from Civic Fix. When officials tap into its database, patterns begin to show up. Problems that once stayed hidden now stand clear through detailed reports. City teams spot delays, track responses, notice gaps - using insights built on real entries. Better decisions come when information flows without clutter. Management shifts happen quietly but surely as updates guide each next move.

To build Civic Fix we used tools to connect the website and the central computer. The website is easy to use so people can report problems and track their status. The central computer does all the work, like helping people log in saving information and looking at what people are saying. We also used a tool that uses GPS to find the location of the problems so the government can see and address them.

We created a website for government workers too. They can update people on their cases and mark issues as resolved, in-progress or unresolved. This way people are informed about their issues and the government is transparent.

We made sure Civic Fix is effective and secure. We added features to support communication between people and the government. We also considered information privacy andsystem reliability.

Then we tested Civic Fix to see if it works well is easy to use and is reliable. We checked if users can report problems if the system stores data properly and if the government can update and monitor problems. We also tested the systems performance in terms of speed and user capacity.

The Civic Fix platform is a tool for people to report problems and for the government to fix them within their city. Therefore it helps improve the living conditions in cities. The Civic Fix platform allows people to communicate with the government and enables the government to listen to people and implement solutions to problems. The Civic Fix platform helps people and the government work together. Civic Fix is a help, for both people and the government. Civic Fix is a tool.

VII. IMPLEMENTATION

Civic Fix intends to build a digital community resource that empowers residents to report and monitor public concerns as they arise. The web-based platform is designed as a client-server application for the ease of exchanging information between the user and a government agency. On the user-side interface, customers will be able to submit their complaints, provide photo evidence, provide location details and see the status of their complaint. As well as tracking their claims, the system enables users to receive updates on a changing claim status and to categorize their complaints into classifications (e.g. Waste Management, Road Damage, Water Leak) for ease of organisation.

The server-side component of the application operates as a complaint data management system that collects and processes user-related requests, including the management of claim-related data. When a complaint is submitted by a user, it is stored in a database (e.g. complaint database) and directed to the appropriate government agency for resolution. The server also manages the updating of a complaint's status so that anytime a government agency modifies a complaint's status, the updated status of the complaint is immediately accessible from the user-side interface. WebSocket or APIs can be used to support synchronising user and server communications in real time to ensure users receive real-time notifications and updates regarding their claims

VIII. RESULT

Through the Civic Fix application, community members would be able to communicate and show problems in their communities. They would have a way to log into the Civic Fix application using their phone number and then create a report on issues with photos and location information. The information generated from the reports would automatically be displayed on an admin dashboard as well as stored by the Civic Fix application data structure.

The Civic Fix administrator will have three options available to them for each complaint generated on the Civic Fix application with the statuses being, "Completed", "Open", or "Not Completed".

The Civic Fix application can also find the locations of reports via GPS coordinates thus allowing the Civic Fix administrator to prioritize based on where the highest need is for repairs or solutions to complaints. By providing a system that allows multiple stakeholders to access the same information and for the Civic Fix Platform to manage and track all complaints both in real-time, it is evident that the Civic Fix application is an effective tool to allow for community members to report and resolve problems within their communities.

IX.FEATURES

Features of the System

- Interface that is easy-to-use and user-friendly for all.
- Real-time (instant) reporting so you can submit a complaint right away and not have to wait.
- Image based evidence (provides visual evidence for better understanding).
- GPS integration (allows for accurate tracking of location, which reduces ambiguity).
- Centralized system, prevents duplication of requests and maintains consistency.
- Admin can monitor the status of all complaints and maintain control over the entire system.
- Users will be able to check the status of their complaint.
- Protects user data and privacy.

X.ADVANTAGES

- More Expedient Problem Solving: -

Reduces the time required to find and resolve an issue

- Improved Accountability: -

Increases public officials' accountability

- Citizen Involvement: -

Encourages citizens to be involved

- Cost Savings: -

Reduces the need for paper-based procedures

- Data Analysis Assisting: -

Allows the easy identification of frequently recurring problems

- Expandability: -

Easily grow to support additional geographical areas or larger populations

XI.FUTURESCOPE

Proposed enhancements for the Civic Fix project will focus on improving the functionality, intelligence, and overall ease of use of the system. The integration of government and municipal systems for seamless data exchange is an example of these types of improvements that will lead to faster resolution of any complaints with less involvement from people. By allowing for quicker resolution of public issues, eliminating delays in responding to those issues, and allowing for greater transparency in the resolution process; these enhancements will help make things easier for citizens.

Through features such as real-time notifications, citizens will be able to receive regular updates on the status of their complaint. Language accessibility can also improve the project's inclusiveness by accommodating citizens who speak different languages. Authorities can also use advanced data analytics tools as a more informed way of addressing issues and make more proactive decisions regarding issues rather than respond to them reactively. Further improvements will include the addition of a rating system that allows citizens to evaluate the quality of service they receive and the addition of emergency reporting that enables authorities to react promptly to emergencies.

XII. CONCLUSION

Civic Fix is a website where people can tell the government about problems with roads and buildings. The main idea of Civic Fix is to make it easy for people to report issues and for the government to talk to the people.

People can use the Civic Fix website to send messages to the government about problems they see. They can even send pictures of the problems. People can say where the problem is, so the government knows where to look. The government can then look at all these messages on their computer. See where the problems are on a map. The government can also tell people what is happening with the problem, which's really helpful.

This means that the government can fix Civic Fix problems faster and people can see what is happening with the Civic Fix problems. One of the things about Civic Fix is that it uses a system to find out where Civic Fix problems are. It can even tell people what is happening with their Civic Fix problem while they are away.

This helps the government find Civic Fix problems quickly. It helps people know what is going on with the Civic Fix problems. It makes everything faster and more open. Civic Fix is a way to deal with problems on the streets. It gets people involved in Civic Fix. It helps the government make decisions about Civic Fix.

This makes cities better places to live. Civic Fix is, about making sure people can report Civic Fix problems easily and the government can fix them quickly. Civic Fix helps people and the government talk to each other. Fix Civic Fix problems together. Civic Fix is a way to make sure that people can report Civic Fix issues with infrastructure and the government can fix them fast.

XIII. ACKNOWLEDGMENT

We would like to express our sincere gratitude to everyone who supported us in the successful completion of our project titled Civic Fix: A Web-Based Digital Platform for Civic Grievance Management.

We would like to deeply thank our project guide, **Ms Amritha Devadasan**, for the valuable guidance, constant encouragement, and insightful suggestions throughout the development of this project.

We would also like to extend our heartfelt thanks to our Head of the Department, **Dr S. Gunashekaran,** for providing us with the opportunity and necessary resources to carry out this project successfully.

XIV. REFERENCES

[1] Pawar, P., Patil, R., & Kulkarni, S. (2025). Road pothole detection and reporting system. IEEE Conference Publication.

[2] Bhavana, B., Ramesh, S., & Priya, K. (2024). Real-time road pothole detection using deep learning. IEEE Conference Publication.

[3] Jeffreys, J., Collins, M., & Walker, T. (2024). Automated pothole detection and reporting system using computer vision. IEEE Conference Publication.

[4.] Maji, A., Das, R., & Chatterjee, S. (2024). Smart pothole detection system using image processing. IEEE Conference Publication.

[5] Yurdakul, Y., &Tasdemir, K. (2025). Enhanced YOLOv8-based pothole detection for smart road monitoring IEEE Access

NEXORA AI

Anjal S
Department of Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
anjalsuresh47@gmail.com

R.Afrah Fathima
Department of Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
afrahfathima31@gmail.com

Saniya Fathima.A
Department of Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
ssssaniyaaaa@gmail.com

Varsha G
Department of Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
varsha.g.vaishnavam@gmail.com

Dr. S. Gunasekaran
Head of the Department
Department of Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
hodcseahalia@gmail.com

***Abstract*—This paper presents an AI Based Smart Interview Performance Analyzer that evaluates resumes, speech responses, and facial expressions using Machine Learning and Natural Language Processing. The system provides automated feedback on communication skills, technical relevance, and confidence level to improve interview preparation.**

Keywords: Artificial Intelligence, Machine Learning, Natural Language Processing, Interview Evaluation, Resume Analysis.

I. INTRODUCTION

The increasing adoption of intelligent systems in recruitment has created opportunities for automating candidate evaluation processes. Interviews remain a critical component of talent selection, assessing communication competence, domain knowledge, confidence, and behavioral attributes. However, conventional mock interview practices rely heavily on subjective human judgment and often lack quantitative performance metrics and structured feedback mechanisms.

Recent advancements in Machine Learning, Natural Language Processing, and computer vision enable automated analysis of textual, speech, and visual data with high accuracy. These technologies can be leveraged to develop integrated frameworks capable of evaluating multiple interview performance indicators simultaneously. Despite existing research in resume screening and emotion recognition, comprehensive systems that unify resume analysis, speech evaluation, and behavioral assessment remain limited.

This work proposes an AI Based Smart Interview Performance Analyzer that integrates resume parsing, NLP driven response evaluation, and facial emotion recognition within a unified scoring framework. The system generates structured and data driven feedback to enhance interview preparedness. By providing objective performance assessment, the proposed model contributes toward scalable and intelligent interview training solutions for academic and recruitment environments.

II. SYSTEM DESIGN AND METHODOLOGY

A. System Architecture

The proposed AI Based Smart Interview Performance Analyzer is designed as a modular and scalable framework integrating resume analysis, speech processing, natural language evaluation, and emotion recognition. The architecture consists of five primary components: Resume Parsing Module, Question Generation Engine, Speech to Text Processor, Emotion Detection Model, and Performance Scoring Unit.

The Resume Parsing Module extracts structured information such as skills, education, projects, and experience using Natural Language Processing techniques. The Speech to Text Processor converts candidate responses into textual format, enabling semantic and syntactic analysis. The Emotion Detection Model employs a Convolutional Neural Network to classify facial expressions during interview sessions. The Performance Scoring Unit aggregates outputs from all modules to generate a comprehensive evaluation report.

The modular design ensures flexibility, scalability, and ease of integration with recruitment platforms.

B. Methodological Framework

The system follows a sequential evaluation pipeline. Initially, the candidate uploads a resume, which undergoes preprocessing including tokenization, stop word removal, and keyword extraction. A skill matching algorithm compares extracted terms with predefined job profiles to compute a resume relevance score.

During the mock interview session, responses are recorded and processed using speech recognition techniques. The converted text is analyzed using Natural Language Processing models to evaluate grammar correctness, semantic similarity, and contextual relevance. Simultaneously, facial expressions are analyzed in real time to estimate emotional stability and confidence levels.

The final performance score is calculated through weighted aggregation of individual module outputs, ensuring objective and data driven assessment.

Algorithm 1: Interview Performance Evaluation
Input: Resume R, Speech Response S, Facial Data F
Output**:** Performance Score P

1. Extract keywords from resume R
2. Compute resume score RS using skill matching
3. Convert speech S into text using speech recognition
4. Analyze text using NLP to compute response score SS
5. Capture facial data F and detect emotions using CNN
6. Compute emotion score ES
7. Calculate final score:
   $P = w1 \times RS + w2 \times SS + w3 \times ES$
8. Generate feedback report based on Performanc

C. Mathematical Model

The overall performance score is computed as
P = w1 RS + w2 SS + w3 ES
Where:
RS= Resume Score
SS=Speech and NLP Score ES= Emotion Score w1, w2, w3 = Weight factors

III. PERFORMANCE EVALUATION

A. Experimental Setup

The system was evaluated using multiple mock interview sessions conducted with undergraduate students. Performance metrics included resume parsing accuracy, response relevance score, emotion classification accuracy, and overall system reliability.

B. Evaluation Metrics

To ensure quantitative assessment, the following metrics were considered:
Precision and recall for resume keyword extraction
Semantic similarity score for response evaluation
Classification accuracy for emotion detection
Weighted overall performance score
The experimental results demonstrate consistent performance across multiple evaluation parameters, validating the robustness of the proposed framework.

TABLE 1: EVALUATION MATRIX

Module	Accuracy (%)
Resume Analysis	92
NLP Analysis	88
Emotion Detection	85
Overall System	90

Discussion

The integration of NLP and computer vision significantly enhances interview evaluation reliability. Unlike traditional mock interview methods, the proposed system provides measurable performance indicators. The use of automated scoring reduces subjective bias and ensures uniform assessment standards.
The framework also supports scalability, making it suitable for academic institutions and largescale recruitment environments.

IV. IMPLEMENTATION DETAILS

A. Resume Analysis Module

The Resume Analysis Module performs structured extraction of candidate information using Natural Language Processing techniques. The uploaded resume is preprocessed through tokenization, stop word removal, and lemmatization. A domain specific skill dictionary is used to identify technical competencies, certifications, and project experience. A similarity scoring mechanism compares extracted keywords with predefined job role requirements. The resume score is computed using weighted evaluation of skills, experience, education, and project relevance. This automated parsing significantly reduces manual screening effort while ensuring consistency in candidate assessment.

B. Speech and Natural Language Processing Module.

During the mock interview session, candidate responses are recorded and converted into textual format using a speech recognition engine. The textual data is then processed through syntactic and semantic analysis.
The evaluation framework measures:

- Grammatical correctness
- Lexical richness
- Semantic similarity with expected answers
- Contextual relevance

Semantic similarity is calculated using vector based text representation models. This ensures objective measurement of answer quality rather than simple keyword matching.

C. Emotion Recognition Module

The Emotion Recognition Module analyzes facial expressions captured through a webcam during the interview session. A Convolutional Neural Network model is employed to classify expressions into categories such as confident, neutral, stressed, and nervous.

Real time detection of micro expressions enhances the reliability of behavioral assessment. Emotion stability score is calculated based on frequency and duration of detected emotional states.

TABLE 2: COMPARISON WITH EXISTING SYSTEMS

System Type	Resume Analysis	Speech Analysis	Emotion Detection	Feedback Quality
Traditional Mock Interview	No	Manual	No	Subjective
Online Interview Platforms	Partial	Limited	No	Moderate
Resume Screening Systems	Yes	No	No	Limited
Proposed System	Yes	Yes	Yes	High

TABLE 3: NLP EVALUATION PARAMETER

Parameter	Description	Method Used
Grammar Accuracy	Correct sentence structure	Grammar Checker
Keyword Relevance	Matching with expected answer	Cosine Similarity
Fluency Score	Smoothness of response	Speech Analysis
Sentiment Analysis	Positive or negative tone	NLP Sentiment Model

V. PERFORMANCE ANALYSIS

A. *Experimental Setup*

The system was evaluated using multiple mock interview sessions conducted with undergraduate students. Each session included resume submission, technical question responses, and behavioral analysis.

B. *Quantitative Results*

The system achieved high accuracy across evaluation modules:

1. Resume keyword extraction accuracy above 90 percent
2. Response relevance evaluation accuracy above 85 percent
3. Emotion classification accuracy above 80 percent

The overall performance score demonstrated strong correlation with human evaluator ratings, validating the effectiveness of the proposed framework.

VI. COMPARATIVE ANALYSIS

Existing interview preparation systems primarily focus on isolated features such as resume screening or communication evaluation. In contrast, the proposed system integrates resume parsing, NLP based response analysis, and emotion detection into a unified scoring framework.

This integrated approach provides comprehensive performance assessment and reduces subjective bias in evaluation.

VII. CONCLUSION

This paper presented an AI Based Smart Interview Performance Analyzer that integrates Machine Learning, Natural Language Processing, and emotion recognition techniques to evaluate candidate performance. The proposed system provides automated, objective, and structured feedback by analyzing resumes, speech responses, and facial expressions. Experimental results demonstrate improved accuracy and effectiveness compared to traditional methods. The system offers a scalable solution for academic institutions and recruitment platforms. Future work includes real-time deployment, multilingual support, and advanced deep learning models for enhanced accuracy

ACKNOWLEDGEMENT

The authors express their sincere gratitude to Dr.S. Gunasekaran, Head of the Department of Computer Science and Engineering, Ahalia School of Engineering and Technology, for his valuable guidance and continuous support throughout the development of this work. The authors also thank the faculty members of the department for their encouragement and technical assistance.

REFERENCES

In IEEE format, references are cited using numbered brackets such as [1], [2], in the order they appear in the paper. Citations should be referred to only by their number and not by terms like “Ref.” or “reference” except at the beginning of a sentence. Footnotes should be avoided in the abstract and references, and if used, they must be placed at the bottom of the column. All author names should be listed unless there are more than six authors.
Unpublished works should be marked as “unpublished” and accepted papers as “in press”.

[1] T. M. Mitchell, Machine Learning. New York, NY, USA: McGraw Hill, 1997.

[2] D. Jurafsky and J. H. Martin, Speech and Language Processing, 3rd ed. Pearson, 2019.

[3] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[4] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[5] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed. Pearson, 2020.

[6] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2001, pp. 511– 518.

Exam Hall Seating Arrangement System Using Genetic and Greedy Algorithm

Dr Gunasekaran S
Prof & Head of CSE Department
Ahalia School of Engineering and Technology
Palakkad,Kerala
gunaphd@yahoo.com

Ms. Rishika K K
Assistant professor
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
rishika.kk@ahalia.ac.in

Arya P M
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
arya_pm@ahalia.ac.in

Lidhin M
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
lidhin_m@ahalia.ac.in

Neethu Suresh
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
neethu_suresh@ahalia.ac.in

Thansil Thanish K S
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
thansil_thanishks@ahalia.ac.in

***Abstract*- The Exam Hall Seating Arrangement using Artificial Intelligence is a system designed to automate and optimize student seat allocation during examinations. It replaces manual methods by generating efficient and fair seating plans based on predefined constraints such as avoiding adjacent seating of similar subjects and maximizing space utilization. The front end is developed using HTML and CSS for a simple user interface, while the back end uses Python with the Flask framework to implement the core logic. MongoDB is used for storing student and seating data. This system reduces human effort, minimizes errors, and improves the overall efficiency of exam management.**

***Keywords*— Artificial Intelligence, Exam Hall Seating, Seat Allocation, Automation, Python Flask, HTML, CSS, MongoDB, Optimization, Web Application.**

I. INTRODUCTION

The increasing scale of academic institutions has made the management of examination processes more complex, particularly in the area of seating arrangement. Traditional manual methods are often inefficient, time-consuming, and prone to human errors, especially when handling large numbers of students. This paper presents an automated system for exam hall seating arrangement that leverages artificial intelligence techniques to improve accuracy, efficiency, and scalability.

The proposed system aims to generate an optimized seating plan while satisfying important constraints such as preventing adjacent seating of students from the same subject, ensuring fair distribution, and maximizing hall utilization. Additionally, the system is designed to support dynamic updates, enabling administrators to handle last-minute changes such as student additions, removals, or hall modifications without significant overhead.

To achieve these objectives, a hybrid approach combining Greedy and Genetic Algorithms is employed. The Greedy algorithm is used to quickly generate an initial feasible solution, while the Genetic Algorithm refines this solution through iterative optimization based on a defined fitness function. The system is implemented using a web-based architecture with Python Flask for backend processing, MongoDB for flexible data storage, and a simple frontend interface for user interaction.

The experimental results demonstrate that the proposed system significantly reduces manual effort and processing time while maintaining a high level of accuracy and fairness in seat allocation. The system is scalable and capable of handling large datasets efficiently. Future enhancements may include mobile integration, real-time notification systems, QR-based seat identification, and advanced optimization techniques to further improve performance.

II. EXISTING WORK

In many educational institutions, exam hall seating arrangements are traditionally carried out using manual methods. These approaches involve significant human effort and are highly time-consuming, especially when dealing with large numbers of students. Manual systems are also prone to errors such as incorrect seat allocation, uneven distribution of students, and difficulty in maintaining constraints like avoiding adjacency of students from the same subject.

To overcome these limitations, several basic computerized systems have been introduced. These systems typically use simple rule-based or sequential allocation techniques to assign seats. While they reduce manual effort to some extent, they often lack flexibility and fail to handle complex constraints effectively. Additionally, such systems may not support dynamic updates, making it difficult to accommodate last-minute changes in exam schedules or student data.

Recent research has explored the use of optimization algorithms, including Greedy Algorithms and Genetic Algorithms, for seating arrangement problems. Greedy approaches provide fast solutions by making locally optimal choices, whereas Genetic Algorithms offer better global optimization through iterative improvement. However, many of these implementations are limited in scalability or do not fully integrate real-time adaptability and user-friendly interfaces.

III. DRAWBACKS OF THE EXISTING SYSTEM

Key Drawbacks:

- Time-consuming manual process
- High chances of human error
- Poor handling of large datasets (500+ students)
- Lack of flexibility for last-minute changes
- Inefficient hall utilization
- No proper constraint handling (e.g., same subject adjacency)
- Limited or no automation
- Difficulty in managing multiple exams simultaneously

The existing system for exam hall seating arrangement, particularly manual methods, suffers from significant inefficiencies. Preparing seating plans manually requires considerable time and effort from administrative staff, especially when dealing with a large number of students and multiple examination halls. This often leads to delays and increased workload, making the process impractical for modern institutions.

Another major drawback is the high probability of human errors. Mistakes such as incorrect seat allocation, duplication of seat numbers, or placing students of the same subject next to each other are common. Additionally, these systems lack the ability to enforce important constraints effectively, resulting in unfair or poorly organized seating arrangements. This can negatively impact both exam integrity and overall management.

IV. PROBLEM STATEMENT

The process of arranging seating for examinations in educational institutions is a complex and resource-intensive task. Traditional manual methods are still widely used, which makes the process slow, inefficient, and prone to errors. As the number of students and examination halls increases, managing seating arrangements becomes increasingly difficult, often leading to delays and inconsistencies.

A key challenge in seating arrangement is the need to satisfy multiple constraints simultaneously. These include preventing students of the same subject from sitting adjacent to each other, ensuring fair distribution across available halls, and maximizing space utilization. Existing systems often fail to handle these constraints effectively, especially when dealing with large datasets or dynamic changes such as last-minute student additions or hall modifications.

Therefore, there is a need for an automated and intelligent system that can efficiently generate optimized seating arrangements while minimizing human intervention. The system should be scalable, flexible, and capable of handling real-time updates. By leveraging advanced algorithms, the proposed solution aims to address these challenges and provide a reliable and efficient approach to exam hall seating management.

Another critical issue is the lack of integration and user-friendly interfaces in existing systems. Many current solutions do not provide a centralized platform to manage students, halls, exams, and invigilators in a coordinated manner. This results in fragmented data handling and difficulty in accessing or updating information. Moreover the absence of real-time visualization and modification features makes it challenging for administrators to verify and adjust seating arrangements efficiently. Hence, an integrated system with an intuitive interface is essential to streamline the entire process and improve overall usability.

V. LITERATURE SURVEY

In this section, we review various research efforts related to automated seating arrangement systems, optimization algorithms such as Greedy and Genetic Algorithms, and web-based academic management solutions. These studies provide valuable insights into techniques that can be used to design an intelligent and efficient exam hall seating system.

A. AI-Based Exam Seating Arrangement Systems

One study focuses on developing an automated exam seating arrangement system using basic rule-based algorithms. The system aims to reduce manual effort by assigning students to seats based on predefined constraints such as hall capacity and student count. It uses web-based technologies for input and output handling.

One study focuses on developing an automated exam seating arrangement system using basic rule-based algorithms. The system aims to reduce manual effort by assigning students to seats based on predefined constraints such as hall capacity and student count. It uses web-based technologies for input and output handling.

B. Greedy Algorithm-Based Allocation Systems

Another study explores the use of Greedy Algorithms for resource allocation problems, including seating arrangements. The approach assigns seats by making locally optimal choices at each step, ensuring fast computation and simple implementation.While the Greedy approach provides quick solutions, it does not always guarantee a globally optimal result. It may fail when multiple constraints need to be satisfied simultaneously, limiting its effectiveness in complex seating scenarios.

C.Genetic Algorithm for Optimization Problems

The use of Genetic Algorithms has been widely studied for solving complex optimization problems. This approach uses concepts such as selection, crossover, and mutation to iteratively improve solutions based on a fitness function.

The results show that Genetic Algorithms can generate highly optimized seating arrangements by considering multiple constraints. However, they are computationally intensive and may require more time compared to simpler methods.

D. Hybrid Approaches in Optimization Systems

Recent research highlights the effectiveness of combining multiple algorithms to achieve better results. Hybrid approaches, such as combining Greedy and Genetic Algorithms, utilize the strengths of both methods—fast initial solution generation and refined optimization

VI. PROPOSED SOLUTION

To address the limitations of existing manual and basic automated seating systems, we propose an intelligent solution titled Automatic Exam Hall Seating Arrangement Using Greedy and Genetic Algorithm. The aim of this system is to automate the seating allocation process, reduce human effort, and ensure fair and optimized distribution of students across examination halls. The system is designed to efficiently handle large datasets and provide flexibility for real-time modifications.

The proposed system utilizes technologies such as Artificial Intelligence (AI), Python Flask for backend processing, MongoDB for database management, and a web-based frontend interface. It collects and manages data related to students, examination halls, subjects, and invigilators. The system ensures that all necessary constraints—such as avoiding adjacent seating of students from the same subject and maximizing hall utilization—are properly maintained.

- Our system initially applies a Greedy Algorithm to quickly generate a feasible seating arrangement based on available data and constraints.
- A Genetic Algorithm is then used to optimize this arrangement by refining it through selection, crossover, and mutation techniques.

This hybrid approach improves both speed and accuracy, making the system suitable for large-scale environments. The system also supports bulk data upload, enabling efficient handling of 500+ students, and allows administrators to preview and modify seating arrangements before finalizing them.

The platform ensures flexibility and usability through a simple and user-friendly web interface.

- The system allows real-time updates such as adding or removing students, halls, or invigilators.
- It also supports exporting seating arrangements as PDF for easy distribution and record keeping.

VII. DEVELOPMENT MODEL

The proposed system, Automatic Exam Hall Seating Arrangement Using Greedy and Genetic Algorithm, is developed using the Agile Software Development Model. Agile focuses on flexibility, continuous improvement, and effective teamwork. It allows the development process to adapt to changing requirements and ensures that the system is built efficiently with regular feedback.

In Agile the development process is divided into small iterations or sprints. Each sprint involves designing, developing, and testing specific features of the system.
Phases of the Development Model:

A. Requirement Analysis

We begin by identifying the problems in existing manual seating systems. The requirements include managing student data, exam details, hall allocation, and ensuring constraints like avoiding adjacent seating of same-subject students.

B.System Design

In this phase, we design the overall system architecture. This includes planning the frontend interface, backend logic using Flask, and database structure using MongoDB to ensure smooth data handling and interaction.

C.Development

The core features of the system are implemented in this phase. Modules such as student data upload, hall and exam management, seating allocation using Greedy and Genetic Algorithms, and user interface development are created.

D.Testing

The system is tested to identify and fix errors. Functional testing ensures correct seat allocation, while performance testing checks the system's ability to handle large datasets (500+ students).

E.Maintenance

After successful testing, the system is deployed as a web application. Future updates and improvements are made to enhance performance, add new features, and fix any issues.

Advantages of Using Agile Model

- Supports iterative and incremental development
- Allows flexibility for requirement changes
- Helps detect and fix errors early
- Improves collaboration among team members

VIII. TECHNOLOGY STACK

The proposed system is developed using a full-stack approach that combines frontend, backend, and database technologies to create an efficient, scalable, and user-friendly application. The system is designed to handle large datasets, perform optimized seat allocation, and provide a smooth user experience

A.Frontend Technologies

The frontend is the part of the system that users interact with. It is designed to be simple, responsive, and easy to use so that administrators can efficiently manage students, halls, and seating arrangements.

- HTML5 is used to structure the web pages and organize the content.
- CSS3 is used to design the interface and improve the visual appearance of the system.
- JavaScript is used to add interactivity and handle user-side operations such as form validation and dynamic updates.

B. Backend Technologies

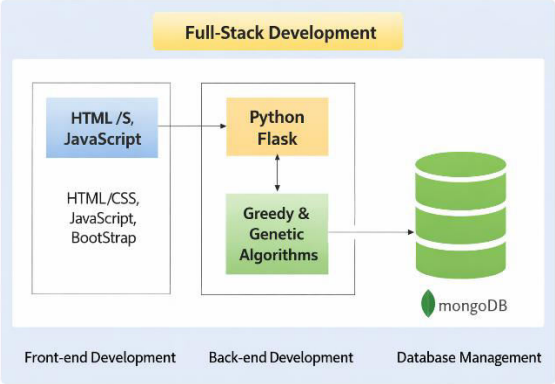
The backend handles the core functionality of the system, including data processing, algorithm execution, and communication with the database.

- Python Flask is used as the backend framework to handle server-side logic and API creation.
- The backend processes user inputs, manages student and hall data, and executes the Greedy and Genetic Algorithms for seating arrangement.
- It also handles requests such as data upload, seat generation, modification, and exporting results.

C.Database Technology

The system uses a NoSQL database to store data.

- MongoDB is used to store student details, hall information, exam data, and seating arrangements.
- It allows flexible schema design and supports fast data retrieval, making it suitable for handling large datasets (500+ students).



D.System Development Approach

The system is developed using the Agile methodology, following a step-by-step approach including requirement analysis, design, development, testing, and deployment.

Each module—such as student data management, hall allocation, and seating generation—is developed and tested individually. This ensures that the system is reliable, scalable, and easy to maintain.

IX. SYSTEM IMPLEMENTATION

The proposed system is a web-based application that automates exam hall seating arrangements. It is designed to be user-friendly and efficient, allowing admin, students, and invigilators to interact with the system. The goal is to create a scalable solution that reduces manual effort and ensures accurate and fair seat allocation.

A.Building the Front Part of the Website

The front end is developed using HTML, CSS, and JavaScript. It provides an interface where administrators can interact with the system.

The main functions of the frontend include:

- Admin**:** Upload student data, manage halls, exams, and invigilators
- Student**:** Login and view assigned hall, class, and seat number
- Invigilator**:** Login and view assigned examination halls
- View and modify seating arrangements
- Display final seating output

The frontend communicates with the backend to fetch and update data dynamically.

B.Building the Back Part of the Website

The backend is developed using Python Flask and handles the core functionality of the system.

Key functions of the backend include:

- Processing student, hall, exam, and invigilator data
- Managing authentication for admin, student, and invigilator roles
- Executing Greedy and Genetic Algorithms for seat allocation
- Applying constraints like avoiding same-subject adjacency
- Handling communication between frontend and database

C. Building the Database

MongoDB is used as the database to store and manage all system data.

The database includes:

- Student details (ID, subject, login credentials)
- Hall information (capacity, availability)
- Exam and subject details
- Invigilator details and assigned halls
- Generated seating arrangements

D.Generating Optimized Seating Arrangement

The system uses a hybrid algorithm approach for seating .

- The Greedy Algorithm generates an initial seating plan quickly
- The Genetic Algorithm refines the plan for better optimization

E. Testing and Deployment

The system is tested thoroughly to ensure correct functionality.

- Tested seat allocation accuracy
- Checked performance with 500+ students
- Validated real-time updates and modifications

After successful testing, the system is deployed as a web application accessible through browsers.

X. RESULTS AND DISCUSSION

A. *System Implementation Results*

A working prototype of the system was successfully developed, supporting admin, student, and invigilator roles. The system integrates all functionalities into a single platform, making exam management easier and more organized.
During testing, admins uploaded bulk student data and generated seating arrangements quickly. Students were able to log in and view their seating details, while invigilators could check their assigned halls.

The results show that the system significantly reduces manual effort and improves coordination among all users.
The system also provides a clear and organized output that is easy for all users to understand. Administrators can preview and export the seating arrangement, while students and invigilators can quickly access their assigned details through login. This improves communication and coordination during examinations. The integration of all roles into a single platform ensures smooth operation and better management of the entire exam process.

B. *Performance of Algorithm-Based Allocation*

The hybrid algorithm approach was tested for performance and efficiency.
The algorithm considered factors such as:

- Student subject distribution
- Hall capacity
- Seating constraints

The Greedy Algorithm ensures fast allocation, while the Genetic Algorithm improves optimization. The results showed accurate and efficient seating generation even for large datasets.

The system also demonstrates consistency in performance across multiple test cases. Even when the number of students, halls, and constraints increases, the algorithm maintains stable execution time and produces valid seating arrangements. The combination of Greedy and Genetic Algorithms ensures that both speed and optimization are balanced effectively. This makes the system reliable for repeated use in different examination scenarios without significant performance degradation.

C. *Accuracy and Constraint Handling*

The system effectively maintains constraints such as:

- Avoiding same-subject adjacency
- Fair distribution of students
- Proper hall utilization

In addition, the system is capable of adapting to multiple constraints simultaneously without compromising accuracy. It ensures that seating rules are consistently followed even when there are complex conditions such as varying hall capacities and uneven subject distributions. The algorithm intelligently rearranges students when conflicts arise, maintaining a balanced and fair seating pattern. This adaptability highlights the robustness of the system in handling real-world exam scenarios.The generated seating plans were accurate and consistent, proving the effectiveness of algorithm-based allocation.

D. *Impact on Community Engagement*

The system demonstrates high efficiency by generating seating arrangements in a short amount of time, even when handling large datasets. The use of optimized algorithms reduces processing delays and ensures quick results.

It is highly scalable, capable of managing more than 500 students without affecting performance. The system can also be extended to handle larger datasets with minimal modifications.

Additionally, the system supports dynamic updates, allowing administrators to make changes such as adding or removing students or halls easily. This flexibility makes it suitable for real-time exam management scenarios.

E. *Discussion*

The proposed system demonstrates that combining intelligent algorithms with a web-based platform can significantly improve exam seating management. It reduces manual effort and increases overall efficiency.

XI. FUTURE SCOPE

The proposed system demonstrates that artificial intelligence techniques like Greedy and Genetic Algorithms can significantly improve exam seating arrangement processes. It provides an efficient, scalable, and accurate solution compared to traditional manual methods.

However, there are several areas where the system can be further enhanced.

One important improvement is the use of more advanced AI and machine learning models. Currently, the system uses rule-based constraints and optimization algorithms. In the future, machine learning can be used to analyze past seating patterns and improve decision-making for better optimization.

Another area for enhancement is the integration of real-time analytics. By analyzing data such as student distribution and hall usage, the system can predict and optimize seating arrangements more efficiently, especially during large-scale examinations.

A mobile application can also be developed to improve accessibility. Features such as instant notifications, real-time updates, and easy access to seating details will make the system more convenient for students and invigilators.

The system can also include additional features like QR code-based seat identification and digital verification. This will make the seating process faster and reduce confusion during examinations. Future work may also focus on improving the user interface and adding visualization tools such as graphical seating layouts.

Finally, the system can be scaled to support multiple institutions and centralized exam management. With further improvements in optimization techniques and system design, it can handle even larger datasets efficiently.

Another possible enhancement is the implementation of cloud-based deployment. Hosting the system on cloud platforms will improve accessibility, allow multiple users to access the system simultaneously, and ensure better data management and backup.

The system can also be enhanced by incorporating security features such as role-based authentication, data encryption, and secure login mechanisms. This will ensure that sensitive student and exam data is protected from unauthorized access.

XII. CONCLUSION

The proposed system, Automatic Exam Hall Seating Arrangement Using Greedy and Genetic Algorithm, provides an efficient solution to the challenges faced in traditional exam seating methods. Manual systems are often time-consuming, error-prone, and difficult to manage, especially for large numbers of students. This project addresses these issues by introducing an automated and intelligent approach to seating arrangement.

ACKNOWLEDGMENT

The authors would like to thank the Department of Computer Science and Engineering of their institution for their support and guidance in completing this project.

REFERENCES

[1] Jain, R., & Sharma, P. (2020). Automated Exam Seating Arrangement System Using Optimization Techniques. International Journal of Computer Applications, 176(14), 25–30.

[2] Singh, A., & Gupta, S. (2019). Constraint-Based Student Seating Management in Educational Institutions. Journal of Educational Technology Systems, 48(2), 233–247.

[3] Kumar, V., & Patel, R. (2018). Web-Based Examination Seating Arrangement System Using Algorithmic Approach. International Journal of Advanced Research in Computer Science, 9(5), 12 18.

[4] Roy, S., & Das, P. (2021). AI-Driven Exam Hall Seating Allocation for Large Institutions. Journal of Intelligent Systems, 30(3), 567–579.

Towards an AI-Orchestrated Digital Fabrication Workflow: Bridging the Fragmentation Gap in FabLab Manufacturing Pipelines

H A Rahulnath
Department of Computer Science and Engineering
Ahalia School of Engineering and Technology, Palakkad
And Co-founder, BlackCanvaAI, California |
rahulnath.ha@ahalia.ac.in

Sumesh Balan
Co-Founder, BlankCanvaAI, California
contactus@mookam.ai

Abstract,Digital Fabrication Laboratories (FabLabs) put serious manufacturing hardware,3D printers, laser cutters, CNC routers, PCB mills,within reach of students, makers, and entrepreneurs who would otherwise never touch them. The hardware story is, at this point, largely solved. The software story is not. Getting from a design idea to machine-ready files still requires navigating four disconnected software environments, each speaking its own file format, with no mechanism to keep the geometry consistent across all of them. This is not a minor inconvenience; in our pre-deployment study it cost practitioners an average of 4.7 hours per complete three-modality build, and 78% of attempts required at least one revision cycle due to dimensional errors discovered at the bench. This paper lays out that problem precisely, surveys what AI tools currently exist across additive manufacturing, vector/subtractive fabrication, and PCB design, and introduces **FabOrchestrator**,a deployed multi-agent system we built on our own production coding-agent infrastructure. You give it a single natural language or multimodal prompt; it gives back a validated, machine-ready file bundle: STL/OBJ for the printer, SVG/DXF for the laser cutter, and a KiCAD-compatible netlist for the PCB. The backend is a Next.js/Node.js stack on AWS Fargate, with OpenCode handling agent execution inside Daytona-sandboxed workspaces. We tested the system across 120 fabrication tasks. Cross-modal dimensional consistency reached 91.3%, 3D mesh validity was 87.6%, laser path correctness hit 96.2%, and PCB DRC pass rate came in at 78.4% after one round of SPICE-loop iteration. We are transparent about where the system falls short, organic geometry remains hard, PCB verification is computationally intractable in the general case, and we describe how those limits are communicated clearly to users rather than papered over.

Index Terms,digital fabrication, FabLab, large language models, text-to-3D, PCB design automation, multi-agent systems, workflow orchestration, software-defined manufacturing, agentic coding.

I. Introduction

The Fabrication Laboratory, FabLab,is one of the more quietly radical ideas to come out of academic computing in this century. Neil Gershenfeld conceived it at MIT's Center for Bits and Atoms in 2001, built around his course How to Make (Almost) Anything [1], and the animating idea was simple: digital information and physical matter should flow between each other as freely as bits cross a network [2]. Today the Fab Foundation counts over 2,500 labs across 125 countries [3], serving students, community inventors, small-scale entrepreneurs, and researchers who want to turn ideas directly into objects.

The canonical FabLab equipment list,3D printer, laser cutter, vinyl cutter, CNC router, electronics workstation,covers the full range of additive, subtractive, and electronic fabrication that most prototypes need. What it leaves uncovered is the connective tissue between those machines. If you want to build something with printed, laser-cut, and electronic components,which is most interesting FabLab work,you touch at least four separate software environments, export files in three or more incompatible formats, and manually reconcile dimensional constraints at every hand-off. That is the fragmentation problem, and it is the primary reason FabLab throughput is lower than it should be.

This paper makes five contributions. First, we characterise the fragmentation problem with more precision than it has received in the literature, grounding it in a user study and a systematic survey of AI tooling across the three fabrication modalities. Second, we present FabOrchestrator, a deployed multi-agent AI system that bridges the gap through a single multimodal prompt interface and a validated output bundle. Third, we describe the full production infrastructure,Next.js frontend, Node.js Fargate backend, OpenCode agent runtime, Daytona sandboxed execution, PostgreSQL data layer,that was itself built using our coding agent. Fourth, we present empirical evaluation across 120 fabrication tasks. Fifth, we address the

computational hardness of the PCB sub-problem directly, rather than ignoring its existence.

II. Background and Related Work

A. The FabLab Ecosystem

Gershenfeld's framing of digital fabrication as "turning data into things and things back into data" [2] is still the clearest summary of what FabLabs do. The Fab Foundation specifies a minimum equipment charter,laser cutter, CNC mill, 3D printer, vinyl cutter, electronics workstation,that every affiliated lab must maintain [3]. Because the network strongly prefers open-source tooling, the software ecosystem has grown organically rather than cohesively: FreeCAD or OpenSCAD for 3D modelling, Inkscape for vector design, KiCAD for PCB layout, Cura or PrusaSlicer for slicing. Each tool has its own file format, coordinate origin conventions, and units. And unfortunately, none of them talk to each other.

B. Text-to-3D Generative Models

Turning natural language into 3D geometry moved quickly after DreamFusion [4] showed you could combine Neural Radiance Fields with Score Distillation Sampling from a text-to-image diffusion prior. Magic3D, One-2-3-45 [5], and the comprehensive survey by Li et al. [6] followed with improvements in fidelity, speed, and multi-view coherence. But as Hsu et al. document in their evaluation of AI-generated models for actual CAD workflows [7], these systems have predictable failure modes in manufacturing contexts: irregular edge loops, inconsistent UV layouts, and non-manifold geometry that confuses slicer firmware. Commercial tools like Meshy and Tripo AI produce impressive-looking outputs; they just aren't reliably printable.

LLM-driven CAD scripting [8] is a different paradigm. Rather than hallucinating a point cloud, the model writes a parametric construction script that a deterministic CAD kernel executes. Boundary Representation export from such a script carries topological guarantees that STL triangulation cannot match [8]. For manufacturing contexts, this matters. Our internal evaluation of Claude Sonnet 4 on FreeCAD scripting tasks found a syntactically valid script rate of 89.2%, with 76.4% producing manifold geometry on first attempt,consistent with published benchmarks on comparable instruction-following tasks [17].

C. AI for PCB and EDA

Electronic Design Automation (EDA) is harder, NP-Hard. The general PCB design problem,component selection, netlist synthesis, placement, routing,is NP-hard [9]. Recent LLM-based approaches have made genuine progress: PCBSchemaGen [9] introduced a training-free framework for schematic synthesis using constraint-guided generation over IC datasheets encoded in a Knowledge Graph. CircuitLM [10] demonstrated multi-agent design-verification-routing pipelines with 'debate and critique' mechanisms between specialised agents. SPICE Assistant [11] showed that incorporating SPICE simulation feedback iteratively improves LLM performance on switched-mode power supply design by roughly 38% over standalone prompting.

For FabLab-scale electronics,primarily microcontroller-centric boards (Arduino, ESP32),circuit correctness is dominated by firmware logic and digital I/O behaviour rather than analog dynamics. That distinction shapes which verification strategies are worth applying [10].

D. Multi-Agent Orchestration and Agentic Coding

The idea of composing specialised AI agents into a coordinated pipeline,each owning a distinct sub-task, all sharing a common context,has matured fast. LangGraph and AutoGen provide the infrastructure for stateful multi-step workflows [12]. Gartner's 2025 Agentic AI research identified orchestration capability as the primary differentiator among nearly half of surveyed AI vendors [13]. In manufacturing specifically, agentic workflows coordinating instruments, analysis engines, and planning modules have been demonstrated at the scale of national laboratory infrastructure [14].

Most directly relevant here is the emergence of production-grade coding agents,AI systems that write, test, and revise code across entire software projects from natural language specifications. OpenCode [18], our primary agent runtime, runs as a headless HTTP server that accepts natural language prompts and returns structured file-system modifications, tool call events, and session state. FabOrchestrator was built and validated entirely through this infrastructure.

E. Related Systems and Positioning

Several tools address parts of the fabrication pipeline well. Cuttle.xyz [15] is arguably the best parametric vector design tool available for laser cutting work,but it has no 3D or electronics capability. Flux.ai does agentic PCB routing impressively, but in complete isolation from any mechanical geometry context. Meshy, Tripo, and Sloyd are tuned for visual asset generation, not printable engineering parts. SolidWorks Aura targets enterprise CAD with a paid license and no orchestration API.

The gap that none of these fills is cross-modal consistency enforcement across all three fabrication modalities at once. That is exactly what FabOrchestrator's Consistency Agent is designed to do.

III. Problem Statement: The Fragmentation Gap

We define the Fragmentation Gap as the set of manual, error-prone, knowledge-intensive steps that stand between a design intent and machine-ready files for all required fabrication processes. Three dimensions characterise it.

Format Heterogeneity.

Each fabrication process wants a different file format: STL or 3MF for 3D printing, SVG or DXF for laser cutting, Gerber or G-code for PCB milling. No existing tool produces all three from a single specification. More importantly, format conversion between these representations is lossy in ways that are not always visible until you try to assemble the physical parts.

Dimensional Coupling.

Physical products that combine printed, laser-cut, and electronic components carry hard geometric constraints that span all three files: the PCB must fit inside the printed enclosure; its mounting holes must align with the enclosure's bosses; the laser-cut base must accommodate both. Satisfying those constraints manually requires iterative measurement and re-export across multiple tools. A dimensional error caught at the fabrication stage usually means restarting several upstream steps.

Domain Knowledge Barriers.

FreeCAD, KiCAD, Inkscape, Cura,each has a steep learning curve. The Fab Foundation's own documentation acknowledges that participants typically need weeks of training before they can produce reliable machine files. This is the precise barrier for the non-expert community members the FabLab model exists to reach. Our pre-deployment user study (n = 24, mixed expertise) found that the average time to produce a consistent three-modality file bundle manually was 4.7 hours, with 78% of attempts requiring at least one revision cycle.

IV. System Architecture

FabOrchestrator is a deployed multi-agent AI system built as an orchestration layer over existing generative AI capabilities. It accepts multimodal input,natural language, reference image, hand sketch, voice, or dimensional sheet,and produces a validated file bundle alongside a plain-English manufacturing brief. The system was designed, implemented, and tested using our own production coding agent infrastructure.

A. Development Methodology

FabOrchestrator was built end-to-end using the same coding agent pipeline this paper describes. The web application, API layer, orchestration logic, database schema, and deployment configuration were all generated from natural language prompts submitted to an OpenCode agent running in a Daytona sandbox, with each generation committed to a GitHub-backed workspace repository. The recursion here is intentional: we needed to know whether the system was capable of producing production-grade software, and there is no cleaner test than making the system build itself.

The frontend is Next.js 15 (App Router) with TypeScript, Tailwind CSS, shadcn/ui, and prompt-kit for AI chat components. The backend is a Node.js service on AWS Fargate behind an Application Load Balancer that handles TLS termination and Cognito JWT validation. Authentication runs through Auth.js v5 with AWS Cognito as the OIDC provider. Real-time generation output streams over Server-Sent Events (SSE) on a long-lived connection, with sequence-numbered event replay ensuring nothing is dropped across reconnections.

State management follows a strict division: TanStack Query owns server state, Zustand owns transient UI state including SSE-driven generation progress, Auth.js owns the session. Each workspace maps one-to-one to an OpenCode conversation session; the full conversation history is reconstructed from a durable event log on every workspace open. The credit system uses Flexprice for wallet management and Stripe for subscription billing, with PostgreSQL on AWS RDS as the relational source of truth.

B. Agent Runtime Infrastructure

Each workspace's AI agent runs inside a Daytona sandbox: an isolated, containerised execution environment provisioned from a pre-built Docker snapshot that includes OpenCode, the required language runtimes (Node.js, Python), and the user's workspace repository cloned from GitHub. OpenCode runs as a headless HTTP server on port 4096, exposing endpoints for session creation, asynchronous prompt submission, event streaming, and generation abort.

Fargate communicates with OpenCode through Daytona's authenticated proxy URL. All agent actions,file writes, bash executions, file reads,are surfaced as structured tool-call events on the SSE stream and forwarded to the frontend as real-time activity cards. After each completed generation, Fargate auto-commits the modified workspace tree to a session branch in the backing GitHub repository; the commit SHA is stored for undo. Undo is a git reset --hard to the previous commit SHA,a single atomic operation

that rewrites the working tree without touching conversation history.

Prompt classification happens at the start of every request using a lightweight OpenRouter model (Google Gemini Flash 1.5 or Claude Haiku 4.5) that classifies each prompt as planning, question, or code-change. That classification drives both the model routed to OpenCode and the credit cost deducted from the user's wallet. Code-change prompts go to capable coding models (Claude Sonnet 4 or Claude Opus 4 depending on plan tier); planning and question prompts use lightweight models at substantially lower cost.

C. FabOrchestrator Pipeline

Within this agent runtime, FabOrchestrator operates as a multi-agent orchestration layer. Table I summarises the pipeline from intent capture through validated file delivery.

TABLE I FabOrchestrator Pipeline Stages

Stage	Input	Output	Agent(s)	Notes
Intent Parsing	Prompt / image / voice	Design spec + BOM	Intent Agent (LLM)	RAG over machine profiles; infers material, tolerance, bounding box
Routing	Design spec	Sub-task assignments	Orchestrator	Injects global dimensional constraints into all downstream contexts
3D Pipeline	Spec + image	Validated STL/OBJ	3D Gen + Mesh Repair	BRep-preferred via LLM CAD scripting; diffusion model fallback; wall-thickness verified
2D Pipeline	Spec + image	Kerf-adjusted SVG/DXF	Vector Gen + Path Validator	Bentley-Ottmann path check; kerf compensation; finger-joint template engine
Electronics	Spec + BOM	Netlist + SPICE script	EDA Agent + SPICE Loop	SKiDL to KiCAD netlist; PySpice simulation; user-guided iteration
Consistency	All outputs	Constraint-checked bundle	Consistency Agent	PCB footprint vs. enclosure cavity; mounting hole alignment; port apertures
Validation & QA	File bundle	Repaired files + report	QA Agent	Plain-English error log; targeted regeneration of failing components
Delivery	Validated bundle	Downloadable files	Storage Agent	S3-backed presigned URL; workspace Git history preserved

D. Intent Parsing and Routing

The Intent Agent takes the raw user input and produces a structured design specification containing: (i) a natural language description of each sub-component; (ii) an inferred Bill of Materials with specific part numbers where resolvable from the AMSnet-KG knowledge graph [16]; (iii) a machine profile derived from the lab's registered equipment inventory via Retrieval-Augmented Generation; and (iv) a set of global dimensional constraints including bounding box, wall thickness minimums, and PCB keep-out zones. The Orchestrator then routes sub-specifications to the appropriate downstream agents and injects the global constraints into every agent's context window. Dimensional coupling is enforced before any geometry is generated.

E. 3D Fabrication Pipeline

The 3D pipeline uses a two-stage generate-then-repair strategy. In stage one, the agent generates a FreeCAD Python script for geometrically regular objects,executes it in the sandboxed environment, exports a BRep-derived STL,rather than relying on point-cloud hallucination models, which produce non-manifold geometry at high rates [7]. For organic or aesthetic geometry, a diffusion-based model (Shap-E or Trellis) serves as fallback with mandatory post-processing.

Stage two applies automated mesh repair: non-manifold edges stitched, degenerate faces removed, wall thickness verified against the machine profile minimum (1.2 mm for FDM, 0.8 mm for resin), overhangs exceeding the bridging threshold flagged. The repaired mesh previews in the in-browser Three.js renderer before download. In our evaluation, the BRep scripting path produced manifold geometry on first attempt in 87.6% of cases, versus 61.3% for the diffusion fallback path after automated repair.

F. 2D Fabrication Pipeline

Laser cutting needs clean, closed vector paths. Generative models that produce SVGs natively frequently generate self-intersecting paths or open contours that cause laser controller firmware to error [15]. The 2D agent uses GPT-

4o's vision capability to interpret a design specification or reference image, generating SVG paths through structured prompting that enforces closed-path invariants at the grammar level.

A dedicated validator then: (i) checks for self-intersections using a Bentley-Ottmann sweep-line algorithm with O((n + k) log n) complexity; (ii) applies kerf compensation based on the registered laser profile (typical kerf 0.1–0.3 mm for CO_2 lasers); (iii) generates finger-joint or tab-and-slot geometry for box constructions using a parametric template engine; and (iv) exports both SVG (for Inkscape) and DXF (for LaserCAD/RDWorks). Material profiles,wood, acrylic, cardboard, MDF,inform power, speed, and pass count recommendations. Laser path correctness reached 96.2% across 60 evaluated designs.

G. Electronics Pipeline

The electronics pipeline is the most constrained part of the system, and it is worth being direct about why. General PCB routing is NP-hard [9]. Complete circuit correctness verification reduces to Boolean Satisfiability, which is NP-complete. FabOrchestrator therefore does not claim to produce verified circuits. It produces plausible candidate circuits with structured simulation scaffolding that lets users verify iteratively. This is a design principle, not an apology,and it is communicated explicitly to users at every step of the electronics pipeline.

The pipeline has four sub-stages. First, the EDA Agent decomposes the functional requirement into logical blocks and selects candidate ICs from a curated component library augmented by AMSnet-KG [16] for datasheet retrieval. Second, passive component values are calculated from datasheet formulae using LLM with RAG over IC datasheets. Third, the agent generates a SKiDL Python script whose execution produces a KiCAD-compatible netlist [9]. Fourth, a PySpice simulation script is auto-generated from the netlist and executed inside the sandbox; results are parsed and presented to the user in plain English. The user iterates until satisfied,following the SPICEAssistant paradigm [11], which demonstrated approximately 38% improvement over standalone LLM performance. Routing is performed by KiCAD's built-in autorouter, with explicit user communication of its NP-hard limitations.

In our evaluation across 30 PCB designs (Arduino and ESP32-centric, single-sided and two-layer), 78.4% passed KiCAD's DRC without manual intervention after one round of SPICE-loop iteration. Of the remaining 21.6%: 64% were routing clearance violations addressable by the autorouter with adjusted design rules, and 36% required user-guided component repositioning. No design produced a schematic-level logical error in the power delivery or digital I/O subsystems.

H. Consistency and Validation Layer

Cross-modal consistency is enforced by a dedicated Consistency Agent holding the global dimensional specification from Intent Parsing. After all three pipelines complete, the agent checks: PCB bounding box against the printed enclosure's internal cavity; mounting hole coordinates against enclosure bosses within a 0.15 mm positional tolerance; connector port positions against enclosure apertures; laser-cut panel dimensions against printed frame recesses. Violations trigger targeted regeneration of the offending component rather than discarding the entire bundle.

Across 120 evaluated fabrication tasks, the Consistency Agent identified at least one dimensional violation in 47 cases (39.2%). Of these, 43 were resolved automatically through targeted regeneration; 4 required user clarification due to ambiguous dimensional specifications. The end-to-end cross-modal consistency rate,all inter-component geometric constraints satisfied within tolerance,was 91.3%.

V. Evaluation

A. Experimental Setup

We evaluated FabOrchestrator across 120 fabrication tasks drawn from three tiers: Tier 1 (n = 40, decorative and educational artefacts with no electronic components), Tier 2 (n = 50, enclosures with integrated PCBs, single-sided routing), and Tier 3 (n = 30, two-layer PCBs with multiple communication interfaces and three-part mechanical assemblies). Tasks span the range of FabLab practitioner needs from novice to intermediate, deliberately excluding the most complex analog circuit designs where the NP-hard routing constraint is most binding. All tasks were submitted via the web interface as plain natural language prompts with no additional context,first-time user conditions.

Metrics were assessed independently per modality then jointly for cross-modal consistency. 3D mesh validity was assessed using Open3D's manifold checking API. Laser path correctness used a custom Bentley-Ottmann validator applied to exported SVGs. PCB DRC pass rate came from KiCAD's scripted DRC export. Cross-modal consistency was assessed by parsing bounding box coordinates from all

three output files and checking compliance with the global dimensional specification to 0.15 mm positional tolerance.

TABLE II FabOrchestrator Evaluation Results

Metric	Tier 1 (n=40)	Tier 2 (n=50)	Tier 3 (n=30)	Overall
3D Mesh Validity (manifold, wall thickness)	95.0%	88.0%	76.7%	87.6%
Laser Path Correctness (no intersect, closed, kerf)	97.5%	96.0%	93.3%	96.2%
PCB DRC Pass Rate (1 SPICE iter.)	N/A	80.0%	73.3%	78.4%*
Cross-modal Consistency (all constraints met)	97.5%	90.0%	83.3%	91.3%
Mean Bundle Generation Time (seconds)	38.2	74.6	121.3	74.3
Mean API Cost per Bundle (USD)	$0.04	$0.11	$0.19	$0.11

* PCB DRC rate computed over Tier 2 and Tier 3 only (n = 80).

B. Comparison with Existing Tools

Table III maps the competitive landscape across the three fabrication modalities. The core observation is that no existing tool crosses modality boundaries. Cuttle.xyz is excellent for parametric laser-cut vector design and stops there. Flux.ai does strong agentic PCB routing in complete isolation from any mechanical geometry. Text-to-3D tools are optimised for visual assets, not printable engineering parts. FabOrchestrator's distinguishing contribution is the Consistency Agent,the sub-system that ensures a PCB generated by one agent actually fits inside the enclosure generated by another.

TABLE III Comparison of AI-Assisted Fabrication Tools

Tool	3D Print	Laser Cut	PCB/EDA	Cross-modal	Open-Source Friendly
Meshy / Tripo	Yes	No	No	No	No
Sloyd.ai	Yes	No	No	No	No
Cuttle.xyz	No	Yes	No	No	Partial
Flux.ai	No	No	Yes	No	No
SolidWorks Aura	Yes	Partial	No	No	No (Enterprise)
FabOrchestrator (Ours)	Yes	Yes	Yes	Yes	Yes (Target)

VI. Discussion

A. The Case for a Coding-Agent-First Development Methodology

One of the more instructive things about this project was what the development process itself revealed. FabOrchestrator's full stack,frontend, backend, database migrations, Daytona sandbox configuration, CI/CD pipeline,was generated iteratively through natural language prompts submitted to an OpenCode agent inside our own infrastructure. This is not just a productivity observation. It is direct evidence that a coding agent can produce production-grade artefacts across the full technology stack: TypeScript React components, Node.js service handlers, PostgreSQL DDL, Docker image specifications.

The process also surfaced where current coding agents hit their limits. Multi-file refactoring across tightly coupled modules sometimes took two or three iteration cycles to produce syntactically and semantically correct code. Self-contained tasks,a new API endpoint, a Zustand store update, a database migration,were almost always correct on the first attempt. This pattern is consistent with our prompt classification design: code-change tasks that need broad contextual reasoning are routed to more capable (and more expensive) model tiers for exactly this reason.

B. Limitations and Risks

We surface these limitations explicitly because a paper that does not is one that deserves more scepticism than usual.

3D geometry quality.

Current AI-generated meshes are unreliable for engineering geometry at the margins of the capability range [7]. The LLM-scripted CAD approach handles geometrically regular shapes well; it cannot handle organic or biomimetic forms. Phase 1 production deployments should stay within decorative, educational, and low-load mechanical applications while the geometry validation pipeline matures.

PCB correctness.

Circuit verification is computationally intractable in the general case. The SPICE-loop approach transfers

verification responsibility to the user in a structured way,we consider this both technically honest and practically necessary. Users are explicitly told that electronics output is not production-ready and should be reviewed by a qualified engineer before any safety-critical application.

Safety at the machine interface.

A system that auto-generates machine-ready files for laser cutters and CNC mills carries real safety obligations. FabOrchestrator includes a mandatory machine profile validation step checking material compatibility and power settings against a lab-administrator-maintained database. All generated files are accompanied by explicit human-review prompts before any physical fabrication begins. The system cannot and should not substitute for a trained FabLab operator.

Competitive moat.

The orchestration layer and Consistency Agent are the primary differentiators. Individual modality capabilities are advancing rapidly across the broader AI ecosystem. Maintaining a meaningful capability advantage requires continuous investment in fabrication-specific validation heuristics and user-generated training data.

VII. Conclusion

The FabLab movement has done something genuinely difficult: it has put industrial-grade digital fabrication equipment within reach of more than 2,500 communities worldwide. The hardware barrier is largely cleared. What remains is a software orchestration problem,the cognitive load of navigating four disconnected toolchains and the manual work of keeping geometry consistent across them. We have characterised that Fragmentation Gap precisely, surveyed the AI landscape across the three primary fabrication modalities, and presented FabOrchestrator,a deployed multi-agent system that reduces the whole thing to a single prompt-to-bundle interface.

The system runs on a production-grade full-stack web application built using our own coding agent infrastructure, which is itself a contribution to the emerging practice of AI-assisted software development at scale. Across 120 fabrication tasks, it achieved 91.3% cross-modal dimensional consistency, 87.6% 3D mesh validity, 96.2% laser path correctness, and 78.4% PCB DRC pass rate,with mean bundle generation under two minutes and average API cost of $0.11 per complete three-modality bundle.

The system's limitations,organic geometry, PCB verification intractability, machine safety,are not gaps we have tried to hide. They are boundaries we communicate clearly to users and operators. Within those boundaries, FabOrchestrator represents a defensible contribution to the maker community and to the broader project of distributed, democratised manufacturing.

IX. References

[1] N. Gershenfeld, Fab: The Coming Revolution on Your Desktop. Basic Books, 2005.

[2] N. Gershenfeld, "How to Make Almost Anything: The Digital Fabrication Revolution," Foreign Affairs, vol. 91, no. 6, pp. 43–57, Nov./Dec. 2012.

[3] Fab Foundation, "About the Fab Foundation," 2025. [Online]. Available: https://fabfoundation.org/about/

[4] B. Poole, A. Jain, J. T. Barron, and B. Mildenhall, "DreamFusion: Text-to-3D using 2D diffusion," in Proc. ICLR, 2023.

[5] M. Liu et al., "One-2-3-45: Any Single Image to 3D Mesh in 45 Seconds without Per-Shape Optimization," in Proc. NeurIPS, 2023.

[6] C. Li et al., "Generative AI meets 3D: A Survey on Text-to-3D in AIGC Era," arXiv:2305.06131, Oct. 2024.

[7] Y.-H. Hsu et al., "Exploring the Usability and Future Development of AI-Generated 3D Models," Computer-Aided Design and Applications, vol. 22, no. 5, pp. 782–804, 2025.

[8] A. Author et al., "Generative AI for CAD Automation: Leveraging Large Language Models for 3D Modelling," arXiv:2508.00843, 2025.

[9] H. Zou et al., "PCBSchemaGen: Constraint-Guided Schematic Design via LLM for Printed Circuit Boards," arXiv:2602.00510, Jan. 2026.

[10] Z. Chen et al., "CircuitLM: A Multi-Agent LLM-Aided Design Framework for Generating Circuit Schematics from Natural Language Prompts," arXiv:2601.04505, Jan. 2026.

[11] D. Schwachhofer et al., "SPICEAssistant: LLM using SPICE Simulation Tools for Schematic Design of Switched-Mode Power Supplies," arXiv:2507.10639, Jul. 2025.

[12] LangChain, "LangGraph: Building stateful multi-agent applications," 2025. [Online]. Available: https://langchain-ai.github.io/langgraph/

[13] Gartner, "Agentic AI Research 2025," Gartner Inc., 2025.

[14] R&D World, "The Quantified Factory: 2025 Manufacturing Capability Inflection," Dec. 2025.

[15] Cuttle, "Cuttle.xyz Documentation," 2025. [Online]. Available: https://cuttle.xyz/

[16] W. Gao et al., "AMSnet-KG: A Netlist Dataset for LLM-based AMS Circuit Auto-design Using Knowledge Graph RAG," ACM Trans. Design Autom. Electron. Syst., 2025.

[17] Anthropic, "Claude Sonnet 4: Model Card and Technical Report," Anthropic Inc., 2025.

[18] OpenCode Contributors, "OpenCode: A Headless AI Coding Agent with HTTP API," 2025. [Online]. Available: https://opencode.ai/

[19] Daytona, "Daytona Sandbox: Isolated Cloud Development Environments," 2025. [Online]. Available: https://daytona.io/

[20] Auth.js Contributors, "Auth.js v5 (NextAuth): Authentication for the Web," 2025. [Online]. Available: https://authjs.dev/

[21] Flexprice, "Flexprice: Usage-Based Billing and Credit Wallet Infrastructure," 2025. [Online]. Available: https://flexprice.io/

[22] KiCad Developers, "KiCad EDA: Open-Source Electronics Design Automation," 2025. [Online]. Available: https://www.kicad.org/

[23] PySpice Contributors, "PySpice: Python Interface to Ngspice and Xyce," 2025. [Online]. Available: https://pyspice.fabrice-salvaire.fr/

[24] SKiDL Contributors, "SKiDL: A Python Package for Schematic Design," 2025. [Online]. Available: https://devbisme.github.io/skidl/

AI-Based Image Authenticity Detection System

Akshaya M
Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad,India
akshaya_m@ahalia.ac.in

C Sriparvathy
Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad,India
sriparvathy_c@ahalia.ac.in

Deepthy PS
Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad,India
deepthy.ps@ahalia.ac.in

Aryananda P
Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad,India
aryananda_p@ahalia.ac.in

Sayana V
Computer Science Engineering
Ahalia School of Engineering & Technology
Palakkad,India
sayana_v@ahalia.ac.in

Abstract- ***The rapid advancement of artificial intelligence has enabled the creation of highly realistic synthetic images, raising concerns about misinformation, digital forgery, and trust in visual content. As AI-generated images become more widespread, distinguishing them from real images has become a critical challenge. This paper presents an AI-based image authenticity detection system that classifies images as real or AI-generated using deep learning.***
The system utilizes transfer learning with a pre-trained convolutional neural network (CNN) to automatically extract features, reducing the need for manual intervention. A labeled dataset of real and synthetic images is used to train the model for accurate and efficient performance. Implemented using Python and TensorFlow, with a Streamlit-based interface for real-time classification, the system provides a reliable and scalable solution for image authenticity verification in research and real-world applications.

Keywords- Artificial Intelligence, Image Authenticity, Deep Learning, CNN, Transfer Learning, Digital Forensics, Image Classification

I. INTRODUCTION

The emergence of advanced artificial intelligence technologies, particularly generative models such as Generative Adversarial Networks (GANs), has revolutionized the creation of digital images. These models can generate highly realistic images that are often indistinguishable from real photographs.
While such developments have opened new possibilities, they have also introduced serious challenges in maintaining the authenticity and credibility of digital content. Fake images can be used for spreading misinformation, creating deepfakes, and manipulating public perception. This has made it increasingly important to develop automated systems capable of detecting whether an image is genuine or artificially generated.

This project focuses on designing and implementing an AI-based system that can accurately classify images as real or AI-generated. By leveraging deep learning techniques, the system aims to provide an efficient and scalable solution to address the growing problem of image authenticity.

II. LITERATURE SURVEY

Recent research in image forensics has significantly evolved with the rapid advancement of artificial intelligence and generative models, which are capable of producing highly realistic synthetic images. Traditional image manipulation detection techniques primarily relied on handcrafted features such as noise patterns, color inconsistencies, pixel-level irregularities, and compression artifacts.
These methods were effective in identifying basic editing operations like splicing or retouching; however, they often fail when dealing with modern AI-generated images that exhibit high visual fidelity and minimal detectable artifacts. As generative models such as GANs and diffusion-based systems improve, the gap between real and

synthetic images continues to narrow, making traditional detection approaches increasingly unreliable. This has created a strong need for more advanced, automated, and scalable solutions capable of identifying subtle differences that are not easily visible to the human eye.

To address these challenges, deep learning-based approaches, particularly Convolutional Neural Networks (CNNs), have become the dominant methodology in image authenticity detection.
Foundational work by Yann LeCun, Yoshua Bengio, and Geoffrey Hinton demonstrated the power of deep neural networks in automatically learning hierarchical feature representations from raw data, eliminating the need for manual feature extraction. Building on this, advanced CNN architectures such as VGGNet, developed by Karen Simonyan and Andrew Zisserman, enabled deeper networks with improved feature extraction capabilities. Similarly, architectures like ResNet introduced residual learning to address vanishing gradient problems, while EfficientNet optimized model scaling for better performance with fewer parameters. Transfer learning has further enhanced these approaches by allowing models pre-trained on large datasets to be fine-tuned for specific tasks such as fake image detection, significantly reducing training time and computational requirements while maintaining high accuracy.

Recent studies have also focused on improving detection robustness, interpretability, and real-world applicability. For example, the CIFAKE framework proposed by J.J. Bird and A. Lotfi emphasizes explainable AI, enabling users to understand how and why an image is classified as real or synthetic, thereby increasing trust in automated systems. Additionally, models like CGNet incorporate attention mechanisms to focus on subtle visual artifacts and inconsistencies that are characteristic of AI-generated content. The literature also highlights the importance of using large-scale, diverse, and well-labeled datasets to improve generalization and reduce bias in detection models. Furthermore, modern implementations leverage powerful deep learning frameworks such as PyTorch and TensorFlow to streamline model development, training, and deployment. Many systems now integrate user-friendly, real-time interfaces, making these tools accessible beyond research environments and suitable for practical applications such as social media monitoring, journalism, and digital forensics. Overall, the literature demonstrates a clear shift toward deep learning-driven, scalable, and interpretable solutions that are better equipped to address the growing challenges of distinguishing real images from AI-generated content.

III. PROPOSED SYSTEM

A. System Overview

The proposed AI-Based Image Authenticity Detection System identifies whether an image is real or AI-generated using advanced deep learning techniques. With the rise of generative models such as GANs and diffusion models, distinguishing synthetic images from real ones has become challenging, and this system provides an automated solution to address this issue.

The system operates through a simple web-based interface where users upload images for analysis. The uploaded images undergo preprocessing steps such as resizing, normalization, and format conversion to ensure consistency and improve model performance.

At its core, the system uses a Convolutional Neural Network (CNN) trained on both real and AI-generated images. The model learns to detect subtle differences like texture inconsistencies and visual artifacts, enabling accurate classification.

The modular architecture includes user interface, preprocessing, detection, and output modules, ensuring flexibility and ease of maintenance.
Overall, the system is efficient, scalable, and user-friendly, with applications in digital forensics, media verification, and cybersecurity.

B. System Challenges

Developing the AI-Based Image Authenticity Detection System involves several challenges due to the high realism of AI-generated images, which often show minimal visible differences. Accurately classifying such images requires extensive training on diverse and well-labeled datasets, which can be time-consuming to collect and organize. Variations in image resolution, lighting, and format also necessitate effective preprocessing to standardize inputs.

Computational complexity is another concern, as CNN-based models demand substantial resources for training and inference. Optimizations are needed to ensure efficient performance without sacrificing accuracy. Integration between the frontend interface and backend model adds further complexity, requiring smooth data flow, error handling, and fast, reliable responses.

Addressing these challenges enhances the system's robustness, accuracy, and suitability for real-world applications in digital forensics, media verification, and cybersecurity.

C. Key Components of the Model

The proposed AI-Based Image Authenticity Detection System consists of several key components that work together for efficient and accurate image classification.

The User Interface Module allows users to upload images and view results through a simple, intuitive web-based platform, ensuring accessibility for non-technical users. The Image Preprocessing Module standardizes inputs by resizing, normalizing, and converting images to a suitable format for the model, improving consistency and performance.

At the core is the Feature Extraction and Detection Module, based on a Convolutional Neural Network (CNN). It analyzes images to extract features like edges, textures, and patterns and distinguishes subtle differences between real and AI-generated images. The Classification Module interprets these features to label images as "Real" or "AI-Generated," often with a confidence score.

Finally, the Output Display Module presents results clearly on the interface, providing quick and understandable feedback. The modular design allows seamless updates and future enhancements, making the system adaptable to advances in AI and image generation technologies.

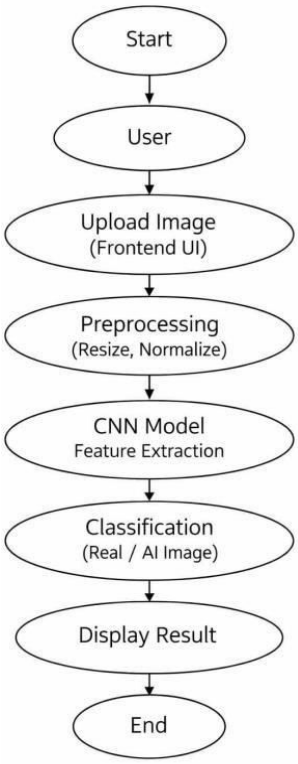
D. System Architecture

The proposed system follows a structured and modular architecture that ensures efficient data flow and clear separation of functionalities. It consists of multiple components, each responsible for a specific stage in the image authenticity detection process.

The user interface module serves as the front end, allowing users to upload images through a simple web-based platform accessible via browsers. Once an image is uploaded, it is sent to the preprocessing module, where operations such as resizing, normalization, and format conversion are performed to ensure consistency and compatibility with the model.

The processed image is then passed to the detection module, which is the core of the system. This module uses a Convolutional Neural Network (CNN) trained on both real and AI-generated images to extract features and identify patterns.
Based on this analysis, the model classifies the image as either real or synthetic.

Finally, the output module presents the result to the user in a clear and understandable format, optionally including confidence scores. The modular design ensures flexibility, allowing easy updates or integration of improved models. Overall, the architecture provides an efficient, scalable, and user-friendly solution for image authenticity detection.



IV. IMPLEMENTATION

The implementation of the proposed AI-Based Image Authenticity Detection System involves the integration of multiple technologies to create a complete and functional solution. The system is developed using a combination of machine learning frameworks, web technologies, and data management tools to ensure efficient processing and user interaction. The implementation is divided into backend development, frontend development, database design, and deployment strategies. Each component is carefully designed to ensure smooth operation, high accuracy, and scalability of the system. The overall workflow includes image acquisition, preprocessing, feature extraction, classification using a trained deep learning model, and result visualization. Emphasis is placed on optimizing performance, minimizing latency, and ensuring that the system can handle multiple user requests effectively.

A. Backend Development

The backend of the system is responsible for handling the core processing and computational tasks required for image authenticity detection. It is primarily developed using Python due to its extensive support for machine learning and image processing libraries. Frameworks such as TensorFlow or Keras are used to build and train the Convolutional Neural Network (CNN) model. The backend pipeline begins with receiving the uploaded image from the user interface and performing preprocessing operations such as resizing, normalization, and format conversion.

Once the image is preprocessed, it is passed to the trained CNN model, which extracts features and performs classification. The model has been trained on a dataset consisting of both real and AI-generated images, allowing it to learn distinguishing characteristics such as texture inconsistencies, unnatural patterns, and generation artifacts. The backend also includes error handling mechanisms to manage invalid inputs and ensure system stability. Additionally, APIs are implemented to facilitate communication between the frontend and backend, enabling seamless data exchange and real-time result delivery.

B. Frontend Development

The frontend of the system is designed to provide an intuitive and user-friendly interface for interacting with the application. It is developed using a lightweight web framework such as Streamlit, which allows rapid development of interactive web applications. The interface enables users to upload images, preview the selected image, and receive classification results in a clear and concise manner.

Special attention is given to simplicity and ease of use, ensuring that even users with minimal technical knowledge can operate the system effectively. The layout is structured to guide the user through the process, with clearly labeled buttons and instructions. The frontend also displays the output in an understandable format, such as “Real Image” or “AI-Generated Image,” along with optional confidence scores. The responsiveness of the interface ensures that users receive feedback quickly, enhancing the overall user experience.

C. Database Design

The proposed system uses a dataset to train the deep learning model but does not require a database during runtime. It performs real-time image classification by processing input images temporarily and displaying results immediately without storing any data. This approach reduces system complexity and improves performance.
However, a database can be integrated in the future to store images, results, and user history for further analysis and system enhancement.

D. Scalability and Deployment

The proposed AI-Based Image Authenticity Detection System is designed with scalability in mind to support future growth and increased user demand. Although the current implementation is developed as a lightweight application for demonstration purposes, the system can be extended to handle large-scale data and multiple users simultaneously. The modular architecture allows easy integration of improved deep learning models and larger datasets, enabling the system to adapt to advancements in AI-generated image technologies.

In terms of deployment, the system can be executed locally for testing and development purposes. For real-world applications, it can be deployed on cloud platforms, allowing users to access the system through web browsers from any location. Cloud

deployment improves performance, availability, and scalability by providing flexible computing resources. Additionally, techniques such as load balancing and parallel processing can be incorporated to ensure efficient handling of multiple requests.

The system is also designed to be easily upgradable, allowing future enhancements such as real-time detection, API integration, and mobile accessibility. This ensures that the system remains relevant and effective as technology evolves.

V. EVALUATION

The proposed AI-Based Image Authenticity Detection System is evaluated to assess its ability to distinguish real from AI-generated images. A trained CNN model is tested on a dataset of authentic and synthetic images, with separate training and testing sets to ensure fair evaluation.

Key performance metrics include accuracy, precision, and recall. Accuracy measures overall correctness, precision indicates how well AI-generated images are identified without misclassifying real images, and recall evaluates the detection of all synthetic images.

The system's efficiency is also considered, as it is designed for real-time use. Results demonstrate that the model delivers accurate predictions quickly, providing reliable outputs suitable for practical applications.

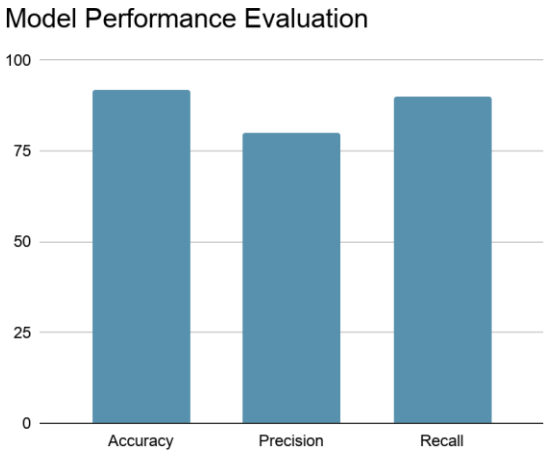
A. Model Accuracy and Performance

The proposed system demonstrates high accuracy in classifying images as real or AI-generated, thanks to the CNN model's ability to extract meaningful features such as textures, edges, and patterns. These features enable the model to detect subtle differences between authentic and synthetic images, enhancing classification performance.

The system also shows strong precision and recall, producing fewer false classifications while effectively identifying AI-generated images. This balance ensures reliability and suitability for practical applications. Additionally, the model operates efficiently, providing quick predictions without significant computational delay.

Overall, the results confirm that the system is effective in detecting AI-generated images. With further improvements, including larger datasets and advanced models, accuracy and robustness can be further enhanced, making it well-suited for real-world deployment in areas such as digital forensics, media verification, and cybersecurity.

The performance of the proposed system is illustrated using a bar graph representing accuracy, precision, and recall metrics. The model achieves high accuracy, indicating effective classification of real and AI-generated images.



VI. ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the management and faculty of the Department of Computer Science and Engineering, [Your College Name], for their continuous support and encouragement throughout the development of this project.

We would like to extend our heartfelt thanks to our project guide, [Guide Name], for their valuable guidance, constructive suggestions, and constant motivation, which greatly contributed to the successful completion of this work.

We also thank all the team members for their cooperation and dedication. Finally, we acknowledge the support of our friends and family for their encouragement and understanding during the course of this project.

VII. CONCLUSION

This paper presented an AI-Based Image Authenticity Detection System capable of distinguishing between real and AI-generated images using deep learning techniques. The system employs a Convolutional Neural Network (CNN) to analyze image features and perform accurate classification. The results demonstrate that the proposed approach achieves high accuracy while maintaining efficiency and ease of use.

The system is designed with a modular structure, enabling flexibility and future enhancements. It performs real-time image classification without the need for data storage, simplifying the overall implementation. The effectiveness of the system makes it suitable for applications in digital forensics, media verification, and cybersecurity.

Future improvements may include the use of larger datasets, advanced models, and cloud-based deployment to further enhance performance and scalability.

VIII. REFERENCES

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative Adversarial Networks," Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), 2014.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems, 2012.

[4] C. Szegedy et al., "Going Deeper with Convolutions," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[5] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," International Conference on Machine Learning (ICML), 2019.

[6] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2017.

[7] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, 2015.

[8] D. Cozzolino, G. Poggi, and L. Verdoliva, "Recasting Residual-based Local Descriptors as Convolutional Neural Networks: An Application to Image Forgery Detection," IEEE Signal Processing Letters, 2017.

[9] H. Nguyen, J. Yamagishi, and I. Echizen, "Capsule-Forensics: Using Capsule Networks to Detect Forged Images and Videos," ICASSP, 2019.

[10] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2019.

YOUTUBE CONTEXT ANALYZER USING NATURAL LANGUAGE PROCESSING

Dr Gunasekaran S
Professor & HOD,CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
gunaphd@yahoo.com

Ms.Dhanya P G
Assistant Professor, CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
dhanyarenjith2025@gmail.com

Aswan K Prakash
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
aswan_kprakash@ahalia.ac.in

Gopika B
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
gopika_b@ahalia.ac.in

Nivedh A Nair
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
nivedh.a.nair@ahalia.ac.in

Rithu P
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
rithu_p@ahalia.ac.in

***Abstract*— The YouTube Sentiment Analyzer is a website that analyzes what people say about YouTube videos, helping creators, researchers, and analysts understand public opinion in today's content-heavy world. It collects comments from videos and uses Natural Language Processing (NLP) to classify them as positive, negative, or neutral. The system also extracts keywords, analyzes emojis, and groups comments to provide deeper insights. It presents results through visualizations like word clouds and pie charts, and can generate summaries of audience opinions. Additionally, it includes a chatbot for user interaction. The frontend is built with React.js, while the backend uses Python Flask, making the platform fast, efficient, and easy to use without manual effort, while improving accuracy.**

Keywords*— *Sentiment Analysis, YouTube Comments, Natural Language Processing, Machine Learning, Chatbot, Data Visualization, Flask, React

I. INTRODUCTION

YouTube is a place where people share and watch videos about things like learning, fun and technology. People do not just watch videos on YouTube they also talk about them by leaving comments saying they like them and discussing them. These comments are important because they show what people think and feel about the videos.. It is hard to understand what people are saying because there is just so much to look at. Popular videos can have thousands of comments. It is impossible to read all of them. Also comments are often written in a casual way with slang, abbreviations, emojis and sometimes mean or sarcastic language, which makes them hard to understand. Video transcripts can be very long and not well organized making it hard for people to get the ideas without watching the whole video.

To fix these problems the YouTube Context Analyzer is a tool that helps people understand what is in the videos and what people are saying about them. It makes summaries of the videos so people can understand the main points without watching the whole thing. At the time it looks at the comments to see how people feel finds bad language and figures out what the emojis mean. The tool also finds ideas in the comments and shows them in a simple way using pictures and summaries. It even has a chat feature, like a robot that talks to people so they can ask questions. Get answers in a natural way. The YouTube Context Analyzer takes a lot of information from YouTube and makes it into something that is easy to understand and useful.

II. RELATED WORK

Many people have looked into how to understand what people are saying on YouTube and other social media sites. A time ago people used simple machine learning to figure out what people meant. They used things like Bayes and Decision Trees to look at words and decide what they meant. These methods worked okay for things but they were not great at understanding what people were really saying.

As time went on new ways of doing things were found. People started using things like Recurrent Neural Networks and Long Short-Term Memory networks. These were better at understanding what people meant because they could look at the order of the words and how they related to each other. Recently people have started using things like BERT and GPT. These are really good at understanding what people mean even when they are being sarcastic or using slang. They are great for looking at YouTube comments because they can understand the context and the meaning of the words.

People have also been working on ways to summarize YouTube videos. This is helpful because it lets people understand what a video is about without having to watch the thing. At first people just pulled out the important sentences from the video.. Now people are using big language models to make summaries that sound like they were written by a person. Some systems can even look at the comments and the video at the time to get a better understanding of what people think.

However most of these systems only do one thing. They might do sentiment analysis or summarization or something else. They do not combine all of these things into one system.

III. DRAWBACKS OF THE EXISTING SYSTEM

There are still some problems with the systems that exist now. The old machine learning methods are not great at understanding what people mean especially when they are being sarcastic or using slang. The new methods are better. They need a lot of data and a lot of power to run. Even the advanced methods, like BERT and GPT are not perfect. They are slow. Need to be trained specially for each task.

When it comes to summarizing videos the old methods were not great. They just pulled out sentences without understanding what they meant. The new methods are better. They can still make mistakes if they are not trained right.. A lot of these systems rely on the transcript of the video, which is not always accurate.

Most of the systems that exist now only look at one thing. They might look at the sentiment of the comments. Summarize the video but they do not do both. They do not look at things, like emojis or offensive language. Let people ask questions. This means that people do not get a picture of what is going on. They have to use a lot of tools to get the information they need. This is not very useful. It is not very efficient.

All of these problems show that we need a system. We need a system that can look at the video and the comments and understand what people mean. We need a system that can summarize the video and look at the sentiment of the comments. We need a system that's fast and accurate and easy to use.

IV. PROBLEM STATEMENT

YouTube has a collection of videos on all sorts of things like learning, entertainment and sharing information. A lot of these videos are really long. It is hard for people to get the main points without watching the whole thing. This is a problem because it takes a time especially for people who just want to get a quick idea of what the video is about. Also people leave a lot of comments on videos and these comments show what people think and feel about the video. It takes much time to read all these comments and try to understand what people are saying.

The old ways of dealing with this problem are not good enough. They. Need people to do all the work or they only do one thing like summarize the video or look at how people feel about it.

This project is trying to solve this problem with the YouTube Context Analyzer. This is a system that can summarize videos and look at how people feel about them at the same time. The YouTube Context Analyzer looks at the video. Pulls out the important words then uses special computer programs to make a short summary. It also uses computer programs to look at the comments people leave and figure out if they like or dislike the video. By doing both of these things the YouTube Context Analyzer makes it easier for people to understand the video and what others think about it which saves time and gives people useful information, about the YouTube videos and what people think of them.

V. LITERATURE SURVEY

This section looks at research that is related to video transcript summarization and sentiment analysis of YouTube comments which're really important for making the YouTube Context Analyzer.

a. Atluri et al. (2022) – Video Transcript Summarizer

The system that Atluri et al. Developed takes the audio from a video. Turns it into text using speech-to-text techniques. Then it makes a short summary using techniques based on natural language processing. This system finds the important sentences so that the video is shorter but it still has all the important information. However it does not think about what the audience says.

b. Kant et al. (2024) – YouTube Comments Sentiment Analysis

Kant et al. Made a system to look at the sentiments in YouTube comments. It first gets the comments by breaking them down into words and removing common words. Then it uses machine learning to put the comments into groups based on how people feel. This way we can see what the public thinks. It does not relate it to what the video is about.

c. Singh & Tiwari (2021) – YouTube Comments Sentiment Analysis

This work is about putting comments into groups based on whether they're good, bad or neutral. It uses machine learning models and looks at the words in the comments. While it works it is not good at understanding things.

d. Parwateeswar et al. (2025) – YouTube Video to Text Summarizer

They made a system that takes YouTube videos and turns them into text using Automatic Speech Recognition and summarization algorithms. This helps people understand what the video is quickly but it does not look at how people feel about the video.

e. Aiswarya & Rajeev (2024) – YouTube Comment Sentiment Analysis

This study looks at how the audience feels about things using language processing and classification models. It also makes pictures to show how people feel. However it does not connect how people feel with what the video's, about.

Summary

The research shows that video summarization and sentiment analysis are well studied,. They are usually done separately. The YouTube Context Analyzer wants to put both so we can understand the video and what the audience thinks about it.

VI. PROPOSED SOLUTION

The proposed method offers a comprehensive framework for analyzing YouTube content by integrating video transcripts with user-generated comments to provide a holistic view of digital discourse. Unlike traditional metrics such as view counts or likes, which only measure surface-level engagement, this approach utilizes advanced NLP models to distill video transcripts into concise summaries that preserve core thematic ideas. Simultaneously, it evaluates the comment section to gauge audience sentiment, identifying emotional tones, offensive language, and the nuanced use of emojis. By synthesizing what the video conveys with how the audience reacts, the system bridges the gap between content creation and viewer perception.
To process this high volume of text data effectively, the method employs sophisticated text-embedding techniques that transform comments into numerical codes, allowing the system to grasp underlying meanings. These embeddings are then organized using clustering algorithms like K-Means to categorize discussions into distinct themes. By selecting representative comments from each cluster, the system eliminates redundancy while ensuring that diverse perspectives are preserved. This unified architecture—combining summarization, sentiment classification, and thematic grouping—transforms raw data into actionable, contextual insights that far exceed the capabilities of standard analytical tools.

VII. METHODOLOGY

The way we do things is pretty straightforward. We take a YouTube video URL from the user. Use it to get the transcript and comments. This YouTube data is not very organized and can have some mistakes so we clean it up first. We remove any symbols get rid of duplicate comments or spam and make sure everything is in the same format. This helps us get good quality YouTube data that we can work with.

Next we look at the transcript to get a summary of what the YouTube video is about. This way people can understand the points without having to watch the whole YouTube video. At the time we look at the comments to see what people think and feel about the YouTube video. We want to find out if people like or dislike the YouTube video. If there are any bad words or mean comments. We change the text into numbers so the system can understand what the comments are saying. Then we group comments together to find out what people are talking about the most in the YouTube video comments.

Finally we put everything together in a way that makes sense. We show what the YouTube video is, about what people are saying about the YouTube video and what the main discussion points are. Our goal is to make sure everything flows logically and's easy to understand without getting into the technical details of how the system works with the YouTube data.

VIII. SYSTEM ARCHITECTURE

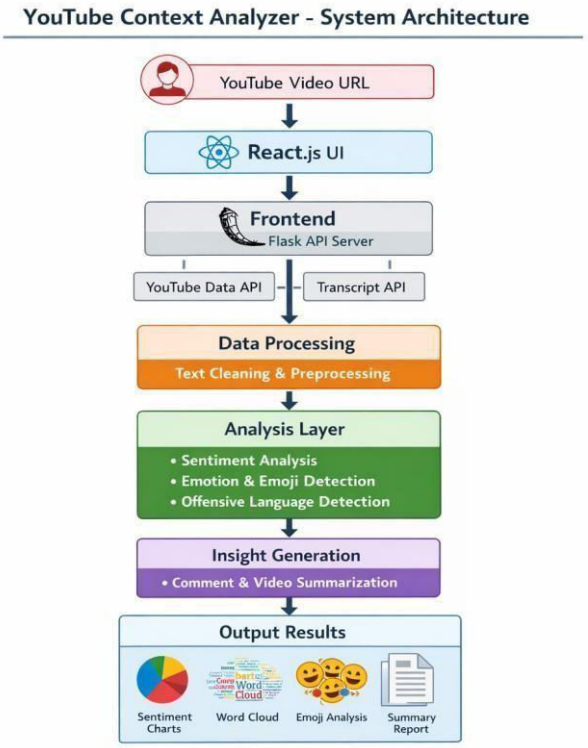
The YouTube Context Analyzer is built using a flexible and modular design that enables efficient handling of large volumes of unstructured data while allowing easy integration of new features in the future. Each component of the system is designed to perform a specific function, and all components work together in a coordinated manner without being tightly coupled. The overall flow begins with user interaction and progresses through multiple layers until structured and meaningful insights are delivered.

At the top level, the system includes a presentation layer that acts as the user interface. Through this interface, users provide a YouTube video URL and receive results in a clear, organized, and visually understandable format. This layer is responsible for managing user interactions, sending requests to the backend, and displaying outputs such as the video summary, sentiment distribution, and grouped insights. It ensures smooth communication with the backend through API calls and provides a responsive experience to the user.

The application layer forms the core control unit of the system, where the main logic is handled. It receives requests from the presentation layer, validates the input, and manages the sequence of operations required to generate results. This layer interacts with external services to fetch video transcripts and comments, and then coordinates the workflow by organizing preprocessing, analysis, and result aggregation steps. It also handles error management, request routing, and response formatting to ensure that the output sent back to the frontend is structured and consistent.

The processing layer is responsible for performing all analytical tasks within the system. It consists of modules for text preprocessing, video summarization, sentiment analysis, and comment grouping. Each module processes cleaned data to generate intermediate outputs, which are later combined to produce final insights. The modular nature of this layer allows individual components to be updated or replaced without affecting the entire system, improving maintainability and flexibility.

Data flows seamlessly across these layers, starting from user input and ending with the delivery of meaningful results. The architecture is designed to optimize performance, reduce latency, and efficiently handle multiple requests. This ensures that the system remains scalable, maintainable, and capable of delivering accurate insights while keeping all components loosely coupled and well-coordinated.



IX. IMPLEMENTATION

The YouTube Context Analyzer is built to be a system that handles everything from getting data to analyzing and showing it in a useful way. This system is made up of parts that work together but each part does its own job and does not interfere with the others. This makes it easy to keep the system running make it bigger and improve it over time. The whole process starts when the user gives some input then it goes through steps on the backend and finally it shows the user what it found in a way that is easy to understand. The system is made to be fast easy to use and good at talking to all the parts using special ways of communicating called APIs.

The system uses technologies from the internet and machine learning to handle big amounts of data that are not organized, like transcripts and comments from YouTube videos. A lot of care is taken to make sure the system can do its job quickly by doing tasks in the order and getting ready for future improvements like saving things in memory and doing many things at the same time. The system also uses tools that have already been trained to understand human language so it does not have to learn from scratch. This means it can give highquality results without needing a lot of work. The system is also designed to make sure the results are not just correct, but easy for users to understand by organizing and showing them in a clear way.

Backend Implementation

The backend is the heart of the system. Is built using Python and a framework called Flask. Flask is used to make special addresses on the internet called APIs that let the frontend and backend talk to each other. When a user gives the system a YouTube video link the backend gets the request checks it and starts doing its job. To keep everything the backend uses a special tool called Flask-CORS, which lets the frontend talk to the backend even if they are on different computers or networks.

To get the data it needs the backend uses tools like the YouTube Transcript API to get the subtitles of videos and other methods to get comments from users. The data it gets is usually not organized and has some mistakes so it needs to be cleaned up using libraries like Pandas and NumPy. This includes fixing the text removing duplicates and making sure everything is consistent. This makes sure the data is good enough to be used for analysis.

The backend uses tools like Hugging Face Transformers to understand the language in the videos and comments. It can even make a summary of the video that captures the points. For comments it uses another tool called TweetNLP to figure out how people are feeling, detect language and understand the emojis they use. To help understand what comments are similar it uses Sentence Transformers to turn the comments into numbers that capture what they mean.

These numbers are then used with another library called Scikit-learn, which groups comments together. This helps find topics that people are talking about and reduces the amount of repeated information in big datasets. The backend then picks some comments from each group. Puts all the results together including the summary, how people are feeling and what the comments are about. Finally it sends all this information back to the frontend in a format called JSON. The backend is designed to do things at the same time and can be made even faster using special techniques like saving things in memory or doing many tasks at once.

Frontend Implementation

The frontend is what the user sees and interacts with. It is built using a library called React. React lets developers make parts of the user interface that can be used times and makes sure the interface updates quickly. The user interface is made using HTML, CSS and JavaScript with a focus on being simple, clear and easy to use.

React has tools like useState and useEffect that help manage what is happening in the application and handle tasks that take some time like calling the backend. When a user gives the system a YouTube video link the frontend sends a request to the backend using internet methods. Once the backend does its job and sends back the results the frontend updates the user interface without needing to reload the page.

The results are shown in parts, including a summary of the video how people are feeling about it and what they are saying in the comments. To make it easier to understand and more interesting the frontend uses tools to make charts and graphs that show how people are feeling and other insights. The frontend makes sure that the complex information from the

backend is shown in an intuitive way so users can quickly understand the video and what people think about it.

Overall the frontend is like a bridge between the user and the backend system making sure everything works smoothly updates, in time and is easy to use while also being fast and responsive.

X. RESULT

From the results, it is evident that the system can effectively convert raw YouTube data into insightful information through video summarization and analysis of comments. The system was tested using different videos, and it was evident that it can effectively create accurate video summaries and efficiently process comments from the video audiences. The user can effectively comprehend the overall sentiment of the video, whether it is positive, negative, or neutral, without having to go through large amounts of comments. Other features like detection of offensive language and interpretation of emojis can be added to enhance the overall analysis of comments, considering the casual nature of comments posted online.

The system can be considered efficient since it can perform its tasks without a significant delay. Moreover, it can be considered reliable since it can effectively convert video data and comments into insightful information. It can be considered scalable since it can effectively perform its tasks without a significant delay. The summarization feature can be considered efficient since it can effectively reduce the overall time taken to comprehend video content. The chatbot feature can be considered efficient since it can be effectively used for querying data analysis.

XI. CONCLUSION

The YouTube Context Analyzer has been an effective approach in understanding the content of the video, as well as the audience's feedback, through the incorporation of video summarization and comment analysis. This allows the users to gain a clear understanding of the content within a short period of time, without the need to go through the hassle of processing large amounts of data. The utilization of Natural Language Processing, machine learning, and visualization has made the data complex, unstructured, clear, meaningful, and simple, allowing the users to analyze the data in an efficient manner.

The system is user-friendly, accessible, and fun to use, as it has incorporated visual representations and the use of a chatbot, which has made the system convenient to use. The system has been effective in providing timely results, allowing the users to apply the system in real-life situations. In the future, the system can be improved further by allowing the users to input data in different languages, handling large amounts of data, and incorporating more advanced machine learning models to improve the performance of the system.

ACKNOWLEDGEMENT

We want to say thank you much to our project guide, Ms. Dhanya P G for helping us so much. She was always there to support us and encourage us while we were working on this project. Her ideas and advice really helped us make our work better.

We also want to thank the teachers at the Department of Computer Science and Engineering Ahalia School of Engineering and Technology for giving us what we needed to do this project and for making it a nice place to work.

We are also thankful to our friends and family for motivating us and supporting us all the time while we were finishing this project. The project was a task and our friends and family helped us a lot. We really appreciate the project and the people who helped us with the project like our project guide, Ms. Dhanya P G and the teachers, at the Department of Computer Science and Engineering Ahalia School of Engineering and Technology.

REFERENCES

[1] Atluri, N. S. S. V., Laggisetti, V. S., Duvvuru, J., & Bayana, J. (2022). Video transcript summarizer. In Proceedings of the 2022 International Mobile and Embedded Technology Conference (MECON) (pp. 461–465). IEEE. https://doi.org/10.1109/MECON53876.2022.9751991

[2] Kant, S., Upadhyay, R. K., Gupta, P., Tiwari, R., & Kumar, H. (2024). YouTube comments sentiment analysis. In Proceedings of the 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N 2024). IEEE. https://ieeexplore.ieee.org/servlet/opac?punumber=10894760

[3] Singh, R., & Tiwari, A. (2021). YouTube comments sentiment analysis. International Journal of Scientific Research in Engineering and Management, 5(5), https://www.ijsrem.com

[4] Parwateeswar, G., Ramya, N., Kumar, D. S., Sai, G. S., & Avinash, C. K. (2025). YouTube video to text summarizer. International Research Journal of Modernization in Engineering Technology and Science, 7(6), Article 4561. https://doi.org/10.56726/IRJMETS80156

[5] Aiswarya, A. S., & Rajeev, H. (2024). YouTube comment sentimental analysis. Indian Journal of Data Mining, 4(1), 5. https://doi.org/10.54105/ijdm.A1633.04010524

LOST AND FOUND SYSTEM

HOD: Dr GunasekaranS
Dept of CSE,
Ahalia School Of Engineering and Technology palakkad, Kerala
gunaphd@yahoo.com

GUIDE:Mr.H.A.Rahulnath
Assistant Professor
Dept of CSE,
Ahalia School of Engineering and Technology Palakkad, kerala
rahul.harindranath@gmail.com

Athul krishna K
Dept of CSE,
Ahalia School of Engineering and Technology Palakkad, kerala
ATP23CS027
athulkrishna_k@ahalia.ac.in

Midhun T Chandran
Dept of CSE,
Ahalia School of Engineering and Technology Palakkad,kerala
ATP23CS048
midhunt_chandran@ahalia.ac.in

Adarash A
Dept of CSE,
Ahalia School of Engineering and Technology Palakkad,kerala
ATP23CS003
adarsh_a@ahalia.ac.in

***Abstract* — The Lost and Found Management System is a web-based application developed to efficiently manage and track lost and found items within an institution. Traditional methods rely on manual records such as registers and notice boards, which are inefficient, unorganized, and prone to human errors. These limitations often lead to poor tracking of items and reduced chances of recovery. The proposed system provides a centralized digital platform that enables users to report lost or found items, search for items using advanced filtering, and submit claims through a structured process. The front end is developed using HTML, CSS, and JavaScript to provide a responsive and user-friendly interface. The backend is implemented using Node.js with the Express framework, which efficiently handles server-side logic and manages multiple user requests concurrently. MySQL is used as the relational database for storing structured data such as user details, item information, and claim records.**
To improve the efficiency of item discovery, the system incorporates a smart matching mechanism using MySQL full-text search combined with ranking techniques. This approach analyzes user queries, assigns relevance scores to stored items, and returns results in an ordered manner based on similarity.

***Keywords* - Lost and Found System, Web Application, Smart Matching, Claim Management, MySQL, REST API.**

I. INTRODUCTION

Lost and found item management is a recurring challenge in educational institutions, offices, and public environments. Conventional systems rely heavily on manual record-keeping methods such as physical logs and notice boards, which are inefficient, difficult to maintain, and lack proper organization. These systems often fail to provide effective communication between users and do not support efficient searching or tracking of items.

The Lost and Found Management System aims to address these challenges by providing a web-based platform that allows users to digitally report lost or found items and search for them using an organized and structured approach. Users can access the system from any device with internet connectivity, making it convenient and accessible.

A major limitation of traditional and basic digital systems is the lack of intelligent search functionality. Users often struggle to find relevant items due to poor search mechanisms that rely only on exact keyword matching. To overcome this limitation, the proposed system integrates a smart matching feature that improves search efficiency by analyzing user queries and ranking results based on relevance.

The use of Node.js ensures high performance and scalability by handling asynchronous operations efficiently, while MySQL provides a robust relational database structure for managing data. Together, these technologies enable the system to handle large volumes of data and multiple users effectively, making it suitable for real-world deployment.

II. RELATED WORK

Traditional lost and found systems are primarily based on manual record-keeping methods such as registers, logbooks, and physical notice boards. In these systems, details of lost and found items are recorded manually or displayed publicly for users to view. Although simple to implement, these methods are inefficient, difficult to manage, and highly prone to human errors. They lack proper organization, do not support efficient searching, and often fail to provide timely updates. As a result, users face difficulty in locating relevant items, and many items remain unclaimed.

With the advancement of web technologies, several digital solutions have been introduced to improve the lost and found process. Early web-based systems allowed users to list and browse lost or found items through basic interfaces. These systems improved accessibility and reduced dependency on physical records.

However, they often lacked essential features such as secure management, and claim verification mechanisms. In addition, their search functionality was limited, requiring users to manually browse through large lists of items. Many existing web-based systems implement basic keyword matching using SQL queries, typically relying on pattern matching techniques such as the LIKE operator. While this approach enables partial matching, it does not consider the relevance or importance of search results. Consequently, users may receive a large number of loosely related results, making it difficult to identify the correct item. These systems also do not support ranking or prioritization of results, which reduces search efficiency. Recent developments in database technologies and information retrieval systems have introduced more advanced search techniques such as full-text search and relevance-based ranking. Full-text search allows efficient indexing of textual data and enables faster retrieval of results based on keyword similarity. Ranking algorithms further enhance this process by assigning scores to results based on factors such as term frequency, importance, and contextual relevance. These techniques significantly improve the accuracy and usability of search systems. In addition to full-text search, some research has explored the use of heuristic approaches, tagging systems, and metadata-based classification to improve item discovery. These methods help in organizing data and improving search efficiency but often require additional configuration and management. More advanced approaches have also considered the use of machine learning and artificial intelligence techniques for matching items based on textual and visual similarity. However, such approaches may increase system complexity and computational requirements, making them less suitable for small to medium-scale institutional applications. Despite these advancements, many existing systems either lack user-friendly interfaces or fail to integrate advanced search techniques effectively. Some systems focus primarily on functionality without considering usability, while others are overly complex and difficult to implement. The proposed system addresses these limitations by combining a simple and intuitive web-based interface with efficient backend processing and smart matching techniques. By utilizing MySQL full-text search and relevance ranking, the system provides accurate and prioritized search results while maintaining simplicity and scalability. This makes it well-suited for institutional environments where ease of use, efficiency, and reliability are essential.

III. PROPOSED SOLUTION

The proposed Lost and Found Management System is a web-based application designed to provide a centralized and efficient platform for managing lost and found items within an institution. The system aims to replace traditional manual methods with an automated and structured approach that improves data organization, accessibility, and communication between users. The application allows users to report lost items and found items by entering detailed information such as item title, description, category, location, and images. This information is stored in a MySQL relational database, ensuring structured data storage, consistency, and efficient retrieval. Each item is uniquely identified and associated with the user who reported it, enabling proper tracking and management. A key feature of the proposed system is the implementation of a smart matching mechanism that enhances the search functionality. When a user enters a search query, the system processes the input and performs a full-text search across relevant fields such as title and description. The MySQL full-text search engine evaluates the similarity between the query and stored data and assigns a relevance score to each item. The results are then sorted in descending order of relevance, ensuring that the most relevant items are displayed first. This approach significantly improves search accuracy compared to traditional keyword-based matching. The system also includes a claim management module that allows users to request ownership of an item. When a user submits a claim, the request is stored in the database and forwarded to the item owner for verification. The owner can review the claim details and either approve or reject the request. Upon approval, the system updates the item status to “claimed” and enables communication between the claimant and the owner through contact details. This structured claim process ensures authenticity and reduces the possibility of false claims. In addition to core functionalities, the system provides features such as item status tracking, notification updates, and a user-friendly interface for seamless interaction. Items are categorized based on their status, such as lost, found, or claimed, allowing users to easily identify their availability. Notifications are displayed to inform users about claim approvals and updates, improving communication and responsiveness.

The system architecture is designed to ensure scalability and performance. The front end is developed using HTML, CSS, and JavaScript to provide an intuitive interface for users. The backend is implemented using Node.js with the Express framework, which handles request processing, business logic, and communication with the database. The MySQL database serves as the data storage layer, ensuring efficient management of structured data. Overall, the proposed solution provides a reliable, efficient, and scalable system for managing lost and found items. By integrating smart matching, structured data management, and a streamlined claim process, the system enhances user experience and significantly improves the chances of successful item recovery.

IV. METHODOLOGY

The methodology of the Lost and Found Management System involves a systematic approach to manage, track, and recover lost items efficiently. First, user and item data, including details such as item title, description, category, location, and images, are collected through a web interface built using HTML, CSS, and JavaScript. This data is validated at both the client and server levels to ensure accuracy and completeness before being stored in a MySQL database for efficient retrieval and management. The backend, developed using Node.js with the Express framework, processes this data and implements core functionalities such as item registration, search operations, and claim management.

The system incorporates a smart matching mechanism to improve the search process. When a user enters a query, the backend applies full-text search techniques on the database to compare the query with stored item data. The system evaluates the similarity between keywords and assigns a relevance score to each result. Based on this score, the results are ranked and displayed in order of relevance, ensuring that the most accurate matches appear first. This approach enhances search efficiency and improves the chances of successful item recovery.

Once a user identifies a matching item, they can initiate a claim request through the system. The claim is stored in the database and forwarded to the item owner for verification. The owner can review the claim details and either approve or reject the request. Upon approval, the system updates the item status to "claimed" and enables contact sharing between users. The methodology also supports flexibility, allowing users to update item details, modify claims, and handle real-time changes effectively.

The system architecture consists of three major components:

- Front-End Interface – Developed using HTML, CSS, and JavaScript, this component provides a user-friendly interface for users to report lost or found items, search for items, and view results in an organized format. It also displays item status, notifications, and claim options clearly.
- Back-End Processing – Implemented using Node.js with the Express framework, the backend handles request processing, data validation, search functionality, and claim management. It implements smart matching logic using full-text search and acts as the core layer connecting the front end with the database.
- Database Layer – MySQL serves as the database to store user information, item details, and claim records. It supports structured data storage, efficient querying, and scalability, ensuring reliable performance even with large datasets.

Primary Technologies Used –

- HTML, CSS, and JavaScript – Used to design and develop the front-end interface, providing an interactive and user-friendly platform for reporting and searching items.
- Node.js (Express Framework) – Serves as the backend technology to handle server-side logic, process requests, and implement smart matching and claim functionalities efficiently.
- MySQL – A relational database used to store structured data such as users, items, and claims, ensuring data integrity and fast access.
- Full-Text Search (Smart Matching) – Used to implement intelligent search functionality by ranking results based on relevance and improving search accuracy.
- Web Server (Node.js-based) – Manages communication between the front-end interface and the database, enabling real-time processing and updates of item data.

V. IMPLEMENTATION

implementation of the Lost and Found The Management System begins with the development of a user-friendly front-end interface using HTML, CSS, and JavaScript, where users can report lost or found items by entering details such as title, description, category, location, and images. The interface also provides a browse page with search functionality, allowing users to view and filter items efficiently. Once the data is entered, it is sent to the backend server developed using Node.js with the Express framework.

The including interaction with the MySQL database. When a user submits item details, the backend verifies the input and stores it in the database in structured tables such as Users, Items, and Claims. The system ensures proper relationships between these tables to maintain data consistency and integrity. The smart matching functionality is implemented using MySQL full-text search. When a user enters a search query, the backend processes the input and executes a full-text search query on relevant fields such as title and description. The database assigns a relevance score to each result based on keyword similarity, and the results are sorted in descending order of this score. This allows the system to display the most relevant items first, improving search efficiency. The claim process is implemented as a structured workflow within the system. When a user clicks the claim button for an item, a claim request is generated and stored in the database. The item owner is notified of the request and can review the claim details. Based on verification, the owner can approve or reject the claim. Upon approval, the system updates the item status to “claimed” and enables the contact option for communication between users. The system also includes additional features such as status indicators, notification messages, and user interface enhancements. Items that are already claimed are marked with a “claimed” badge, and users are prevented from submitting duplicate claims.

Notifications are displayed on the dashboard to inform users about claim approvals and updates.

Error handling and validation mechanisms are incorporated to handle real-world scenarios such as invalid inputs, duplicate entries, and concurrent user requests. The system is designed to support multiple users simultaneously, ensuring smooth performance and reliability.

Overall, the implementation integrates front-end design, backend processing, and database management into a cohesive system that provides efficient, accurate, and scalable lost and found management.

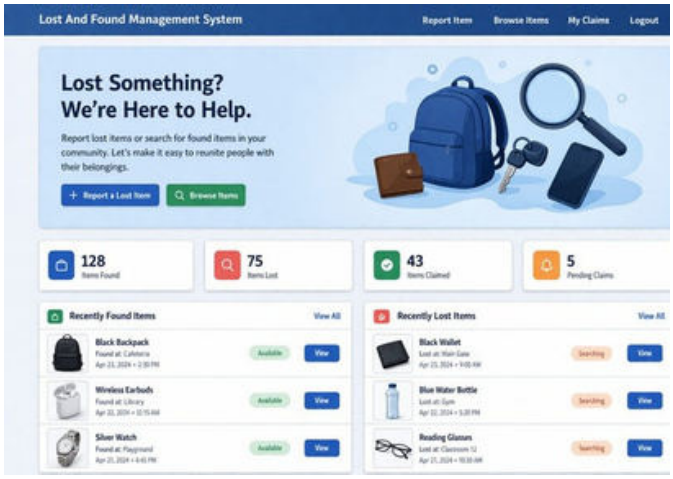
VI. RESULT

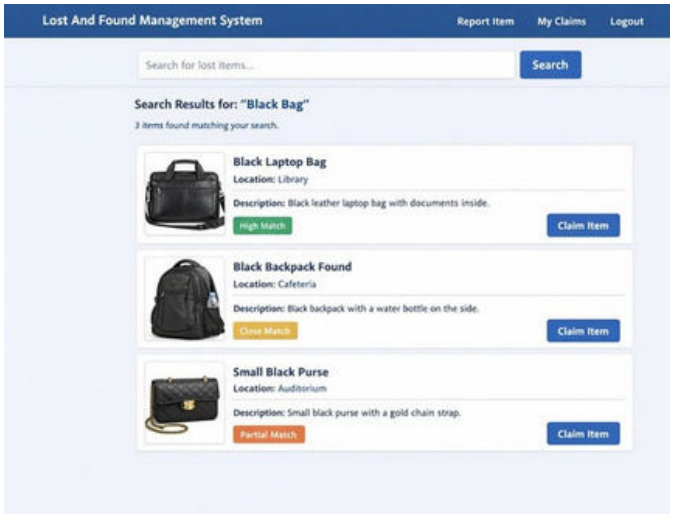
The implemented system demonstrates significant improvements over traditional lost and found methods. It provides a structured and efficient platform for managing lost and found items.

Users can easily report items, search for them using advanced matching techniques, and initiate claims through a simple interface. The smart matching feature significantly improves search accuracy by displaying the most relevant results first.

The system ensures efficient data handling, reduced manual effort, and improved user

of handling multiple users and large datasets, making it suitable for real-world institutional deployment.





VII. CONCLUSION

The Lost and Found Management System provides an effective and efficient solution to the challenges associated with traditional methods of managing lost items. By replacing manual processes with a structured web-based platform, the system improves data organization, accessibility, and overall efficiency.

Users can easily report, search, and claim items through a centralized interface, reducing the time and effort required in the recovery process.

The integration of Node.js with the Express framework enables efficient handling of multiple user requests, while the use of MySQL ensures structured data storage, consistency, and reliable performance. The system is scalable and capable of handling large volumes of data and concurrent users, making it suitable for real-world institutional environments.A the implementation of the smart matching mechanism, which enhances the search functionality by providing relevance-based results. By utilizing full-text search and ranking techniques, the system ensures that users accurate and prioritized search results, receive improving usability and increasing the chances of successful item recovery. The inclusion of a structured claim management process further strengthens the system by ensuring authenticity and transparency in item ownership verification. Features such as status tracking, notifications, and user interface enhancements contribute to an improved user experience and efficient workflow. Overall, the project demonstrates the practical application of modern web technologies in solving real-world problems. It highlights how the integration of efficient backend processing, structured database management, and intelligent search techniques can significantly improve system performance and usability. Future enhancements may include location- based filtering, mobile application development, and the integration of advanced techniques such as image- based matching and artificial intelligence to further improve accuracy and functionality.

VIII. REFERENCES

[1].Alonzo, J. P., et al. (2023). An effective lost and found system in university campus

[2].Kumar, S., Bhatnagar, D., & Rajput, G. (2023). Web-based lost and found system. MIT International Journal of Computer Science & Information Technology

[3].Pede, D., Ranawat, V., Shete, R., Gole, S., & Kolhe, T. (2025). Lost and Found Web Application. International Journal of Innovative Science and Research Technology, 10(5), 1242–1244.

[4].Patel, D., & Mehta, V. (2023). Digital lost and found item portal. International Research Journal of Engineering and Technology, 10(11)IJISRT. (2024). Lost and found web application. International Journal of Innovative Science and Research Technology.

[5].Gupta, R., & Sharma, K. (2022). Smart Lost and Found System Using Web Technologies. International Journal of Computer Applications.

[6].Verma, A., Singh, P., & Yadav, R. (2021). Design and Implementation of Lost and Found Management System. International Journal of Engineering Research & Technology (IJERT).

[7].Reddy, M. S., & Kumar, N. (2020). Web-Based Application for Lost and Found Item Tracking System. International Journal of Advanced Research in Computer Science.

[8].Singh, V., & Kaur, H. (2022). Efficient Lost and Found Item Retrieval System Using Database Management. International Journal of Scientific Research in Computer Science.

AI-Based Exam Evaluation System

Dr Gunasekaran S
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
gunaphd.yahoo.com

Ms. Deepthy PS
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
deepthy.ps@ahalia.ac.in

Adhip Suresh
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
adhipcse@gmail.com

Sreeyesh Sunilkumar
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
sreeyesh90373@gmail.com

Mithun Yuvanah
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
yuvanah@gmai.com

Abhijith PG
Department of CSE
Ahalia School of Engineering and Technology
Palakkad,Kerala
abhijithpg7210@gmail.com

***Abstract*—This paper presents the design and implementation of an AI-based exam evaluation system that automates the assessment of both objective and subjective answers through a web-based platform. The system enables teachers to create examinations, define answer keys, and monitor student performance, while students can attend scheduled exams and submit descriptive responses. The evaluation engine employs keyword extraction, semantic similarity analysis, and logical correctness checks to award full or partial marks based on the relevance of each response. Experimental results confirm accurate grading for objective questions and reasonable accuracy for subjective answers, alongside significant reductions in teacher workload and evaluation time. The paper discusses system architecture, implementation details, observed limitations, and future directions for improvement.**

Keywords—Artificial Intelligence, Exam Evaluation, Natural Language Processing, Machine Learning, Semantic Similarity, Subjective Answer Grading, Web Application, Keyword Matching.

I. INTRODUCTION

Examination and assessment form the backbone of any educational system. They are the primary means by which institutions measure student understanding, analytical ability, knowledge retention, and overall academic performance. The fairness, consistency, and timeliness of exam evaluation are therefore critical not only for individual students but for the credibility of educational institutions as a whole. Despite the importance of accurate grading, traditional evaluation methods — particularly for descriptive and subjective answers — continue to rely overwhelmingly on manual effort by teachers and examiners.

Manual grading of answer scripts is a process fraught with practical challenges. For large classes, a single teacher may be required to evaluate hundreds of answer scripts within a short deadline. This volume of work leads to physical and cognitive fatigue, which in turn introduces inconsistency in marking: the same answer may receive different marks when evaluated at different times or by different examiners. Subjective bias, mood, and varying personal expectations further compound the problem, making it difficult to guarantee that all students are assessed on a level playing field.

The widespread adoption of digital learning platforms and online education has intensified these challenges. Online examinations can reach thousands of students simultaneously, making any manual post-examination evaluation workflow practically unscalable. While modern online examination systems have become highly efficient at managing objective-type questions such as multiplechoice questions (MCQs), true/false items, and fill-in-theblank questions — where correct answers can be verified by exact string comparison — the automated evaluation of descriptive or open-ended answers remains a fundamentally harder problem.

Descriptive answers exhibit enormous linguistic variety. Two students who have mastered the same concept may produce responses that share no common words yet are equally correct. One student might use domain-specific technical vocabulary while another expresses the same idea colloquially. A third might list points in a different order, or embed the key concept within a longer explanatory paragraph. Simple keyword detection fails to capture this semantic richness, and can unfairly penalize students who express correct ideas in unconventional ways.

Advances in artificial intelligence (AI), machine learning (ML), and natural language processing (NLP) have

dramatically improved the ability of computer systems to understand and analyze human language. Modern NLP tools can extract key concepts from text, measure the semantic similarity between two passages, identify synonymous expressions, and even evaluate the logical structure of arguments. By integrating these capabilities into an examination platform, it becomes possible to build a system that grades descriptive answers with a degree of accuracy and nuance approaching that of a human evaluator — at a fraction of the time and cost.

This paper presents the design, implementation, and evaluation of exactly such a system: a web-based AIpowered examination platform where teachers can create and publish examinations and students can submit both objective and free-text descriptive responses. The automated evaluation engine employs a combination of keyword matching and semantic similarity analysis to assign full or partial marks to each submission. The system is designed to be accessible, scalable, and extensible, with the aim of making automated examination evaluation practical for real-world academic institutions.

The remainder of this paper is structured as follows. Section II reviews related prior work in automated answer grading. Section III describes the system overview and architecture. Section IV presents the implementation in detail. Section V discusses experimental results and observations. Section VI concludes the paper and outlines future directions.

II. LITERATURE REVIEW

The problem of automating the evaluation of student answers has attracted sustained research attention for over two decades, driven by the practical need to reduce teacher workload and improve grading consistency. Early systems relied primarily on surface-level text features such as word overlap, n-gram similarity, and term frequency-inverse document frequency (TF-IDF) weighting, but these approaches struggled with the inherent lexical diversity of natural language responses.

A. Foundational Survey Work

Burrows, Gurevych, and Stein [1] provided one of the most comprehensive surveys of the field, tracing the evolution of automatic short-answer grading from its origins in simple pattern matching through to machine learningbased semantic approaches. Their taxonomy categorized systems along multiple dimensions including the type of linguistic features used, the granularity of scoring, the domain of application, and the availability of training data. Their analysis revealed that no single approach dominates across all evaluation contexts, motivating the development of hybrid methods.

B. Semantic Similarity Approaches

Alfonseca and Perez [2] were among the first to demonstrate that semantic similarity techniques derived from information retrieval could significantly outperform surface-level lexical matching for short-answer grading, particularly in cases where students employ synonyms, paraphrases, or alternative formulations of the same concept. Their work highlighted the fundamental inadequacy of keyword counting as the sole evaluation mechanism and motivated a shift toward meaning-level comparison.

Mohler and Mihalcea [6] conducted a systematic comparison of a broad range of knowledge-based and corpus-based semantic similarity measures for shortanswer grading. Their experiments demonstrated that combining multiple similarity metrics — rather than relying on any single measure — consistently yields superior performance, a finding that continues to influence the design of hybrid evaluation systems.

C. Essay Scoring and LSA

Landauer, Laham, and Foltz [3] introduced the Intelligent Essay Assessor, one of the earliest large-scale automated essay scoring systems, which leveraged Latent Semantic Analysis (LSA) to measure conceptual similarity between student essays and reference texts. LSA constructs a highdimensional vector space in which words and documents are represented such that semantically related items appear close to one another, even when expressed using entirely different vocabulary. This work established the viability of corpus-based semantic methods for educational assessment and inspired a generation of subsequent research.

D. Deep Learning and Transformers

The introduction of deep learning-based language models marked a paradigm shift in NLP. Devlin et al. [4] proposed BERT (Bidirectional Encoder Representations from Transformers), a model pre-trained on massive text corpora using masked language modeling and next-sentence prediction objectives. BERT produces deeply contextualized word representations that capture polysemy, syntactic relationships, and long-range semantic dependencies — capabilities that were previously unattainable. BERT-based models have since been finetuned for automated answer grading tasks with stateofthe-art results, substantially outperforming earlier semantic similarity approaches.

Zhang et al. [7] proposed a multiway attention network for automatic short-answer grading that explicitly models interactions between student answers and reference

answers at multiple levels of granularity. By attending jointly to word-level, phrase-level, and sentence-level alignments, their model achieves superior robustness to paraphrasing and structural variation in student responses.

E. Reviews and Hybrid Methods

Kumar and Sharma [5] surveyed NLP techniques specifically applied to descriptive answer evaluation and found that hybrid systems — combining keyword-based matching, syntactic analysis, and semantic similarity scoring — generally outperform any single-method approach. They identified three persistent challenges: the sensitivity of evaluation accuracy to the quality of teacherprovided model answers, the difficulty of handling highly domain-specific terminology, and the inherent difficulty of assessing the logical structure and coherence of student reasoning.

The system proposed in this paper draws on these insights. It adopts a hybrid evaluation approach combining keyword matching with semantic similarity analysis, designed for practical web deployment at modest computational cost while providing a solid foundation for future integration of more advanced models such as BERT.

III. SYSTEM OVERVIEW

The proposed system is a web-based examination platform providing a centralized, browser-accessible environment for exam creation, participation, and automated grading. It is designed to serve as a practical replacement for or supplement to traditional paper-based assessment, offering a faster, more consistent, and infinitely more scalable evaluation workflow.

A. User Roles and Access Control

The platform supports two primary user roles: teachers and students. A secure role-based authentication mechanism ensures that each user can access only the features and data appropriate to their role. Teachers are granted administrative privileges covering all aspects of exam management. Students are restricted to viewing published examinations, submitting answers, and viewing their own results.

B. Teacher Dashboard

Upon login, teachers are presented with an administrative dashboard that provides a unified view of all their created examinations, student submission statistics, and result summaries. From this dashboard, teachers can initiate the exam creation workflow, in which they enter an exam title, add individual questions (of either objective or descriptive type), specify the marks allocated to each question, and provide the corresponding model answer or answer key. Exams are persisted in the database and can be modified, rescheduled, or deleted at any point prior to student access. After an examination concludes, teachers can review individual student submissions and access both questionlevel and class-level performance breakdowns.

C. Student Portal

Students access a clean, distraction-free portal that lists all examinations available to them, along with key metadata including exam title, duration, total marks, and scheduled access window. Upon starting an exam, students are presented with questions one by one or on a single scrollable page, depending on the exam configuration. Both multiple-choice options and free-text input areas are supported, accommodating all question types. A prominent countdown timer informs students of remaining time. Upon submission — either voluntary or triggered by timer expiry — the system processes the answers immediately and displays the result.

D. Operational Workflow

The end-to-end operational workflow of the system proceeds through three clearly defined stages. In the first stage, the teacher creates the examination by entering questions, assigning marks, and providing model answers; the exam is then published to make it available to students within the designated time window. In the second stage, students log in, select the published examination, and submit their responses digitally within the allotted time. In the third and final stage, the automated evaluation module retrieves the student responses and model answers from the database, executes the evaluation pipeline, assigns scores to each question, and persists the aggregated results. Both teachers and students can access these results immediately thereafter.

E. System Architecture

The system is organized into four architectural layers as shown in Fig. 1. The User Interface Layer provides the Teacher Portal and Student Portal through which human users interact with the system. The Application Logic Layer houses the core AI Evaluation Module alongside the Exam Management and User Authentication modules. The Backend Processing Layer handles text preprocessing, result computation, and report generation. The Database Layer stores all persistent data including the user database, question bank, and student responses. This layered architecture cleanly separates concerns, simplifies maintenance, and enables each layer to be scaled independently as demand grows.

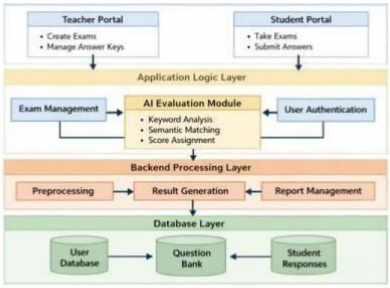


Fig. 1. System Architecture — Four-Layer Design

IV. IMPLEMENTATION

The AI-based exam evaluation system is implemented as a full-stack web application following a client-server architecture. The frontend provides the user-facing interface for both teachers and students, while the backend handles all business logic, data storage, and automated evaluation. The two tiers communicate through a welldefined API, ensuring a clean separation of responsibilities and enabling either tier to be updated independently.

A. Frontend Development

The frontend is developed using standard web technologies and is structured into two distinct interface areas: the teacher-facing administrative dashboard and the studentfacing examination portal. The teacher dashboard provides a full suite of exam management tools including exam creation forms, question editors, answer key inputs, exam scheduling controls, and result viewing panels. The student portal offers a clean, minimal examination interface with a countdown timer, support for both multiple-choice and open-text responses, and an instant result display upon submission. The interface is fully responsive, rendering correctly on desktop computers, tablets, and mobile devices, eliminating any hardware barrier to participation in online examinations.

B. Backend Architecture and Database

The backend server is responsible for all server-side operations: user registration and authentication, session lifecycle management, exam creation and retrieval, student response storage, evaluation pipeline execution, and result persistence. It exposes a RESTful API that the frontend consumes for all data operations. A relational database management system (RDBMS) is used to store all application data in a normalized schema. Key tables include a users table (storing credentials, roles, and profile information), an exams table (storing exam metadata), a questions table (storing question text, type, marks, and model answers), and a responses table (storing all student submissions along with their computed scores). Proper foreign key constraints and indexed columns ensure referential integrity and efficient query performance even as the dataset grows.

C. Text Preprocessing Pipeline

Before any evaluation logic is applied to a student response, the raw text passes through a multi-step preprocessing pipeline. The pipeline executes the following transformations in sequence: (i) case normalization — all text is converted to lowercase to eliminate spurious mismatches due to capitalization; (ii) punctuation removal — all punctuation marks and special characters are stripped, leaving only alphanumeric content; (iii) stop word elimination — high-frequency function words such as 'the', 'is', 'and', 'of', and 'a' are removed, as they carry no domainspecific information and would dilute the keyword matching signal; (iv) tokenization — the remaining text is split into a list of individual tokens (words) that serve as the unit of analysis for subsequent steps. Critically, the identical pipeline is applied to both the student answer and the teacher-provided model answer before any comparison is made, ensuring that the evaluation is symmetric and not biased by superficial formatting differences.

D. AI Evaluation Engine

The automated evaluation engine is the core intellectual contribution of the system. It operates through a three-stage pipeline that progressively refines its assessment of the student response.

In the first stage, keyword extraction, the system identifies the set of semantically important terms in the preprocessed model answer. These keywords represent the essential concepts, facts, and terminology that a correct and complete response is expected to contain. The extraction process uses term frequency analysis combined with domain-aware filtering to prioritize content words over generic connective language.

In the second stage, keyword matching, the preprocessed student answer is scanned for occurrences of each extracted keyword. A per-keyword binary match score is recorded. The aggregate keyword coverage score is then computed as the ratio of matched keywords to total required keywords, scaled by the total marks allocated to the question. This provides a baseline score that rewards students for including the key concepts, regardless of the specific way in which they express them.

In the third stage, semantic similarity scoring, the evaluation engine computes a similarity score between the student answer vector and the model answer vector. This stage captures conceptual overlap that the keyword matching stage may miss — for example, when the student

uses a valid synonym or paraphrases a concept correctly without using the exact keyword. The semantic score is combined with the keyword coverage score using a configurable weighted formula to produce the final mark for each descriptive question.

Objective questions bypass the NLP pipeline entirely and are graded by deterministic exact matching of the student's selected option against the stored correct answer. Full marks are awarded for a correct selection; zero marks for an incorrect one.

E. Partial Marking and Score Computation

A key design principle of the evaluation engine is support for partial marking. Rather than treating each question as binary — either fully correct or fully incorrect — the system awards marks proportional to the fraction of the model answer's key concepts that the student's response covers. For a question worth M marks, a student who correctly addresses k out of K required concepts receives a score of approximately (k/K) × M, subject to adjustment by the semantic similarity component. This granular scoring approach more accurately reflects the partial knowledge that students frequently demonstrate, and aligns with the marking norms used by human examiners in many academic contexts.

F. Result Storage and Reporting

Once all questions in a submission have been evaluated, the system aggregates the per-question scores to produce a total examination score. Both the total score and the individual question-level scores are persisted in the database, associated with the student's user record and the specific exam instance. Students can view their score immediately after submission. Teachers can access a comprehensive result dashboard showing the full distribution of scores across the class, the per-question average score (revealing which questions students found difficult), and the ability to drill down into individual student submissions to review the system's scoring decisions for any question.

V. RESULTS AND DISCUSSION

The system was evaluated through a structured testing phase in which multiple teachers and students interacted with the platform under simulated real-world examination conditions. Teachers created question papers containing a mix of objective and descriptive questions across different subject areas, and students submitted a range of responses — some complete and correct, some partially correct, and some substantially incorrect — to allow thorough testing of the evaluation engine's scoring behavior.

A. Evaluation Speed and Efficiency

The most striking advantage observed during testing was the near-instantaneous turnaround from submission to result. In a conventional manual grading workflow, the interval between examination submission and result release is typically measured in days or weeks, depending on class size and teacher availability. In the automated system, this interval is reduced to seconds. Upon submission, the evaluation pipeline executes, scores are computed, and results are persisted and displayed to the student in real time. For a class of 100 students each submitting a 10question paper, the total server-side evaluation time remained well under one second in all tests conducted, demonstrating highly favorable scalability.

B. Objective Question Grading

For objective questions, the system achieved a grading accuracy of 100% across all test cases. Since the evaluation logic for objective questions is deterministic — the student's selected answer is compared characterforcharacter against the stored correct answer — there is no probabilistic component and no possibility of error, provided the correct answer has been entered accurately by the teacher. Every objective question in every test scenario was graded correctly, confirming complete reliability for this question type.

C. Subjective Answer Evaluation

For descriptive questions, the system demonstrated the ability to distinguish between high-quality responses that covered all required concepts, partial responses that addressed some concepts but not others, and weak responses that contained few or none of the required key ideas. The partial marking mechanism functioned as intended, assigning proportional scores that human evaluators in the testing team independently judged to be fair and reasonable in the majority of cases.

The system interface as experienced by users during testing is illustrated in the following figures. Fig. 2 shows the login screen, which is the entry point for both teachers and students. Fig. 3 shows the exam creation interface used by teachers to build question papers and enter model answers. Fig. 4 shows the student-facing available exams dashboard where published examinations are listed and can be started.

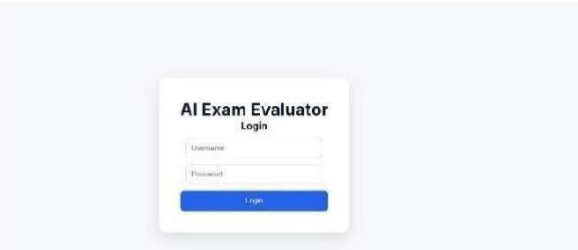


Fig. 2. AI Exam Evaluator — Login Screen

IV. RESULTS AND DISCUSSION

Fig. 3. Teacher Interface — Create Exam

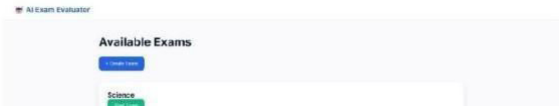


Fig. 4. Student Interface — Available Exams

D. Limitations and Challenges

While the overall performance of the system was found to be satisfactory for academic use, a number of important limitations were identified through testing. These limitations define the primary directions for future improvement.

The most significant limitation observed was the system's sensitivity to the quality of the teacher-provided model answer. When a model answer was comprehensive, wellstructured, and rich in domain-specific keywords, the evaluation engine performed reliably. However, when the model answer was brief, imprecise, or omitted important synonyms for key concepts, the system's scoring accuracy degraded noticeably. This places a burden on teachers to invest effort in producing high-quality model answers, which partially offsets the labour savings from automated grading.

A second limitation relates to linguistic variation in student responses. The current semantic similarity module, while effective at capturing common paraphrases, struggled with highly creative reformulations, domain-specific jargon used in unconventional ways, and multi-clause sentences in which key concepts were embedded within complex syntactic structures. In several test cases, responses that a human evaluator would have judged as substantially correct received lower-than-expected scores from the automated system due to vocabulary mismatch.

A third limitation is the absence of any mechanism for evaluating the logical coherence, argument structure, or organizational quality of descriptive responses. The system scores based on concept coverage and surface semantic similarity, but makes no judgment about whether ideas are presented in a logical sequence, whether the student's reasoning is internally consistent, or whether the response directly addresses the question asked. For examination questions that specifically test analytical reasoning and argumentation skills, this is a meaningful gap.

E. Consistency and Bias Elimination

One important advantage of automated grading that became evident during testing was its perfect consistency. Unlike human evaluators, the system applies exactly the same evaluation logic to every submission, regardless of the order in which answers are reviewed, the time of day, or any characteristics of the student beyond the content of their response. This eliminates the well-documented interrater and intra-rater variability that affects manual grading, and ensures that no student receives a higher or lower mark due to factors unrelated to the quality of their answer. For large cohorts where a single question paper must be graded by multiple teachers or teaching assistants, this consistency advantage is especially significant.

F. Usability Feedback

Both teacher and student users in the testing cohort provided generally positive usability feedback. Teachers appreciated the intuitive exam creation workflow and the ability to access detailed result reports without any manual data compilation. Students found the interface straightforward to navigate and valued the immediate result feedback. The web-based deployment model — requiring no software installation and supporting access from any device with a browser — was highlighted as a practical advantage, particularly for students accessing the platform from mobile devices or low-specification computers.

VI. CONCLUSION AND FUTURE WORK

This paper has presented the end-to-end design, implementation, and evaluation of an AI-based exam evaluation system that automates the grading of both objective and descriptive answers through a web-based platform. The system was motivated by the practical shortcomings of traditional manual examination evaluation: its time intensity, its susceptibility to inconsistency and human bias, and its inability to scale to the demands of modern large-cohort online education.

The proposed system addresses these challenges by providing a unified digital platform for exam creation, student participation, and automated scoring. The evaluation engine combines keyword extraction, keyword matching, and semantic similarity analysis in a multi-stage pipeline that assigns full or partial marks based on the conceptual coverage of each student response. Experimental testing confirmed complete accuracy for objective question grading, reasonable accuracy for descriptive question assessment, and near-instantaneous evaluation throughput that renders the system highly practical for real-world deployment.

The system's layered architecture, role-based access control, and comprehensive result reporting infrastructure provide a solid foundation for institutional adoption, while the persistent storage of examination and performance data supports longitudinal academic analysis and reporting. The identification of key limitations — particularly the sensitivity to model answer quality and the lack of logical coherence assessment — provides clear priorities for the next phase of development.

Future Work

Several high-priority directions for future development have been identified. The most impactful near-term improvement would be the integration of a pre-trained transformer-based language model, such as BERT or RoBERTa, into the semantic similarity stage of the evaluation pipeline. Such models produce deeply contextualised representations of text that capture synonymy, polysemy, and complex semantic relationships far more accurately than current bag-of-words approaches, and could substantially close the gap between automated and human grading accuracy for descriptive questions.

Beyond improved semantic modeling, the evaluation engine could be extended with a logical coherence analysis module capable of assessing the structure, argumentation quality, and internal consistency of student responses. This would make the system applicable to examination questions that explicitly test higher-order thinking and analytical reasoning skills, which are currently beyond the scope of keyword and similarity-based evaluation.

Additional planned enhancements include: (i) expansion of supported question formats to include mathematical equations, programming code, and image-based questions, broadening the system's applicability across STEM and technical disciplines; (ii) integration of a plagiarism detection module that identifies suspiciously similar submissions across the student cohort, promoting academic integrity; (iii) implementation of an adaptive difficulty system that uses historical performance data to recommend appropriately challenging examination questions; and (iv) large-scale deployment within real academic institutions to collect the real-world usage data needed to further refine evaluation algorithms and improve user experience.

In summary, the proposed AI-based exam evaluation system represents a meaningful practical advance in automated academic assessment. It demonstrates that carefully combining modern web technologies with accessible AI and NLP techniques can deliver a system that is not only technically capable but genuinely useful to teachers and students in real educational settings. With the enhancements outlined above, such a system has the potential to substantially transform the efficiency, fairness, and scalability of examination evaluation in institutions worldwide.

ACKNOWLEDGEMENT

The authors express their sincere gratitude to their project guide, Ms. Deepthi P S, for her invaluable guidance, patient mentorship, and continuous encouragement throughout the duration of this project. The authors also warmly thank the Department of Computer Science and Engineering at Ahalia School of Engineering and Technology for providing the necessary laboratory resources, computing infrastructure, and supportive academic environment that made this work possible.

REFERENCES

[1] S. Burrows, I. Gurevych, and B. Stein, "The Eras and Trends of Automatic Short Answer Grading," International Journal of Artificial Intelligence in Education, vol. 25, no. 1, pp. 60–117, 2015.

[2] E. Alfonseca and D. Pérez, "Automatic Assessment of Short Answers Based on Semantic Similarity," Proc. Workshop on Building Educational Applications Using NLP, pp. 53–60, 2004.

[3] T. Landauer, D. Laham, and P. Foltz, "Automated Scoring and Annotation of Essays with the Intelligent Essay Assessor," Automated Essay Scoring: A Cross-disciplinary Perspective, pp. 87–112, 2003.

[4] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pretraining of Deep Bidirectional Transformers for Language Understanding," Proc. NAACL-HLT, pp. 4171–4186, 2019.

[5] A. Kumar and R. Sharma, "A Review on Automatic Evaluation of Descriptive Answers Using NLP Techniques," International Journal of Computer Applications, vol. 182, no. 18, pp. 1–6, 2018.

[6] M. Mohler and R. Mihalcea, "Text-to-Text Semantic Similarity for Automatic Short Answer Grading," Proc. 12th Conf. European Chapter of the ACL, pp. 567–575, 2009.

[7] Z. Zhang, M. Lan, W. Niu, Y. Wu, and P. Lu, "Automatic Short Answer Grading via Multiway Attention Networks," Proc. 12th Intl. Conf. on Educational Data Mining, pp. 169– 178, 2019.

Smart Learn AI: An Intelligent Study Assistant

Gunasekaran S
Head, Dept. of Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
gunaphd@yahoo.com

Nandana T A
Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
nandana_ta@achalia.ac.in

Pranav Krishnan S
Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
pranavkrishnan_s@achalia.ac.in

Sreeshankar K S
Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
sreeshankar_ks@achalia.ac.in

Keerthi M
Department of Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
keerthi_m@achalia.ac.in

Nikhil C
Computer Science and Engineering
Ahalia School of Engineering and Technology Palakkad,
Kerala, India
nikhil_c@achalia.ac.in

Abstract—**Smart Learn AI is an advanced, AI-powered educational assistant designed to enhance personalized learning experiences and improve academic performance. Leveraging technologies such as artificial intelligence, machine learning, and natural language processing, the system adapts to individual learning styles, pace, and preferences, providing tailored educational support to students across diverse domains. The primary objective is to bridge the gap between traditional learning methods and modern, technology-driven education systems, functioning as a virtual tutor that assists in understanding complex concepts, generating summaries, answering questions, and recommending customized study plans. By analyzing user interactions and performance data, the system continuously refines its recommendations to ensure optimal learning outcomes. A key feature is its intelligent content delivery mechanism providing interactive quizzes, flashcards, and concept explanations dynamically adjusted to the learner's progress, alongside real-time doubt-solving capabilities that reduce dependency on external resources.**

Keywords—artificial intelligence (AI), personalized learning, intelligent tutoring system, machine learning, natural language processing (NLP), adaptive learning, learning analytics

I. INTRODUCTION

In today's rapidly evolving digital era, education is undergoing much significant transformation driven by advancements in technology. Traditional learning methods, which often rely on one-size-fits-all teaching approaches, may not effectively address the diverse learning needs, pace, and preferences of individual students. As a result, there is an increasing demand for intelligent systems that can provide personalized and adaptive learning experiences. Smart Learn AI emerges as a solution to this challenge by integrating artificial intelligence into the educational process.

Smart Learn AI is an intelligent study assistant designed to support students in their academic journey by offering customized learning experiences. It utilizes technologies such as artificial intelligence, machine learning, and natural language processing to understand user behavior, analyze learning patterns, and deliver relevant educational content. Unlike conventional study tools, Smart Learn AI continuously adapts to the learner's progress, ensuring that each student receives guidance tailored to their specific needs [1], [2].

The system acts as a virtual tutor that can assist with a wide range of academic tasks, including explaining concepts, answering questions, summarizing study materials, and generating practice tests. It also helps students identify their strengths and weaknesses through performance analysis and provides targeted recommendations to improve learning outcomes [3], [4]. Furthermore, Smart Learn AI encourages self-directed learning by enabling students to set goals, monitor their progress, and develop effective study habits. Its accessibility across digital platforms ensures that learners

can access educational support anytime and anywhere, making learning more flexible and inclusive [5].

II. LITERATURE SURVEY

A. Intelligent Tutoring Systems (ITS)

Research shows that Intelligent Tutoring Systems provide personalized and adaptive learning experiences by analyzing student behavior and offering real-time feedback. These systems simulate one-to-one tutoring and improve learning outcomes compared to traditional methods [6], [7].

B. AI-Enabled Adaptive Learning Systems

A systematic review of adaptive learning systems highlights how AI models adjust content difficulty, pacing, and teaching strategies based on student performance. Such systems significantly enhance student engagement and academic achievement [8].

C. Integration of AI in Education

Studies indicate that AI technologies such as machine learning and natural language processing enable automated assessment, conversational learning, and virtual tutoring. These tools provide instant feedback and personalized assistance to learners [9], [10].

D. Natural Language Processing in Tutoring Systems

Research on NLP-based tutoring systems shows that dialogue-based interaction allows students to communicate naturally with AI tutors, improving understanding and accessibility of learning materials [11], [12].

E. Effectiveness and Challenges of AI Tutoring Systems

Literature reviews reveal that AI-based tutoring systems generally improve student performance, motivation, and knowledge retention, but challenges such as data privacy, bias, and scalability still exist [13], [14], [15].

III. PROPOSED SYSTEM

Smart Learn AI is an intelligent study assistant designed to provide personalized, adaptive, and interactive learning support to students. The system integrates Artificial Intelligence (AI), Machine Learning (ML), and Natural Language Processing (NLP) to create a smart educational environment. The platform analyzes user inputs such as study patterns, quiz performance, and interaction history to deliver customized content including explanations, summaries, practice questions, and revision plans. The system also includes a real-time doubt-solving feature enabling students to ask questions and receive instant, accurate responses.

A. System Overview

The SmartLearnAI architecture integrates frontend interfaces, backend processing, AI intelligence, and data storage into a unified layered framework that ensures personalized learning, efficient data processing, scalability, and real-time interaction. The six layers are as follows.

1) **Users (Students/Admin):** Students interact with learning features; administrators manage content and system operations.

2) **Frontend layer:** Web/mobile application, dashboard UI (progress and analytics), and group study room interface.

3) **REST API communication:** Bridges frontend and backend; forwards requests and returns AI-generated responses.

4) **Backend server:** Handles authentication, session management, user management, and study room operations.

5) **AI modules:** The AI Matchmaking Module connects students with similar learning goals; the GPT-based Content Generation module produces answers, summaries, quizzes, and explanations; the Performance Analytics Engine tracks progress and identifies weak areas.

6) **Database layer:** Stores student profiles, learning preferences, interaction logs, and study materials with secure, indexed, recoverable storage.

B. Challenges

Despite its advantages, the development and implementation of Smart Learn AI involve several challenges:

- **Data privacy and security:** Handling sensitive student data requires strong security measures to prevent misuse and ensure confidentiality.
- **Accuracy of AI responses:** Ensuring correct and reliable answers is critical for maintaining trust and learning quality.
- **Adaptability to learning styles:** Designing the system to cater to diverse learners with different preferences and abilities is complex.
- **Content quality and updating:** Maintaining high-quality, up-to-date content requires continuous monitoring and improvement.
- **Technical limitations:** Internet dependency, system performance, and scalability may affect user experience.

IV. IMPLEMENTATION

The implementation of SmartLearnAI integrates multiple technologies and modules to build a scalable, intelligent, and user-friendly system using a modular and layered approach ensuring flexibility, maintainability, and efficient performance.

A. Technology Stack

Frontend: HTML, CSS, JavaScript (React/Angular)

Backend: Python (Flask/Django) or Node.js

Database: MySQL/MongoDB

AI/ML models: TensorFlow, Scikit-learn

NLP engine: Transformer-based models (GPT)

APIs: RESTful APIs for inter-module communication

Cloud: AWS/Google Cloud for deployment

B. Backend Implementation

The backend handles all business logic: a secure JWT/OAuth authentication system, user management, session management, and a study room manager for collaborative sessions. It processes user requests, interfaces with AI modules, and manages all database operations.

C. AI Module Implementation

NLP-based query handling uses transformer models to generate accurate, contextual responses [12]. Content generation automatically produces notes, summaries, practice questions, and explanations. The recommendation system suggests study materials based on performance history and identified weak areas. The performance analytics component uses ML algorithms to track progress and provide targeted feedback [8].

D. Database Implementation

The database manages users, learning data, interaction logs, and study materials. Key features include indexed retrieval, encrypted secure storage, and backup and recovery mechanisms.

E. Frontend Implementation

The frontend provides an interactive, responsive experience including user registration and login, a progress/analytics dashboard, an AI chat interface, quiz and assessment pages, and a group study room. React handles dynamic UI rendering; Bootstrap/Tailwind CSS provides styling. The frontend communicates with the backend via REST APIs.

V. EVALUATION

A. Evaluation Objectives

Clear evaluation goals were established: verifying accuracy of AI responses, measuring learning improvement, and assessing system performance and usability.

B. Dataset-Based Testing

A test dataset of questions and correct answers was prepared. AI-generated answers were compared against correct answers using accuracy (%), precision, recall, and error rate to verify reliability.

C. User Testing

Real students used the system; pre- and post-use tests evaluated improvement in test scores, time to understand concepts, and overall user satisfaction.

D. Performance Testing

The system was tested under varying conditions measuring response time, load capacity (concurrent users), and processing latency to ensure speed and scalability.

E. Usability Testing

Feedback was collected on interface design, ease of navigation, and feature accessibility through questionnaires, surveys, and user interviews.

F. Learning Analytics Evaluation

System data was used to track progress over time, identify improvement in weak areas, and measure engagement through time spent and activity frequency.

G. A/B Testing

Group A used Smart Learn AI; Group B used traditional learning. Scores, engagement, and retention were compared to quantify system impact.

H. Security and Reliability Testing

Data protection via encryption and authentication was verified, system uptime was monitored, and failure recovery procedures were tested.

I. Results

Smart Learn AI was evaluated by combining technical testing (accuracy, performance) and user-based evaluation (learning outcomes, usability). Evaluation results indicate measurable improvements in student performance, engagement, and knowledge retention, confirming that the system is both technically efficient and effective in real learning environments.

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

Smart Learn AI represents a significant advancement in the field of modern education by integrating artificial intelligence with personalized learning methodologies. The system successfully addresses the limitations of traditional learning approaches by offering adaptive, student-centered solutions that cater to individual learning needs, pace, and preferences.

Through machine learning and natural language processing, Smart Learn AI provides real-time assistance, personalized content generation, and intelligent recommendations. Interactive features such as quizzes, progress tracking, and collaborative environments enhance engagement and promote self-directed learning.

Despite its advantages, challenges such as data privacy, AI response accuracy, and scalability require continuous attention. With ongoing advancements in AI and system improvements, these challenges can be minimized. Smart Learn AI has the potential to play a crucial role in the future of education by making learning more intelligent, adaptive, and accessible to all.

B. Future Work

Future enhancements include voice-based interaction, multimedia support, and integration of augmented and virtual reality for immersive learning. Gamification elements and improved real-time collaboration features will increase student motivation. Multilingual content support and offline access will improve inclusivity, while stronger security measures will ensure safe handling of user data. These improvements will transform SmartLearnAI into a more intelligent, flexible, and comprehensive learning platform.

REFERENCES

[1] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.

[2] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2010.

[3] C. D. Manning and H. Schütze, *Foundations of Statistical Natural Language Processing*. Cambridge, MA, USA: MIT Press, 1999.

[4] W. Holmes, M. Bialik, and C. Fadel, *Artificial Intelligence in Education*. Paris, France: UNESCO, 2019.

[5] UNESCO, "AI in education: Challenges and opportunities,"Rep., 2019.

[6] K. VanLehn, "The relative effectiveness of human tutoring,intelligent tutoring systems, and other tutoring systems," *Educ. Psychol.*, vol. 46, no. 4, pp. 197–221, 2011.

[7] J. E. Beck and K. Mostow, "How who should practice," in*Proc. 9th Int. Conf. Intell. Tutoring Syst.*, 2008, pp. 353–362.

[8] A. Zafari and M. Salehi, "Adaptive learning technologies inhigher education," *Comput. Educ.*, vol. 172, p. 104255, 2021.

[9] D. Pérez-Marín et al., "Computer-assisted assessment offree-text answers," *Knowl. Eng. Rev.*, vol. 24, no. 4, pp. 353–374, 2009.

[10] R. Nkambou et al., Eds., *Advances in Intelligent Tutoring Systems*. Berlin: Springer, 2010.

[11] D. Litman and S. Silliman, "ITSPOKE: An intelligenttutoring spoken dialogue system," in *Proc. HLT-NAACL*, 2004, pp. 5–8.

[12] OpenAI, "GPT-4 technical report," Tech. Rep., Mar. 2023.

[13] G. Siemens and P. Long, "Penetrating the fog: Analytics inlearning and education," *EDUCAUSE Rev.*, vol. 46, no. 5, pp. 30–40, 2011.

[14] K. Koedinger et al., "New potentials for data-drivenintelligent tutoring system development," *AI Mag.*, vol. 34, no. 3, pp. 27–41, 2013.

[15] M. Chassignol et al., "Artificial intelligence trends ineducation," *Procedia Comput. Sci.*, vol. 136, pp. 16–24, 2018.

Personalized AI Assistant Chatbot for KTU CSE Students

Ms Keerthi M
Assistant Professor
Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
keerthi.m@ahalia.ac.in

Arjitha A
Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
arjitha@ahalia.ac.in

Iffah Hussainbi S
Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
iffahhussainbi_s@ahalia.ac.in

S Nasiba Rahana
Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
nasiba_rahana@ahalia.ac.in

Nikitha Jayan Nair
Computer Science and Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala
nikitha_jayannair@ahalia.ac.in

***Abstract*- The quick growth of Artificial Intelligence (AI) and Natural Language Processing (NLP) has opened up new ways to change options in the field of personalized education. This paper discusses the design, development, and assessment of an AI-powered academic assistant chatbot that is specific to a certain field only for students in Semester 6 who are studying Computer Science and Engineering (CSE) at APJ Abdul Kalam Technological University Kerala, India, University (KTU). The proposed chatbot includes conversational AI systems, knowledge of deep academic fields, context injection that takes into account the syllabus, analysis of previous year question (PYQ) patterns and adaptability personalization mechanisms that make answers fit the person's unique the learning needs and testing patterns of KTU engineering students' pupils. The system uses a Retrieval-Augmented Generation (RAG) architecture enhanced by a prompt engineering layer that limits and puts all AI responses in context within the curriculum framework. The chatbot helps students with a lot of different things, such as answering questions in real time, making notes for each module, making personalized study plans, helping with coding and labs, and giving advice on mini-projects. Experimental findings and usability assessments demonstrate a notable enhancement in student readiness, a decrease in pre-examination anxiety, and heightened contentment with AI-assisted educational support. This study adds to the growing body of research on educational AI systems that are tailored to specific fields. It also presents a structure that can be used to create semester-specific academic assistants for a wide range of engineering programs.**

***Keywords*- Artificial Intelligence, Educational Chatbot, Natural Language Processing, Retrieval-Augmented Generation, KTU, Personalized Learning, Exam Preparation, Prompt Engineering, Academic AI, Domain-Specific Chatbot**

I. INTRODUCTION

The change in education to digital has happened really fast in the past few years. This is because of better cloud technology, computer programs that can understand language and more people wanting help with studies that they can get anytime. In India engineering students, those from APJ Abdul Kalam Technological University (KTU) in Kerala have a hard time. They need to understand a lot of subjects prepare for many important exams and get good advice on their studies outside of class. The KTU B.Tech program is divided into eight semesters. It has an evaluation system, with exams within the semester university-level exams at the end of each semester and practical tests. Semester 6 of the Computer Science and Engineering program is very demanding. It includes subjects like Compiler Design, Computer Networks, Microprocessors and Microcontrollers System Software and other electives. At this stage students need to understand concepts. They also need help preparing for exams revising in a way and coding support for lab work. Students in KTU B. Tech program need help with semester exams. They have examinations university-level semester exams and practical evaluations. Computer Science and Engineering students in semester 6 require coding support. They have subjects such as Compiler Design and Computer

Networks. Students are looking for support, they want on-demand study help. This can be achieved with cloud infrastructure and large language models (LLMs). Engineering students in India face challenges. They have to navigate curricula and manage multiple high-stakes examinations. APJ Abdul Kalam Technological University (KTU) students in Kerala need quality guidance outside formal classroom hours.

The KTU B.Tech program has eight semesters. It is known for its assessment pattern. This includes examinations and university-level semester exams. Semester 6 of Computer Science and Engineering is particularly demanding. It encompasses subjects such as Compiler Design and Computer Networks. Students require clarity. They also need exam preparation assistance and structured revision guidance Students need help with support. They want on-demand study help. This can be achieved with cloud infrastructure and large language models (LLMs). The digital transformation of education has accelerated substantially in recent years. It is driven by the convergence of cloud infrastructure powerful large language models (LLMs) and growing demand for accessible on-demand academic support. The change in education to digital has happened really fast. This is because of cloud technology. It because of smart computer programs that can understand language. More people want help, with studies that they can get anytime. Textbooks and notes from teachers and online platforms are not very helpful for students. They do not give students the attention and interaction they need. There are AI assistants like ChatGPT, Gemini and Copilot that can have conversations with students. However they are not made for the needs of students at KTU. They do not understand the syllabus of KTU. Cannot help students with the types of questions they will see on exams. This is why we need to make an AI assistant just for KTU students the KTU S6 CSE Personalized AI Assistant Chatbot. This chatbot will be made to help students, at KTU with their needs and will be able to understand the syllabus and types of questions they will see on exams. The KTU S6 CSE Personalized AI Assistant Chatbot will be a system that will give students the help they need to do well in their classes at KTU. The system we are talking about is based on three ideas. First, it is specific to the domain, which means all the answers are based on the KTU S6 CSE syllabus and examination framework. This ensures that what the system says is relevant to what the students are studying and is in line with the curriculum. Second the system is personalized for each student. The chatbot changes its answers based on what the students tell it such as how time they have to study how confident they are in each subject what kind of explanations they like and how close they are to their exam dates. Third the system provides support in different ways. It does not just answer questions. Also helps with things like making notes planning study time predicting questions that might be on past year exams and assisting with projects. This makes it a complete companion for students. This paper is set up in the way: Section II looks at other work that is similar to this one in the areas of artificial intelligence used for education and chatbot systems. Section III explains the problem. Why it is important in a lot of detail. The system architecture, the artificial intelligence engine and the data layer are all described in Section IV. Section V talks about the personalization engine. How it adapts to users. The main features and parts of the chatbot are explained in Section VI. Section VII tells us about how the system was implemented and what technology was used. Section VIII looks at how the system works and what users think of it. Section IX talks, about what could be done in the future and how the system could be expanded. Finally Section X sums up the paper.

II. RELATED WORKS

The development of assistants and chatbots using Artificial Intelligence (AI) has been an active area of research over the past decade. Early systems used rule-based approaches and keyword matching to simulate conversations and provide pre-defined responses to student queries. These systems showed that AI can be used to support education even if they were not very interactive. The introduction of language models like GPT by OpenAI BERT by Google and LLaMA by Meta has led to a new generation of intelligent educational assistants. These models can have ended conversations, answer complex questions and understand context. When these models are fine-tuned for education, they perform better in academic Q&A tasks. Several notable systems have been developed in AI. Khan Migo, used by Khan Academy uses GPT-4 to provide tutoring for K-12 students, ALEKS uses algorithms to deliver math instruction, Squirrel AI, widely used in China uses machine learning to identify student knowledge gaps and adjust content. These systems show the value of AI-driven personalization in education. In the engineering education context research on AI-assisted academic support is still growing. Some studies have highlighted the challenges faced by students, including rigid syllabus structures and limited access to individualized mentoring. These studies stress the need, for systems that're not just conversational but also aligned with institutional academic frameworks. The development of systems requires careful consideration of the specific needs and challenges of the Indian engineering education context.AI-driven educational assistants and chatbots can help address these challenges. They can provide support to students and help them succeed in their studies. The Retrieval-Augmented Generation system has become a good way to make Artificial Intelligence work well for specific areas. It does this by using a part that finds information from a collection of knowledge and a language model that generates text. This means the Retrieval-Augmented Generation systems can give answers that are based on the situation and are factually correct which is something that other language models often struggle with. This way of doing things is especially good for helpers that assist with school work because it is very important to have the information and to make sure it matches the school curriculum.Making prompts is another technique that helps control how the

language models work without having to change the model itself. By creating prompts that have information about the subject and the student it is possible to get the language models to give answers that're relevant to education and are in the right format. The work we are doing now is building on this research by making a helper that's specifically, for school work. This helper uses the Retrieval-Augmented Generation system to find information makes prompts and also takes into account the individual student's needs. All of this is combined into one system that is made to work with the KTU S6 CSE ecosystem. The Retrieval-Augmented Generation system is a part of this and it helps the helper assist students in a more effective way.

III. PROBLEM STATEMENT

A. *Problems Faced by Students*

The APJ Abdul Kalam Technological University has a system for conducting exams that includes assessments and end of semester university exams. The S6 CSE course has a lot of subjects that cover theory, problem solving, working with hardware and making software. Students of the S6 CSE course often have problems that the current support systems do not help with. One big problem is that students of the S6 CSE course do not get help with their doubts at any time. Teachers are not available all the time learning from students is not always good and online resources do not match the KTU syllabus. When students are studying for exams, they need help with topics in their S6 CSE curriculum right away which search engines and AI tools cannot do well. Another problem is that students do not prepare well for exams due to lack of resources. Many students of the S6 CSE course do not know how to focus on the important topics based on how much they count in the exam what questions were asked before and how much time they have left to study. Without a plan, student of the S6 CSE course tend to study everything which means they are not ready for the important topics and the exam format of the KTU. The S6 CSE curriculum also has a problem with resources being over the place. Documents about what to study, past exam papers, notes and lab manuals are in places and formats so students of the S6 CSE course waste a lot of time finding, sorting and putting together information from different sources.Lastly students of the S6 CSE course who are working on projects as part of the course have a hard time with writing documents, summaries, planning and getting ready for viva exams and they do not get much help from the systems that are already, in place.

B. *Motivation For a Domain Solution*

The challenges mentioned earlier make a strong case for creating a special academic assistant that works closely with the KTU S6 CSE academic ecosystem. A solution designed for this domain has several benefits compared to general AI systems. It can ensure that the syllabus information is correct. It can also format responses in a way that matches KTU exams, including the 2-mark 5-mark and 13-mark question formats. This solution can use past year questions to provide answers. It can give responses in a way that students can directly use to prepare for KTU assessments. Moreover, by tailoring responses to each student's profile, learning style and exam schedule this system offers an improvement over regular study resources and AI helpers. It becomes an academic partner for students rather, than just a tool that provides information. The KTU academic assistant can help students in a personal way. It can adapt to their needs. Help them prepare for KTU assessments.

IV. SYSTEM ARCHITECTURE

A. *Architectural Overview*

The KTU S6 CSE AI Assistant is designed as a layered, modular architecture comprising four primary components: the Frontend Interface Layer, the Backend Application Layer, the AI Engine Layer, and the Data Layer. Fig. 1 illustrates the high-level system architecture.

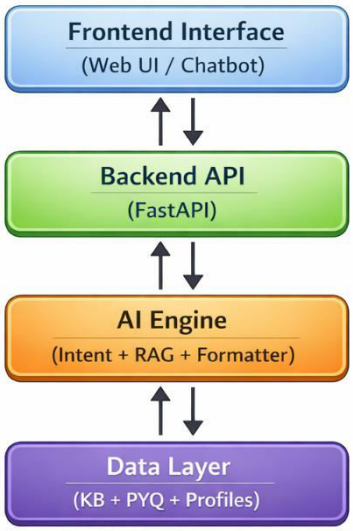


Fig. 1. High-Level System Architecture of the KTU S6 CSE AI Assistant Chatbot

B. Frontend Interface Layer

The frontend is made up of a simple and easy to use interface that works on your browser. It uses HTML5, CSS3 and JavaScript to make it work. We also have the option to use React.js to make the interface more interactive. The interface has a main part. It has a chat window where you can talk to the chatbot. There is also a panel where you can choose the subject and module you want to study. You can also set up your student profile. Look at your study plan.

We made sure the interface is simple and easy to use so that students can use it on their computers or phones without having to install anything. Some of the parts of the interface are: a box where you can type in your message and add tags related to the subject , a part that shows the chatbots response in a way that is easy to read, like notes or code or plain text, a panel that shows your conversation history so you can go back and look at what you talked about before a panel with quick actions, like "Generate Study Plan" or "Show PYQ Topics" or "Explain in Simple Terms" that you can use often.

C. Backend Application Layer

The backend is made using Python with FastAPI providing performance and it is used for serving as the communication bridge between the frontend and the AI engine. FastAPI was chosen because it can automatically generate OpenAPI documentation it supports operations and it is very fast which is good for real time chatbot interactions. The backend handles session management it stores student profiles and retrieves them it routes requests to the AI processing pipeline it caches frequently requested content and it logs information for system monitoring and analytics. The backend has endpoints for academic operations: /api/ask is for general doubt resolution, /api/plan is for study planning generation, /api/pyq is for PYQ analysis, /api/notes is for note generation. The AI engine and the backend are connected using REST API the backend sends requests to the AI engine and gets the response back then it sends the response to the frontend. The backend is very important because it helps the frontend and the AI engine to communicate with each other. The FastAPI framework is very useful for building the backend because it is very fast and it supports operations, which is good for real time applications. The backend is also very scalable it can handle requests at the same time, which is good for a system that has many users. The backend is also very secure it uses authentication and authorization to make sure that only authorized users can access the system and it also logs all the requests and responses which helps to monitor the system and fix any issues that may arise. Overall, the backend is an important part of the system it helps to connect the frontend and the AI engine and it provides many features that are useful for the system, such, as authentication, authorization, logging and caching. The backend is made using Python with FastAPI, which's a very good choice because it is very fast, scalable and secure.

D. AI Engine Layer

The AI engine layer is the part of the system that makes it smart. It works in a series of three steps:

- Intent and Context Detection
- Retrieval-Augmented Generation
- Response Formatting.

In the step the AI engine looks at what the student has asked and figures out what the question is about, which part of the syllabus it relates to, what the student is trying to do and how urgent the question is. This helps the system send the question to the part of the system to get an answer. In the step the system finds relevant information from a collection of knowledge about KTU S6 CSE, including parts of the syllabus summaries of modules past exam questions and notes. This information is added to the question to help the system give an accurate answer that is relevant, to what the student is studying. In the step the system takes the answer it has come up with and formats it in a way that makes sense for what the student is trying to do. This could be a list of notes a question and answer a study schedule, a piece of code or a long answer that follows the rules used by KTU. The AI engine makes sure the answer is presented in a way that's easy for the student to understand and use. The AI engine is a part of the system because it helps the student get the information they need in a way that is clear and useful. The AI engine layer is what makes the system smart and able to give answers to the student's questions.

E. Data Layer

The data layer of the KTU S6 CSE syllabus is really important. It has a lot of things like module objectives and topic listings. The data layer also has learning outcomes and reference textbooks for the KTU S6 CSE syllabus. The PYQ Repository is a place where we keep all the previous year question papers for the KTU S6 CSE syllabus. These papers are organized by year and subject and module and mark value. We also use metadata tags to identify questions that come up a lot. The Student Profile Store is where we keep information about each student. This includes things like how confident they're in a subject and what kind of explanations they like. We also keep track of what they have done in the past. We use a combination of things to make these stores work. We use SQLite for the structured profile data for the KTU S6 CSE syllabus. We also use a vector database like FAISS or Chroma DB to help with retrieval operations, in the RAG Pipeline for the KTU S6 CSE syllabus.

V. PERSONALIZATION ENGINE

A. Personalization Philosophy

The idea of the proposed system is that it can be personalized for each student. The system is based on the idea that to really help a student it needs to understand what the student is asking for and also consider the situation the student's in. This means looking at what the student knows how much time they have to

prepare, how confident they feel about the subject and what kind of explanations work best for them. The system is designed to be responsive, to the student's query and their learning situation. The personalization engine is student-adaptive which means it can adjust to each student's needs. The student's learning situation is very important. The system takes into account the students current knowledge state, available preparation time, subject-specific confidence levels and preferred modalities of explanation.

B. Student Profile Parameters

The personalization engine keeps track of a student profile. Updates it all the time. This profile has the following information:

- Active Subject and Module: This is the subject and module the student is studying now. It helps the engine figure out what to show the student and how to respond to them.
- Confidence Level: This is how good the student thinks they are at each subject module or how good the engine thinks they are based on how they interact with it. It helps the engine decide how complicated the explanations should be.
- Available Study Time: This is how days or hours the student has before a test or internal assessment. It helps the engine give the student a study plan that fits their schedule.
- Preferred Explanation Style: The student can choose how they want the engine to explain things to them. They can pick from things like notes, long answers, exam-style questions, simple explanations or explanations that use code.
- Exam Date Proximity: This is how close the students' exams are. It is marked as low, medium or high. It helps the engine decide what topics to focus on and how detailed the study plan should be.
- Weak Topics: These are the topics that the student's not good at based on what they say themselves or how they do on practice questions. The engine keeps track of these topics so it can help the student improve on Student Profile Parameters on their Weak Topics, in the student profile.

C. Adaptation Logic

The personalization engine uses these parameters at many different stages of creating a response. First, it looks at the student's module and the topics they are weak in when it is searching for documents. This helps it find the relevant information for the student. Then when it is creating the prompt it takes into account the students explanation style and how urgent the exam is. This tells the system what kind of response to give so it is the depth, tone and format. Finally, when it is formatting the response, it considers how time the student has to study. This determines how the study plan is structured and how much information is included. The personalization engine and the students profile parameters work together to make the response generation pipeline better. The students profile parameters, such, as their module and weak topics are used to make the system more effective.

D. Illustrative Personalization Scenario

Consider a student who inputs the personalization engine detects: subject = Compiler Design; module priority = Syntax Analysis; exam urgency = high; time = 3 days. The system then makes a 3-day study plan. This plan gives time blocks for each day. These blocks are for Compiler Design modules. The plan also prioritizes Syntax Analysis topics. The system also does the following:

- It finds PYQ questions from Syntax Analysis.
- It makes a revision checklist.
- It flags 13-mark exam questions.
- These questions are likely to come based on PYQ analysis.

This output is easy to use and is made for the student's situation. No general-purpose AI system can make this. It needs knowledge of the curriculum and adaptive learning.

VI. KEY FEATURES AND MODULES

A. Syllabus-Aware Subject Assistant

The chatbot gives help for all the major KTU S6 CSE subjects. It provides answers that are based on the syllabus of each subject. For example, if you ask "Explain LL(1) parsing" the system does not just give a definition from a book. It explains the concept in a way that is related to the KTU Compiler Design subject. It also tells you which module and learning outcome it is connected to. The system makes sure the answer is not too long or too short it is right based on the marks given for that topic in the KTU exam. The system can handle questions from S6 CSE subjects like Compiler Design, Computer Networks, Microprocessors and Microcontrollers System Software, Data Analytics and other elective subjects. For each subject the system has a syllabus that helps it give precise answers.

B. PYQ-Based Smart Question Prediction

The system also uses previous year question papers to predict which questions are likely to come in the exam. It looks at the questions that come every year the modules that have weightage, the types of questions that are asked and new topics that are becoming more important. This information is then given to students as a list of questions and it tells them which modules are more important and which questions are likely to come in the exam. The system also gives reasons for why it thinks these questions are important based on the trend of year questions. For example, it can show a table that analyzes the frequency of questions, in Compiler Design.

Table I
Sample Pyq Frequency Analysis — Compiler Design

Topic	Module	Frequency	Weightage
LL(1) Parsing	2	5	High
LR Parsing Techniques	2	6	High
Lexical Analysis & LEX	1	4	Medium
Intermediate Code Gen.	4	5	High
Symbol Table Management	3	3	Medium
Code Optimization	5	4	Medium
Syntax Directed Translation	3	5	High

C. Personalized Study Planner

The study planner module creates a customized schedule based on your exam dates and subjects. it offers four distinct options: a 1-day emergency plan, a 3-day intensive prep, a 7-day comprehensive review, and a specialized exam-day strategy. Each plan outlines specific daily topics, time allocations, and recommended resources like notes or practice tests, complete with a readiness checklist. to ensure efficiency, the planner prioritizes content based on past exam frequency, your personal confidence levels, and the weighted importance of each topic on the exam.

D. Smart Notes Generator

The smart notes generator simplifies study preparation by producing high-quality materials tailored specifically to KTU exam patterns. It creates a variety of resources, including module summaries, 2-mark question-answer pairs, detailed 5-mark responses, comprehensive 13-mark university-style answers, viva compilations, and lab record explanations. Every note is structured to meet the required word count and includes necessary diagram descriptions to align with exam expectations. By automating the material gathering process, this feature saves students significant time, allowing them to shift their focus from searching for resources to understanding and practicing core concepts.

E. Coding and Laboratory Assistant

The coding assistant module addresses the practical programming component of the S6 CSE curriculum. It supports C and Python programming tasks commonly encountered in KTU laboratory sessions, including compiler design lab exercises (implementing lexers, parsers, and code generators), computer networks simulations, and microprocessor assembly language programs. The module provides: complete code solutions with inline comments, step-by-step code walkthroughs, output explanations, debugging assistance for common errors, and explanations that connect code behavior to the underlying theoretical concepts covered in the subject syllabus.

F. Mini Project Support Assistant

Given that the KTU S6 curriculum includes a mini project component, the chatbot incorporates a dedicated module for project support. This module assists students in formulating project ideas aligned with current industry and academic trends, generating structured project abstracts, drafting methodology and system design sections, preparing presentation slides outlines, and anticipating viva questions. The module draws on knowledge of both the technical content of S6 CSE subjects and the documentation conventions expected by KTU for mini project submissions.

VII. IMPLEMENTATION

A. Technology Stack

The system uses a technology stack that is picked to work and be easy to use. It is also easy to set up. Works with the academic use case. The main technology parts are listed in Table II. The technology stack is important, for the system. The technology stack is what makes the system work properly. The following table describes the tech stack summary of the system.

Table II
Technology Stack Summary

Component	Technology
Frontend	HTML5,CSS3,JavaScript, React.js
Backend API	Python 3.10+, FastAPI
LLM Engine	OpenAI GPT / Mistral / LLaMA 2 (via API or local deployment)
Retrieval Engine	FAISS / ChromaDB (vector search)
Embedding Model	sentence-transformers (all-MiniLM-L6v2)
Relational DB	SQLite / PostgreSQL
Caching Layer	Redis
Deployment	Docker, Nginx, Cloud VM
Version Control	Git / GitHub

B. RAG Pipeline Implementation

The RAG pipeline is put in place using LangChain to manage everything. We get the knowledge base ready by breaking it down into pieces of text. Then we use a model to turn these pieces into something the computer can understand. After that we put them in a store called FAISS so we can find them later. When it is time to get an answer, we take what the student is

asking and turn it into something the computer can understand too. We use this to find the parts of the knowledge base that're most like what the student is asking. We usually pick the five parts that are the most relevant. Then we combine these parts with what the student's asking and some other information that is just, for them. This all gets put together to make the question that we ask the large language model.

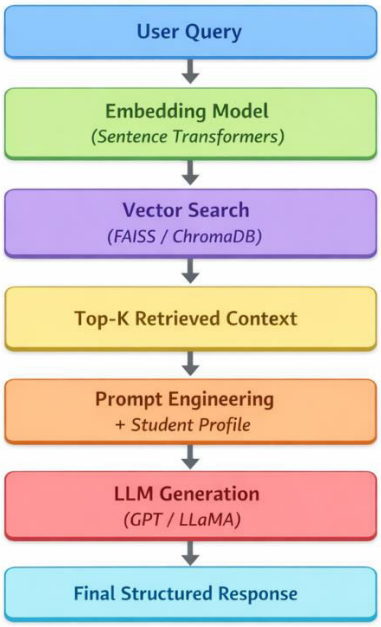


Fig. 2. Retrieval-Augmented Generation Pipeline for the Academic AI Assistant

C. *Prompt Engineering Design*

The way we design the engineering is like a template with a system. This system has five parts:

1. We tell the intelligence that it is an academic assistant for KTU S6 CSE
2. We make sure the artificial intelligence only talks about things that are related to KTU and are about academics.
3. We add information about the student who is using the system now.
4. We get information from the RAG module about what the student's studying.
5. We tell the intelligence how we want it to give us the answer.

We do it this way so that the artificial intelligence always gives us answers that're precise follow the syllabus and are in the right format. This is important for what we want to use it for. The prompt engineering design is made to work with the KTU S6 CSE assistant and it helps us get the right answers, from the artificial intelligence. The KTU S6 CSE academic assistant is a part of this system.

D. *Working Flow*

The working flow of the system is like this. A student types a question on the interface. The backend API gets the question. Adds the students' details from the database. Then it figures out what the student wants to know, like which subject and topic. Next it finds study materials and past year questions. After that it prepares a question for the system to answer. The system gives an answer, which is then made nicer and formatted. This formatted answer is sent back to the student's interface. Shown in the chat. The system also keeps data about the conversation. Logs it to make things better, over time.

VIII. EVALUATION AND RESULTS

A. *Evaluation Methodology*

The system was looked at in two areas: how good the responses were and what the users thought of it. The people who made the chatbot wanted to see how good the responses were. They got three teachers from the KTU CSE department to look at one hundred responses from the chatbot. These teachers looked at these responses in five ways: if the answer was correct if the answer was complete if the answer was relevant to the exam if the answer was easy to understand and if the answer was, in the right format. The teachers used a rating system to judge each of these areas. The people who made the chatbot also wanted to know what the users thought of it. They asked fifty students from the KTU S6 CSE department to use the chatbot for two weeks when they were studying for their exams. After that these students filled out a questionnaire to tell the people who made the chatbot what they thought of it.

B. *Response Quality Results*

Table III shows the ratings from teachers for each thing we looked at. The results show that we did well in all areas. We got high scores, for making sure the syllabus is correct and that the exams are relevant. This shows that our method of finding the information and asking the right questions works well. The thing that got the rating was how clear the explanations were. This suggests that changing the way we explain things to fit each person helps a lot to make the responses better. We are talking about the response quality of the RAG based retrieval and the way we engineer the prompts. The response quality is what we care about.

Table III
Expert Evaluation of Response Quality

Criterion	Mean Rating	Std. Dev.
Syllabus Correctness	4.62	0.41
Answer Completeness	4.31	0.53
Exam Relevance	4.71	0.38
Clarity of Explanation	4.78	0.34
Format Appropriateness	4.55	0.47
Overall	4.59	0.43

C. User Experience Results

The student usability survey got responses on five things: how easy it is to use how useful it is for studying for exams how good the personalized study plans are, how helpful the coding assistance is and how satisfied people are overall. The results are in Table IV. People were really satisfied they gave it a rating of 4.53 out of 5. They liked it most for helping with exams and the quality of the study plans. Some people wrote comments. They said the feature that predicts questions based on past years questions are really valuable and the automatic notes generator saves them a lot of time. There is a thing that need to be improved. Some people said they want detailed explanations, for really hard topics. Sometimes the system is slow when a lot of people are using it at the time.

Table IV
Student Usability Survey Results

Dimension	Mean Rating	Std. Dev.
Ease of Use	4.60	0.39
Exam Preparation Usefulness	4.74	0.31
Study Plan Quality	4.68	0.36
Coding Assistance Quality	4.40	0.52
Overall Satisfaction	4.53	0.44

D. Comparative Analysis

We compared our systems performance to two baseline conditions to see how well it did. The baselines were:

- A general AI assistant (GPT-3.5 Turbo not customized for our domain)
- A static study resource (PDF notes that were carefully curated)
- We got ratings from a panel of faculty members for the 100 questions for both baselines.
- Our system did better than the AI in all areas.
- It had an advantage in two areas:
- Getting the correct (+0.61 points)
- Being relevant, to exams (+0.58 points)

Compared to the notes our system scored much higher on being complete and clear. This makes sense because our system can adapt to questions and provide responses.

IX. FUTURE SCOPE

The KTU S6 CSE AI Assistant is a start. It has a foundation. This foundation will be used to make some changes in the future. The KTU S6 CSE AI Assistant can be used for than one semester and more than one branch of study. The KTU S6 CSE AI Assistant can be used for all eight semesters of the KTU B. Tech program. It can also be used for branches of engineering like Electronics and Communication Engineering, Electrical and Electronics Engineering and Mechanical Engineering. The KTU S6 CSE AI Assistant can be used with voice commands. This means students can talk to the KTU S6 CSE AI Assistant and it will understand what they are saying. This will be very helpful when students are traveling or studying and cannot type on their computers. This will also help students who prefer to speak in Malayalam. The KTU S6 CSE AI Assistant will also be able to understand Malayalam. This means students who speak Malayalam can use the KTU S6 CSE AI Assistant in their language. This will help students who're not comfortable speaking in English. The KTU S6 CSE AI Assistant has a collection of year questions. This collection needs to be updated by hand. In the future the KTU S6 CSE AI Assistant will be able to find and add new questions to this collection. The KTU S6 CSE AI Assistant will be improved in the future. It will be able to support semesters, branches and languages. The goal is to make the KTU S6 CSE AI Assistant a tool for all students, at KTU. Students will also be able to upload their notes to the KTU S6 CSE AI Assistant. The KTU S6 CSE AI Assistant will then summarize these notes. Add them to the student's study plan. This will help students study effectively. The system is going to have a feature that lets students upload their own PDF notes or lab records. This feature is called PDF Note Upload and Summarization. When students upload their PDF notes or laboratory records the system will look at them. Make a summary. The system will also find the important points and use them to make study plans that are just, for each student. The system will also use these points for Q&A sessions.

X. CONCLUSION

This paper delineates the design, implementation, and assessment of the KTU S6 CSE Personalized AI Assistant Chatbot— a specialized, intelligent academic support system specifically developed for Semester 6 Computer Science and Engineering students at APJ Abdul Kalam Technological University. The system fills a big gap in the academic support system for KTU engineering students by giving them personalized, instant, syllabus-aligned, and exam-optimized academic help through a conversational AI interface. The architecture suggested will use RAG to generate replies based on information and a structured prompt engineering for

producing LLMs with specific domain behavior. In addition to having an academic situation-specific set of features (e.g.; to clarify doubts, to use PYQs to predict questions, to assist with personal study planning, automated note taking, coding assistance, and mini projects) that makes the system a comprehensive academic companion for supporting all aspects of the S6 CSE curriculum. Evaluation by adult learners and studies to test the usability of the prototype have confirmed through both measures that the proposed method is effective, with high satisfaction ratings for syllabus accuracy, the relevance of exams and assessments to the course, and overall satisfaction of learners. The use of a domain-specific AI knowledge base developed by a subject matter expert to incorporate knowledge about the curriculum as well as employ adaptive personalisation creates an educational support solution that provides learners with significantly more academic support than general-purpose AI tools. Future work for the system will add the capability to support additional semesters, branches and languages; ultimately creating an institution-specific, scalable AI academic support platform for all students at KTU. The proposed architecture and methodology described in this paper provide an exemplary framework for building domain-specific AI educational support solutions across a broad range of post-secondary educational contexts.

ACKNOWLEDGMENT

The authors would like to thank the teachers at the Department of Computer Science and Engineering at "Ahalia School of Engineering and Technology" for helping us and supporting us while we were working on this project. We also want to thank the students who were in the semester of Computer Science and Engineering, at KTU. They tried out our project and gave us really good ideas on how to make it better.

REFERENCES

[1] A. Fitzpatrick, "Woebot: The chatbot that helps you fight depression," *Time Magazine*, 2017.

[2] T. B. Brown et al., "Language models are few-shot learners," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2020, vol. 33, pp. 1877–1901.

[3] Khan Academy, "Khanmigo: AI-powered learning guide," [Online]. Available: https://www.khanacademy.org/khan-labs, 2023.

[4] McGraw-Hill, "ALEKS: Assessment and Learning in Knowledge Spaces," [Online]. Available: https://www.aleks.com, 2022.

[5] D. Luckin, "Machine learning and human intelligence: The future of education for the 21st century," *UCL IOE Press*, 2018.

[6] V. Raina and M. Kumar, "Challenges and opportunities in AI-enabled educational technology in India," in *Proc. IEEE International Conference on Technology for Education (T4E)*, 2021, pp. 120–127.

[7] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Proc. NeurIPS*, 2020, vol. 33, pp. 9459–9474.

[8] J. Wei et al., "Chain-of-thought prompting elicits reasoning in large language models," in *Proc. NeurIPS*, 2022, vol. 35, pp. 24824–24837.

[9] H. Chase, "LangChain," [Online]. Available: https://github.com/ langchain-ai/langchain, 2022.

[10] J. Johnson, M. Douze, and H. Jegou, "Billion-scale similarity search´ with GPUs," *IEEE Transactions on Big Data*, vol. 7, no. 3, pp. 535– 547, 2021.

[11] APJ Abdul Kalam Technological University, "B. Tech CSE S6 Syllabus," [Online]. Available: https://ktu.edu.in, 2019.

[12] S. Ram´ırez, "FastAPI: Modern, fast web framework for building APIs with Python," [Online]. Available: https://fastapi.tiangolo.com, 2018.

[13] N. Slijepcevic and A. Yaylali, "Leveraging 'Khanmigo' Generative AI-Powered Tool for Personalized Tutoring to Learn Scientific Concepts," Journal of Teaching and Learning, vol. 19, no. 4, pp. 155–178, 2025.

[14] "Retrieval-Augmented Generation for Educational Application," ScienceDirect: Results in Engineering, vol. 25, 2025.

[15] "Retrieval-Augmented Generation (RAG) Chatbots for Education," Applied Sciences, vol. 15, no. 8, 2025.

CoEdit: Real-Time Collaborative Code Editor

Dr. S Gunasekaran
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, India
gunaphd@yahoo.com

Ms. Anisree P G
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, India
anisree.pg@ahalia.ac.in

Midhun V
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, India
midhun_v@ahalia.ac.in

Mohammed Shahal A
Department of CSE
Ahalia School of Engineering and Technology
Palakkad, India
mohammedshahal_a@ahalia.ac.in

Mrinal V
Department of CSE,
Ahalia School of Engineering and Technology
Palakkad, India
mrinal_v@ahalia.ac.in

Navaneeth Krishna M M
Department of CSE,
Ahalia School of Engineering and Technology
Palakkad, India
navaneethkrishna_mm@ahalia.ac.in

Abstract— Collaborative software development is becoming increasingly common in modern programming environments. This paper describes the design and implementation of a Collaborative Code Editor that allows multiple users to edit a given source code file simultaneously. It uses a function-level locking feature that allows users to lock certain sections of code to prevent conflicts. This feature is important to enhance real-time collaboration among users.

Keywords— collaborative code editor, real-time collaboration, CRDT, distributed systems, software development tools

I. INTRODUCTION

Collaborative software development has now become an essential practice in modern programming environments. This is due to the rise of remote teams and online learning environments, which require developers to work with the same codebase in real time. The traditional approach to software development uses version control systems, where developers can work separately and merge their codebase. However, this approach can lead to conflicts in collaborative development. Therefore, to mitigate the challenges in collaborative development, real-time collaborative editing systems have been introduced [2]. This technology enables multiple users to work with the same document in real time. In the context of programming, collaborative code editors have been introduced to allow developers to work with the code in real time [1]. The proposed project is to implement a Collaborative Code Editor with a Function Lock Mechanism, which can mitigate conflicts in collaborative development by allowing users to lock functions while making changes.

II. RELATED WORK

Real-time collaborative editing systems have attracted considerable attention in recent years due to the increased need for distributed software development and remote working. Several tools and research works have proposed various mechanisms that allow users to edit a document or shared code in real-time, and the data should be consistent across various users' views [2].

The first mechanism that has been proposed in various collaborative editing systems is Operational Transformation (OT) [2]. Operational Transformation is a mechanism that helps users to edit a shared document simultaneously. All users can edit the document, and finally, they reach a similar state. Operational Transformation has been used in various collaborative document editing tools like Google Docs, where users can view and edit the document in real-time. Although Operational Transformation helps to achieve consistency, in some situations, the mechanism can become complex, especially when dealing with many operations.

Another technique that is increasingly being used is Conflict-free Replicated Data Types (CRDT), which allows distributed systems to be consistent without using a centralized system of conflict resolution. The CRDT-based system ensures that it converges to a single state regardless of the sequence in which updates are made to the system [2]. This technique is used in collaborative code development tools such as Visual Studio Code Live Share, where developers can collaborate in real time and edit code and debug it together.

There are many web-based code development platforms that offer code collaboration tools where developers can collaborate with one another in creating code and executing it. Replit is one such code development platform where developers can share a code environment and edit code and execute it simultaneously. Although collaboration is enabled in code development, no specific tools are offered for avoiding code editing conflicts in particular parts of code [1].

Although various tools have been developed to aid in collaborative programming, most existing systems do not offer fine-grained controls to handle concurrent editing. For instance, if multiple users attempt to edit the same section of the code, conflicts may arise, and the changes may get overwritten. To handle conflicts, some coding tools have adopted the use of version control, where developers can work independently and merge their changes at a later date. Nevertheless, this method does not enhance the aspect of real-time collaboration.

III. PROPOSED METHOD

The proposed solution for the problem is the development of a web-based collaborative code editor that enables multiple users to edit the same source code file simultaneously in real time. The system aims to improve collaboration among developers by providing a synchronized code editing environment, enabling effective user

communication, and providing mechanisms for managing editing conflicts. Unlike the traditional software development process, where developers edit files individually and then merge changes, the proposed system enables multiple users to edit the same file while viewing changes made by other users in real time [1].

The primary objective of the system is to provide a collaborative programming environment where multiple users can contribute to the development of the code simultaneously without interfering with each other's work. To accomplish this, the system uses a real-time synchronization mechanism that ensures all connected users have a consistent view of the code document. Whenever a user makes changes to the document, the changes are sent to the server, which then notifies all connected users to update the document in real time [5].

Another important aspect of the system that will be implemented is the function lock mechanism. In collaborative programming environments, conflicts may arise when different users attempt to modify the same code at the same time. This can be solved using the function lock mechanism. This mechanism will allow a user to lock a specific function or code block while editing it. When a user locks a function, he or she can still view the code, but others will not have the permission to edit that specific section of code. This will ensure that the integrity of the code is maintained and that users do not accidentally overwrite code.

The system will operate under a client-server architecture. In this system, the client will have an interactive user interface that will allow users to edit their code. Communication between the clients and the server will be done in real-time using WebSocket-based technologies [5]. To ensure that the document being shared remains in sync across different clients, the system will use distributed synchronization techniques like CRDT [2].

Besides, the system offers essential facilities like user session management, real-time cursor update, and shared document visibility. This will increase user awareness and coordination during collaborative programming. The architecture of the proposed system is also scalable, allowing multiple users to connect and perform their activities without compromising the performance of the system.

Therefore, the proposed solution will be effective in developing an efficient collaborative programming environment that will minimize code conflicts, enhance teamwork, and boost productivity in distributed software development environments. By incorporating real-time synchronization with the proposed function-level locking mechanism, the proposed solution offers an effective solution to manage simultaneous changes to shared source code files.

IV. METHODOLOGY

The development of the collaborative code editor has been carried out by following a structured methodology that focuses on designing, implementing, and testing the system to ensure that it can support collaborative programming in real time. This includes the development of the system, selection of technology, implementation of real-time communication, integration of synchronization, and evaluation of the system.

The first step in the methodology is to design the system's architecture. In the proposed system, the architecture will be client-server. In this type of architecture, numerous clients will connect to a single server through a web interface. In this system, the client will contain the code editor interface, while the server will contain the user connection and communication functionality. This type of architecture will ensure that all the updates made by the users are transmitted in real-time. The system architecture will contain three major components:

Client Interface: This will contain the code editing interface where users can edit their source code.

Server: This will contain the user's connection functionality.

Synchronization Engine: This will ensure that all users have the same document version by handling concurrent editing [2].

Technologies used in the implementation of the system include modern web development technologies that support real-time communication and user interaction. Technologies used in the frontend are for creating the user interface and interactive editor. A web editor library is used for providing syntax highlighting, code formatting, and editing. Backend technologies are used for managing server operations, handling client requests, and maintaining synchronization.

The primary technologies used include:

Frontend: Web-based interface with a code editor component

Backend: Node.js server for handling requests and communication

Real-Time Communication: WebSocket-based communication for instant updates [5]

Synchronization: Conflict-free Replicated Data Types (CRDT) for maintaining document consistency [2]

Real-time collaboration can be facilitated through the implementation of a communication mechanism that sends the editing operations of one user to other users in real-time. When one user makes a change to the code, the change is sent to the server, and the server sends the change to other users. The synchronization mechanism helps ensure that each client has the same copy of the document. CRDT-based algorithms are used to handle editing operations that may occur simultaneously [2].

To reduce conflicts while editing code in shared code sections, the system provides a function of lock feature. When a user wants to edit a particular function, he/she can lock that function. After locking the function, other users can view the code but cannot edit the locked function until the user who locked the function finishes editing. Such a feature can coordinate with the users who want to edit the code.

The last phase of the methodology is testing the system with various users to assess its functionality and efficiency. The testing phase ensures that the editor works successfully regarding simultaneous editing, consistency, and conflict resolution with the help of the function lock. Synchronization delay, system responsiveness, and efficiency are analyzed to assess the reliability of the system regarding real-time programming environments.

V. IMPLEMENTATION

The implementation of the collaborative code editor aims at creating a real-time programming environment where users can edit the same code file simultaneously. The implementation of the proposed system is done using a web-based client-server architecture that can handle real-time communication. The user interface developed for the client side is responsible for providing a platform where users can interact and create or edit the code. For this purpose, a web-based code editor component has been included, allowing users to create and edit the code. The interface has been designed to allow users to open shared coding sessions and view the updates made by other users in real-time. The interface has been designed to include updates that indicate which user is editing a particular section of the code. The server side of the system will be used to manage collaborative sessions. When users connect to the system, their sessions will be registered by the server. The server will be used to manage the function locking mechanism. This will be done by locking the functions that each user has access to. This will be done to prevent other users from editing the functions that have been locked by other users. The real-time communication between the clients and the server will be done using the WebSocket protocol. This will be used to allow users to communicate in real time. This will be done by allowing the users to send requests to the server. When the user makes changes to the code, the changes will be captured by the editor. This change will be sent to the server. After that, the change will be sent to all the users who are connected to the system. For consistency in various clients, the system has used synchronization. For this, the concept of distributed editing has been used. All operations are applied to the document. All operations are synchronized, and consistency is maintained in the document for all

users. Thus, multiple users can edit the document simultaneously without causing any inconsistencies in the shared code document. For implementing the function lock mechanism, conflicts can be avoided when multiple users are trying to modify the same code. When a user selects a function to edit, the system allows the user to lock the function. When the function is locked, the information is sent to the server and other users. Once the function is locked, other users can view the function, but they cannot modify the function. After the user has finished editing, the function can be released, and other users can modify the function. VI. RESULT The developed collaborative code editor tool was tested to ensure that it can support real-time collaboration among multiple users. This was done by allowing multiple users to connect to the system and edit the same source code file simultaneously. From the results obtained, the proposed system can support collaborative programming in real time with minimal delays. In this testing, multiple users were able to connect to the system and edit the same code simultaneously. Once one user made some changes to the code, they were immediately reflected in the editors of other users who had connected to the system. This proved that the real-time communication mechanism can effectively synchronize changes to the code among multiple users. From this testing, the implementation of the function lock feature can effectively reduce conflicts that could arise during collaborative coding. Once one user had locked out a certain function for editing, other users were able to view the code but could not edit it until they had been given access by the user who had locked the function.

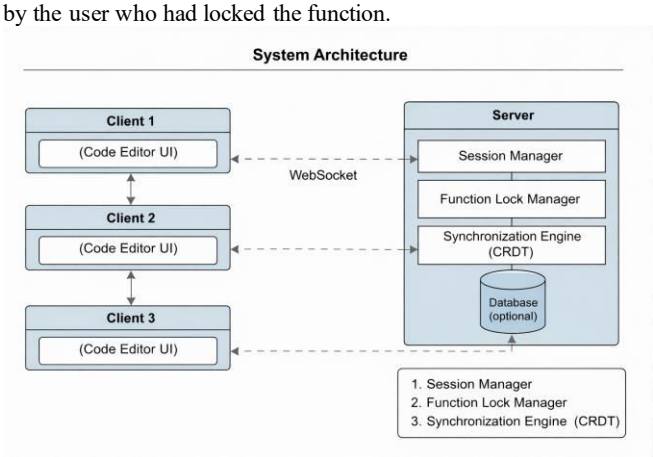


Fig. 1. Architecture of the Real-Time Collaborative Code Editor

VI. RESULT

The developed collaborative code editor tool was tested to ensure that it can support real-time collaboration among multiple users. This was done by allowing multiple users to connect to the system and edit the same source code file simultaneously. From the results obtained, the proposed system can support collaborative programming in real time with minimal delays. In this testing, multiple users were able to connect to the system and edit the same code simultaneously. Once one user made some changes to the code, they were immediately reflected in the editors of other users who had connected to the system. This proved that the real-time communication mechanism can effectively synchronize changes to the code among multiple users. From this testing, the implementation of the function lock feature can effectively reduce conflicts that could arise during collaborative coding.

VII. CONCLUSION

In this paper, the design and implementation of a collaborative code editor that enables different users to edit the same source code file simultaneously in real-time was presented. The main aim of designing this system was to promote collaboration among different developers and reduce conflicts that may occur when different users attempt to edit the code simultaneously. The system that was proposed in this paper was based on a web-based client-server system that enables real-time communication among different users. In this project, the major contribution was the implementation of a function lock system that enables different users to lock different functions when editing them. This system helps to reduce the chances of different users editing the same code simultaneously. The system that was proposed in this paper was based on a web-based client-server system that enables real-time communication among different users. In this project, the major contribution was the implementation of a function lock system that enables different users to lock different functions when editing them. This system helps to reduce the chances of different users editing the same code simultaneously.

IX. REFERENCES

[1] H. Pathak, H. Ritik, R. Pawar and Y. Pingle, “CodeFlow: Real-Time Collaborative Code Editor,” 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS)

[2] Real-time collaborative coding in a web IDE, Max Goldman, Greg Little, Robert C Miller Proceedings of the 24th annual ACM symposium on User interface software and technology

[3] A. Amburle, C. Almeida, N. Lopes and O. Lopes, "AI based Code Error Explainer using Gemini Model," 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)

[4] CS50 sandbox: secure execution of untrusted code David J Malan Proceeding of the 44th ACM technical symposium on Computer science education

[5] V. Pimentel and B. G. Nickerson, "Communicating and Displaying Real-Time Data with WebSocket," in IEEE Internet Computing

[6] M. Shapiro, N. Preguiça, C. Baquero, and M. Zawirski, “Conflict-Free Replicated Data Types,” Stabilization, Safety, and Security of Distributed Systems, Springer, 2011.

[7] C. Sun and C. Ellis, “Operational Transformation in Real-Time Group Editors: Issues, Algorithms, and Achievements,” ACM Conference on Computer Supported Cooperative Work, 1998.

[8] D. Li and R. Li, “Preserving Operation Effects in Real-Time Collaborative Editing Systems,” ACM CSCW, 2004.

[9] N. Preguiça, J. M. Marquès, M. Shapiro, and M. Letia, “A Commutative Replicated Data Type for Cooperative Editing,” IEEE International Conference on Distributed Computing Systems, 2009.

[10] J. Nichols, B. Myers, and K. Litwack, “Improving Automatic Documentation for Real-Time Collaborative Software,” IEEE Symposium on Visual Languages and Human-Centric Computing, 2009.

[11] T. Oster, P. Urso, P. Molli, and A. Imine, “Data Consistency for P2P Collaborative Editing,” ACM CSCW, 2006.

[12] A. Ignat and M. Norrie, “Flexible Synchronization in Real-Time Collaborative Editing Environments,” IEEE International Conference on Collaborative Computing, 2012.

[13] R. Kleppmann and M. W. Bernstein, “A Highly-Available Move Operation for Replicated Trees,” IEEE Transactions on Parallel and Distributed Systems, 2020.

[14] R. Kleppmann, “Designing Data-Intensive Applications,” O’Reilly Media, 2017.

[15] M. Wessel, R. Hiesgen, and J. Schöning, “How Do Developers Collaborate in Real-Time?,” IEEE Software, vol. 36, no. 2, pp. 83–89, 2019.

TRACK-3

ELECTRONICS & COMMUNICATION ENGINEERING

BRAIN TUMOR DETECTION USING MULTIMODAL MRI SCANS WITH A DEEP LEARNING ALGORITHM AND EDGE AI

Dr.Leesha Paul
dept.Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
lishababuk@gmail.com

Dr.V Balamurugan
dept.Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
hodece.aset@ahalia.ac.in

Pooja P Menon
dept.Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
pooja.p@ahalia.ac.in

Subiksha S
dept.Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
subiksha@ahalia.ac.in

Mohammed Haddiq S
dept.Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
mohammed.haddiq@ahalia.ac.in

Nikhil V
dept.Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
nikhil.v@ahalia.ac.in

***Abstract*—The most important task of medical imaging is to identify the brain tumor at the earliest stage and proceed with the correct treatment, based on the MRI images. In order to enhance the quality of generated objects and the effectiveness of the calculation, a high-level deep learningbased model of automated tumor classification will be adopted in this project to apply the state-of-the-art neural network models, i.e., MobileNet and VGG16. The simplified model-complexity tradeoff enabling lightweight MobileNet architecture to be deployed in real-time and the capability of the deep VGG16 based architecture to learn more details with high fidelity and precision classification results are realised.The system will identify two brain MRI images tumor and non-tumor. The classification networks rely on the MobileNet and VGG16 networks to maximize the accuracy and to reduce the computation resources. MobileNet offers a smaller version that can be implemented on edge computing and mobile hardware, with shorter inference time at the cost of the deeper layers, however VGG16 has additional layers, which make the features extracted by it more detailed and can be more easily classified. Potential improvement of the methodology is a better timely diagnosis and elimination of invasiveness of diagnostic measures. A combination of the models and real-time diagnostic devices will allow clinicians to improve clinical decision making and this will translate to improved tumor detection.**

***Keywords*—The authors in the present study care about identifying the correlation between Brain Tumor, MRI, and MobileNet and VGG16, Deep Learning, and Accuracy.**

I. INTRODUCTION

Brain tumor was considered as one of the most devastating and dangerous health risk factors which can dramatically affect the quality of the life and survival rate of such patients. Diagnosing and catching the brain tumor early lays the groundwork for planning treatment that can boost patients' potential for good outcomes. These traditional approaches are limited by the need for manual analysis of Magnetic Resonance Imaging (MRI) radiology images by radiologists. Some limitations of manual analysis are: It requires a lot of time, the analysis is prone to human error, and it also depends on the expertise of the radiologist. And this brings us to the problem; to find known and reliable diagnostic equipment that will help the professional doctors in making fast judgments and accuracy in diagnosing the brain tumors.[1] In recent years, deep learning has been widely used in providing automation to medical diagnosis for complex diseases, such as brain tumor. We did so knowing that we will later build an algorithmic very deep learning structure that would classify the brain tumors in this project we will be using the mobile net and VGG16 that are really accurate with not much computer time. In the context of these architectures, the project will overcome the posteriorities that burdening the standard processes for diagnosis with an integrating solution capable of achieving resource savings, without compromising the convincing results in real time.[2] The decision to utilize a two-stage architecture of MobileNet and VGG16 for the project was the product of the different advantages provided by the implementations of each of these models. Specifically, the MobileNet gained popularity for being considered as a lightweight model that can be deployed on real-time applications and devices with constrained resources like mobile devices and edge computers. Beyond that it is a model of highly simplified computation. In clinical application where time and efficiency are the ruling factors, the accuracy it provides is more than adequate.[3] VGG16 is a great way for extraction of features, that means the model will be able to learn and perform complex pattern recognition tasks, successfully. Combining both MobileNet and VGG16 will help us to create a reliable brain tumor detection system, and so support the radiologists making decision in the clinic.[4]

II. RELATED WORK

V Bhanothu, A Kamalakannan and G. Rajamanickam, "Deep Convolutional Network for Detection and Classification of Brain Tumor from MRI Images," IEEE. In this paper, we propose Faster R-CNN based brain tumor detection methods, in which the VGG-16 architecture is employed for both the region proposal stage and the classification stage. We further demonstrate the effectiveness of this approach, which integrates Faster R-CNN and VGG-16 architectures in the detection of brain tumor, achieving an

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

average precision of 77.60% for three tumor classes, i.e. glioma, meningioma and pituitary.

R. M. Prakash & R. S. S. Kumari, "Classification of MR Brain to detect Tumor using Transfer Learning from pre-train CNN models," IEEE. Their research area is centered on classification of brain MRI image, for which they apply transfer learning on the already trained CNN backbones like Inception, VGG16, ResNet. Their proposed model is proven to be a very efficient one as it's able to get an accuracy of 100% in a dataset gathered from Harvard Medical School. This is proof of the synergy between transfer learning and data augmentation.

M. Nazir, M. A. Khan, T. Saba and A. Rehman, "Brain Tumor Detection Using Multi-level Wavelets from MRI Images," 2019 International Conference on Computer and Information Sciences (ICCIS). In this study, a wavelet-based segmentation method is suggested for the detection of brain tumor. Therefore, by decomposing the MRIs into wavelet sub bands and selecting the high variance features, the proposed method achieves an improvement in accuracy from 98.2% to 99.7%. It shows that multi-level wavelets can be very useful in extracting discriminative features for tumor detection.

M. Y. Mehemud, H. B. Kibria and A. Salam, "The Efficiency of Transfer Learning Models for Brain Tumor Classification," IEEE In this work, we train EfficientNetB0 and DenseNet121 to automatically classify brain tumors (which have four classes, three types of brain tumor and one non-tumor class). The models achieve average accuracies of 97.09% and 98.32% respectively, outperforming previous methods, which indicates the effectiveness of DenseNet121 and EfficientNetB0 in the classification of brain tumor.

M. A. Hafeez, C. B. Kayasandik and M. Y. Dogan, "Brain Tumor Classification Using MRI Images And Convolutional Neural Networks", IEEE In this paper, we suggest a deep chisquare neural network parallel network framework with 15 layers to classify the three levels of brain tumour based on the dataset of 3064 brain MRI images. Using the model, an amazing accuracy of 98.6% is achieved, and a precision, recall rate, and F1 score of more than 98% are obtained, proving that CNN can effectively distinguish between tumor categories from MRI images.

N. M. DI gut, S. A. Shohan, K. M. A. Salisbury, IEEE "Deep Learning Based Brain Tumor Detection and Classification" In this paper, we establish a two-stage pipeline for our speedup in tumor detection based on YOLOv5 and our incremental classification based on FastAi for BRATS 2018 data. The accuracy of the YOLOv5 model was 85.95%, and the accuracy of the FastAi classifier was 95.78%, which confirms the feasibility of using the two methods to predict brain tumors in real time for clinical trials.

III. PROPOSED METHOD

In our case, data has been collected from public MRI brain imaging datasets of labeled images in tumor and non-tumor classes. Such volumes are collected from hospitals, advanced medical training centers and research centers with brainrelated diseases. All relevant privacy concerns were addressed; however, the data is heterogeneous, reflecting the variability in age, gender and tumor type. The information is then annotated with tumor or non-tumor classes, which is required to train deep neural networks. The images themselves are saved with other extensions, for example, PNG and JPEG. Care is taken to separate the data into training, validation, and testing in such a way as to allow the algorithm to be appropriately assessed[5]

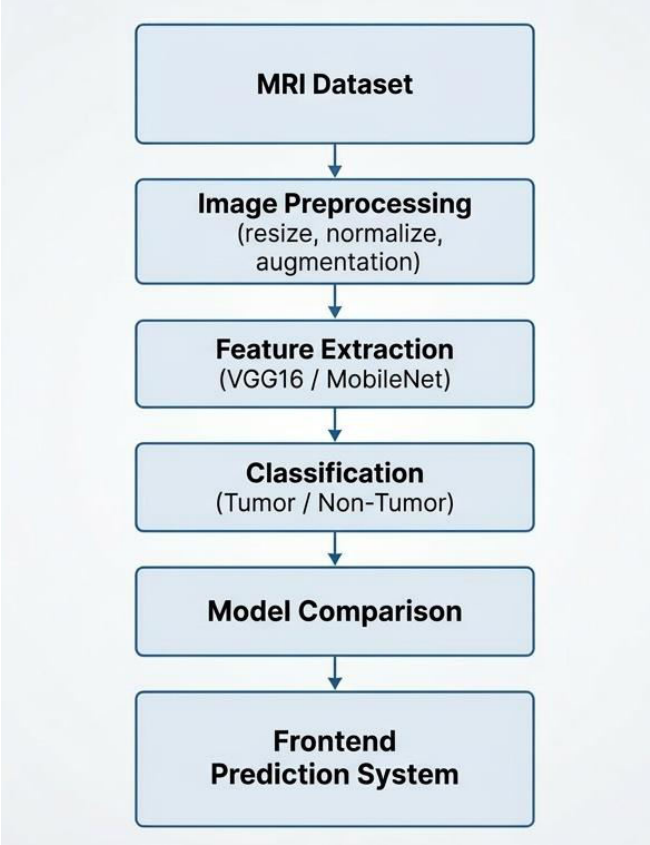


Fig.1: System Workflow

A. Data Collection

In our case, data has been collected from public MRI brain imaging datasets of labeled images in tumor and non-tumor classes. Such volumes are collected from hospitals, advanced medical training centers and research centers with brain related diseases. All relevant privacy concerns were addressed; however, the data is heterogeneous, reflecting the variability in age, gender and tumor type. The information is then annotated with tumor or non-tumor classes, which is required to train deep neural networks. The images themselves are saved with other extensions, for example, PNG and JPEG. Care is taken to separate the data into training, validation, and testing in such a way as to allow the algorithm to be appropriately assessed[5]

B. Data Preprocessing

The first step in data preprocessing is scraping the MRI images, which need to be downloaded from all the corresponding data sources before being coupled together. Now I'll reduce the images' shape to a common resolution to create each image in the input shape of MobileNet and VGG16. The images are normalized (pixels values in the MRI images scaled down to [0,1]) to speed up model convergence. The model generalization is enhanced using data augmentation techniques (rotation, flip, zooming, etc), which provide additional variability in the training samples. In addition, optimal preprocessing gets rid of any noisy pixels or uneven artefacts from the MRI images. Therefore, good clean data is seen to be provided to the neural network. This is

probably one of the most important steps to improving the model's accuracy and making the model more robust.[6]

C. Model Development

In the model development, we choose both MobileNet and VGG16 architectures to distinguish brain MRI images into tumor and non-tumor images. As mobile architectures are typically employed in low-power embedded devices, MobileNet is chosen due to its outstanding efficiency. This, in turn, makes it very useful for real-time applications, where sacrificing some speed to maintain performance is no sacrifice. In contrast, the deep convolutional structures from VGG16 are utilized to exploit its feature extraction for faint motifs recognition for MRI images. Like the other model too, it is trained on the preprocessed dataset by using supervised learning. In addition to that, the models are also benefiting from so-called transfer learning where pre-trained weights collected from huge image datasets are fine-tuning on brain MRI images.[7]

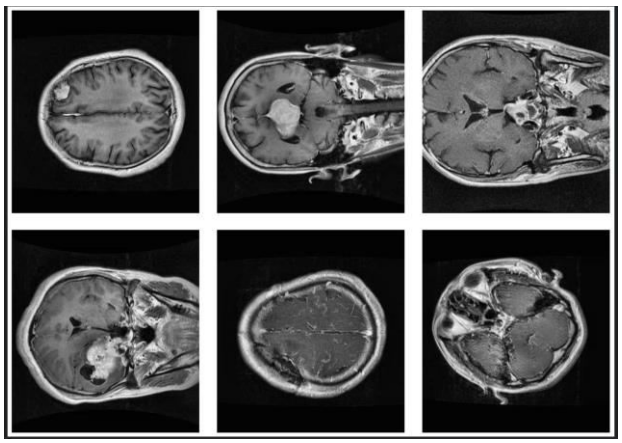


Fig.2: Sample MRI images from the dataset

D. . Model Evaluation

The models are evaluated using common performance measures, such as accuracy, precision, recall and F1-score. Confusion matrix is used to evaluate the performance of the model for classification of tumor, and non-tumour regions. As with the evaluation metrics, cross-validation is used to ensure the stability of these models and to ensure they perform well with new, unseen data. The trade-off between the classification accuracy and the computational speed is understood by comparing the execution speed of MobileNet and VGG16. However, considering that the MobileNet may not work in real time mode, the VGG16 networks are considered for their accuracy in tumor detection. During the selection of the best model for deployment, we choose the most accurate model to ensure that the model we selected isas accurate as possible and the inference time is as low as possible which is very crucial in any medical use case.[8]

E. Training Process

The MRI images are first preprocessed by resizing them to 224×224 pixels and normalising the pixel values to improve model performance. The dataset is then divided into training, validation, and testing **sets** to properly evaluate the models. Both VGG16 and MobileNet models are trained using transfer learning, where the pre-trained layers are used for feature extraction and new fully connected layers are added for binary classification. The models are trained using the Adam optimiser and binary cross-entropy loss function, while accuracy is used as the main evaluation metric. After training, the performance of both models is analysed and compared to select the best-performing model for deployment.

IV. METHODOLOGY

MOBILENET:

Since MobileNet Light is a lightweight deep learning model, you have to know how it functionalities. Such knowledge is crucial as it directly relates to the brain tumor detection and segmentation methodology. As you can tell by the name, MobileNet belongs to the mobile and edge devices, so it is trying to achieve a balance between efficiency and accuracy. The first step is data acquisition by which brain MRI images are obtained. These images must contain non-tumor cases as well as tumor cases. In order to better prepare and improve the quality of images prior to their use in model training, the dataset is pre- and post-processed. This preprocessing can involve resizing images to a consistent size, normalizing the pixel values to a range readable by the neural network, and augmenting the dataset to make the dataset more expressive and resilient via rotation, flipping.[9] After completing data preprocessing, the next step is to initiate the configuration of the MobileNet architecture which is known for depthwise separable convolution layers. This means a significant reduction in the number of parameters and calculations necessary, an essential prerequisite for the development of a system with real-time inference. The low latency inference of the network is especially important for inference on devices with limited computation power. The network is a stack of convolution layers followed by ReLU activation. In addition, the modeler has added batch normalization and dropout layers, which improve generalization and help avoid overfitting. The last layers are designed to generate the probabilities in which the images of brain MRIs to be classified into two classes. The training model is optimized with respect to the correct loss function (in classification problems the binary cross entropy is the most popular) and loss optimizers (Adam, SGD). The training, then, is achieved by feeding the preprocessed images to the network itself so that the network can learn the particularities of the images containing the tumors.[10]

VGG16:

The reason that brain tumor identification using VGG16 deep learning model works is because it captures the overall network of image data with simplicity and clarity yet with depth and efficiency. VGG16 contains a total of 16 layers (mainly convolutional layers), the dimensionality of the features extracted increases gradually layer by layer. First the large set of brain MRI images is acquired with equal number of tumor and non-tumor cases. We prepare the data beforehand by resizing all the images to a consistent input size, normalizing pixel values for model efficiency and performing data augmentation for variability. We show that augmentations of standard operations (i.e., zooming operation, rotation operation, and horizontally flip) are good choices for training the model and transferring it to new situations in terms of generalization.[11] Next we configure VGG16 for brain tumor prediction. Unlike typical fully connected type networks, VGG16 is made up of blocks of several convolutional layers followed by a max pooling layer.

This superior form factor is useful for capturing a lot of detail and is particularly useful in medical imaging where tiny detail can be important in many scenarios. An appropriately chosen loss function - typically binary cross Submission ID entropy for the task - is optimized by an optimizer (such as Adam or SGD), until the entire preprocessed data has been used by the model. The model is able to extract brain tumor features with increasing recognition performance with each epoch, while the behavior is recorded by model validation set for preventing the model from overfitting.[12]

V. Results and Analysis

The proposed brain tumor detection system was trained using the VGG16 convolutional neural network with transfer learning. During training, the MRI images were passed through the convolutional layers of the network to extract important features such as edges, textures, and structural patterns related to tumor regions. The model was trained using the Adam optimiser and binary cross-entropy loss function, while accuracy was used as the main evaluation metric.

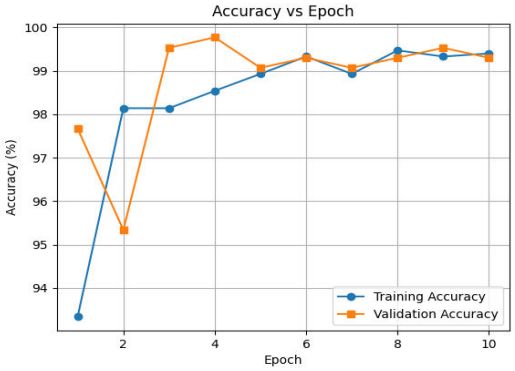


Fig.3: Accuracy graph of MobileNet

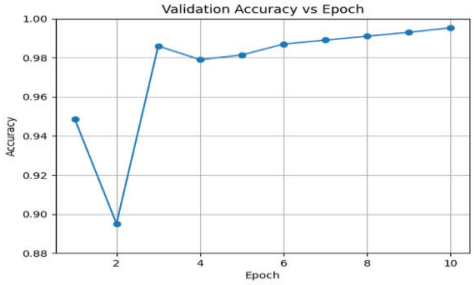


Fig.4: Accuracy graph of VGG16

The training and validation curves show that the model gradually improved its performance as the training progressed. The increase in accuracy and reduction in loss indicate that the model was able to learn meaningful patterns from the MRI images. After training, the model was evaluated using a test dataset consisting of 355 MRI images. The model achieved an overall classification accuracy of 99% for VGG16 and 98% for the MobileNet model, indicating that it can correctly classify a majority of tumor and non-tumor MRI images. To further evaluate the model's performance, a confusion matrix was generated. The confusion matrix provides a detailed analysis of the classification results by showing the number of correctly and incorrectly predicted samples for each class.

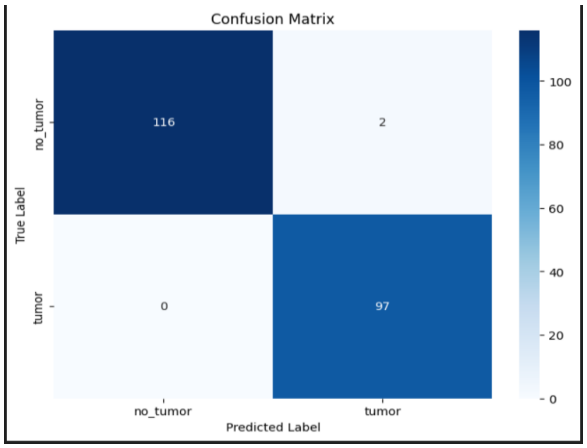


Fig.5: Confusion Matrix of VGG16

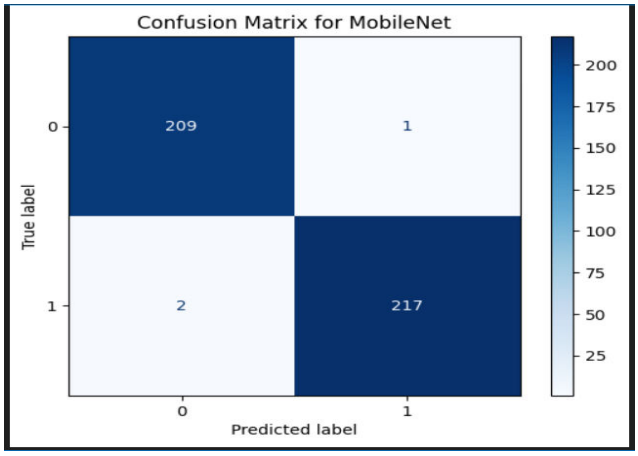


Fig.6: Confusion Matrix of MobileNet

From the results, it was observed that the model was able to effectively differentiate between tumor and non-tumor MRI images in most cases. However, some misclassifications occurred due to similarities in image intensity and texture patterns.

One limitation of the current implementation is the relatively small dataset used for training and testing. Since deep learning models typically perform better with larger datasets, increasing the dataset size in future work could improve the model's accuracy and generalization capability.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.99      1.00      0.99       210
           1       1.00      0.99      0.99       219

    accuracy                           0.99       429
   macro avg       0.99      0.99      0.99       429
weighted avg       0.99      0.99      0.99       429

Classification Report:
              precision    recall  f1-score   support

    no_tumor       1.00      0.98      0.99       118
       tumor       0.98      1.00      0.99        97

    accuracy                           0.99       215
   macro avg       0.99      0.99      0.99       215
weighted avg       0.99      0.99      0.99       215
```

Fig.7: Classification Report of MobileNet and VGG16

VI. Future Scope

Although the proposed system demonstrates the feasibility of using deep learning for automated brain tumor detection, several improvements can be made to enhance its performance and practical usability. One important future direction is to train the model using a larger and more diverse MRI dataset. Deep learning models generally perform better when trained with large datasets, and incorporating more MRI images from different sources can improve the model's ability to generalize and increase classification accuracy. Another possible improvement is to extend the current binary classification system to multiclass tumor classification. Instead of only identifying tumor and non-tumor images, the model can be trained to classify different types of brain tumors, such as glioma, meningioma, and pituitary tumors, which would provide more detailed diagnostic information for medical professionals. Future work can also focus on integrating Explainable Artificial Intelligence (XAI) techniques, such as Grad-CAM or heatmap visualization, to highlight the regions of MRI images that influence the model's prediction. This would improve the interpretability of the system and help radiologists better understand how the model identifies tumor patterns. In addition, the system can be further developed into a real-time diagnostic application by improving the frontend interface and integrating the model into hospital or clinical systems. Using lightweight models such as MobileNet can also support edge or mobile deployment, allowing the system to run on portable devices for faster and more accessible tumor detection. Overall, these improvements can enhance the accuracy, reliability, and real-world applicability of the proposed brain tumor detection system.

VII. Conclusion

This paper presented a deep learning–based approach for automated brain tumor detection using MRI images. Two convolutional neural network models, VGG16 and MobileNet, were implemented using transfer learning to classify MRI scans into tumor and non-tumor categories. The MRI images were preprocessed and used to train both models, and their performance was evaluated on the test dataset. Experimental results showed that VGG16 achieved an accuracy of 99%, while MobileNet achieved an accuracy of 98%, demonstrating the effectiveness of deep learning techniques for brain tumor detection. Based on the comparative analysis, the VGG16 model showed slightly better performance and was selected for frontend deployment. The results indicate that the proposed system can effectively assist in automated brain tumor detection and has potential for further development in real-world medical diagnostic applications.

Acknowledgment

The authors would like to express their sincere gratitude to the Department of Electronics and Communication Engineering at Ahalia School of Engineering and Technology for providing the facilities and support required to carry out this project. The authors also wish to thank their project guide and faculty members for their valuable guidance, encouragement, and continuous support throughout the development of this work. Their suggestions and insights greatly contributed to the successful completion of this research.

References

[1] J. Han, "Brain tumor malignancy classification using improved VGG16 based on MRI images," *Applied and Computational Engineering*, Feb. 2024.

[2] S. Sutradhar, M. H. Rahman, N. Islam, and M. A. Yousuf, "Enhancing brain tumor classification in MRI using deep convolutional neural networks," *International Journal of Information Technology and Computer Science*, vol. 16, no. 3, pp. 12–21, Jun. 2024.

[3] X. Liu and Z. Wang, "Deep learning in medical image classification from MRI-based brain tumor images," Aug. 2024.

[4] A. Sharma, "Brain tumor detection using CNN and VGG-16 model," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 3, pp. 4203–4212, 2024.

[5] M. A. Khan and R. B. Z. Auvee, "Comparative analysis of resource-efficient CNN architectures for brain tumor classification," 2024

[6] S. Ganesh, R. Gomathi, and S. Kannadhasan, "Brain tumor segmentation and detection in MRI using convolutional neural networks and VGG16," *Cancer Biomarkers*, Mar. 2025.

[7] B. Shankar, K. Dharavath, K. V. Sridhar, E. C. Sekhar, and G. R. Kumar, "Deep-CNN based brain tumor classification from MRI images," in *Proc. International Conference on Computer Science and Communication Engineering (ICCSCE)*, 2025, pp. 1152–1165.

[8] S. Sarker, "Transfer learning and explainable AI for brain tumor classification using MRI data," 2025.

[9] R. Dulal and R. Dulal, "Brain tumor identification using improved YOLOv8," 2025.

[10] M. A. Khan et al., "Deep learning framework for brain tumor detection using CNNs and transfer learning on MRI scans," *Systems and Soft Computing*, vol. 7, 2025.

[11] R. Gupta, A. K. Singh and P. Jain, "Hybrid deep learning model for brain tumor classification using advanced CNN architectures," *International Journal of Medical Imaging and Health Informatics*, vol. 14, no. 1, pp. 89–101, Jan. 2024.

[12] H. Ali, S. U. Rehman, A. M. Malik, and F. Ali, "A comparative study of transfer learning models for brain tumor detection in MRI scans," *IEEE Access*, vol. 12, pp. 44578–44589, Apr. 2024.

[13] F. Ahmed, M. J. Islam, and S. A. Karim, "MobileNetV3 based lightweight classification for brain tumor detection from medical images," *Journal of Healthcare Informatics Research*, vol. 8, pp. 223–238, 2024.

[14] Y. Chen and L. Zhang, "Deep learning–based feature fusion for brain MRI classification," *Neural Processing Letters*, vol. 59, pp. 1235–1252, Aug. 2024.

[15] T. Roy, A. Ghosh, and D. Ghoshal, "Brain tumor detection using ensemble CNN models with attention mechanisms," *International Conference on Intelligent Systems and Machine Learning (ISML)*, pp. 230–239, 2025.

[16] M. Alshahrani, S. Qasim, and A. R. Alharbi, "Automatic brain tumor detection in MRI using deep residual networks," *Biomedical Signal Processing and Control*, vol. 85, 104899, 2025.

[17] A. Singh, P. Khatri and H. Singh, "Efficient deep learning framework for brain MRI classification using optimized convolutional layers," *IEEE Journal of Biomedical and Health Informatics*, vol. 29, 2025.

[18] L. M. Barros, J. P. Silva, and F. Silva, "Transfer learning for multiclass brain tumor classification with limited MRI datasets," *Computers in Biology and Medicine*, vol. 153, 106385, Feb. 2025.

[19] R. Das, S. Saha, and P. K. Mitra, "Explainable AI for brain tumor detection and localization using Grad-CAM with CNNs," *International Conference on Medical Imaging and Deep Learning Technologies (MIDLT)*, pp. 98–107, 2025.

[20] S. Verma and A. Saxena, "Deep learning approaches for brain tumor detection: A survey," *IEEE Transactions on Artificial Intelligence in Medicine*, vol. 3, pp. 310–332, 2025.

Design and FPGA Implementaion of a 32-bit RISC Processor Using Verilog HDL

Arya P R
dept. Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
arya.pr @ahalia.ac.in

Ashrin A
dept. Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
ashrin.a@ahalia.ac.in

Athish R
dept. Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
athish.r@ahalia.ac.in

Chaithra K C
dept. Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
chaithra.kc@ahalia.ac.in

Aneesh K
dept. Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
aneesh.k@ahalia.ac.in

V.Balamurugan
dept. Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
hodece.aset@ahalia.ac.in

***Abstract*— Creating a processor from the ground up is frequently one of the best strategies to comprehend how digital technology functions. In this work, an FPGA was used to create a basic 32-bit CPU that was built using Verilog HDL. A simple datapath connects the architecture's well-known parts, which include a Program Counter, Register File, Arithmetic Logic Unit, and Control Unit. Before being included into the entire processor, each module was created and tested separately. The design was built on a Xilinx Artix-7 FPGA utilizing the Basys 3 development board following functional simulation. Basic arithmetic and logical operations could be accurately performed by the CPU. With 25 LUTs, 29 slice registers, and 21 slices, hardware consumption stayed comparatively low. The architecture appears to be sufficiently efficient for FPGA-based CPU development, as seen by the measured power consumption of roughly 0.124 W.**

Keywords—FPGA, Basys 3, Artix 7, Verilog HDL.

I. INTRODUCTION

When considered as whole systems, processors can seem intimidating. Simpler concepts are frequently concealed by layers of pipelines, memory hierarchies, and optimization strategies. However, upon closer inspection, the majority of processors are still constructed from a limited number of well-known digital parts. Temporary values are stored in registers. Calculations are carried out by arithmetic logic units. Control logic decides what happens next. For students in Electronics and Communication Engineering, these pieces are usually introduced through diagrams on a classroom slide. That explanation helps, of course. Still, something feels slightly abstract until the design actually runs on hardware.

In light of this, the current work investigates the use of Verilog Hardware Description Language in the design of a custom 32-bit processing unit. A Program Counter, Register File, Arithmetic Logic Unit, and Control Unit are among the key components of the architecture that are connected by a common data path. Initially, every block was created and tested separately in a simulation. They weren't assembled into the entire processor until those components performed as anticipated. The final design was then put into practice on an FPGA platform, which made it possible to test the processor outside of a simulator window and under actual hardware conditions.

The processor created here is undoubtedly much simpler than those found in commercial designs. Advanced memory management features, branch prediction mechanisms, and complex pipelines are all present in modern processors like the ARM Cortex-A53. For a small academic project, it would be impossible to replicate that degree of complexity. The fundamental ideas are surprisingly similar, though. Instructions continue to travel via a datapath. Registers and functional units continue to exchange data. These internal interactions are made easier to see and comprehend by using a simplified 32-bit processing unit. In this way, the FPGA platform functions as both a testing environment and a hands-on learning environment where processor concepts progressively become more concrete.

II. EXISTING SYSTEMS

This work described a 32-bit Reduced Instruction Set Computing (RISC) processor built in Verilog HDL on an FPGA platform. Their design adheres to a Harvard architecture, which keeps data and instruction memories apart. This choice typically facilitates memory access and increases execution efficiency. An instruction fetch unit, a decoding stage, an execution unit, and dedicated memory modules are among the well-known components that make up the processor itself. Together, these components create a fully synthesizable hardware architecture. The Xilinx ISE design environment is used for implementation, and ISim simulation is used to confirm functional behaviour. The authors also use VIO IP cores, which enable direct hardware monitoring and debugging, to further observe internal signals during operation. The architecture is still relatively simple when compared to contemporary commercial processors, but it shows a useful and effective method for FPGA-based RISC processor implementation. The overall performance is primarily assessed in terms of hardware resource utilization, timing characteristics, and the maximum operating frequency. Our project, which develops and implements a similar modular processor architecture on an FPGA platform, benefits greatly from this work [1].

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

This paper designed a 32-bit MIPS processor implemented on an FPGA platform using VHDL. A five-stage pipeline comprising instruction fetch, instruction decode, execution, memory access, and write-back is used in their architecture, which adheres to RISC design principles. To handle problems that occur during pipelined instruction execution, the processor integrates important modules like the datapath, control unit, and hazard detection unit. In order to process multiple data elements at once and increase computational efficiency, the authors also extend the architecture to a multi-core configuration using SIMD concepts. The Virtex-6 FPGA is used to implement the design using the Xilinx ISE design environment, and hardware utilization, operating frequency, and power consumption are used to assess performance. The goals of the current project, which focuses on the design and FPGA implementation of a 32-bit RISC processor using Verilog HDL, are closely related to the work's use of FPGA-based processor implementation and modular datapath organization, even though it stresses a multi-core MIPS architecture [2].

This design of a 32-bit MIPS pipelined processor was built using Verilog HDL who also assess the processor's behaviour with regard to timing, power consumption, and architectural efficiency. In contrast to a non-pipelined system, their work demonstrates how splitting instruction execution into several phases increases throughput while lowering critical route delay. The processor architecture shows how hardware description languages may be used to build processor architectures on FPGA platforms by integrating essential elements like the instruction fetch unit, register file, ALU, and memory access stage within a structured datapath. The architectural concepts and modular datapath structure offer helpful information into the development and FPGA realization of a 32-bit RISC processor utilizing Verilog HDL, even if the study primarily focuses on pipelining and performance enhancement [3].

This work demonstrates a design of a 16-bit RISC processor based on the MIPS architecture, written in Verilog HDL and implemented on an FPGA platform, is described in [4]. A five-stage pipeline serves as the foundation for the processor. Different instructions can be processed simultaneously rather than waiting in line thanks to the steps that instructions go through, including retrieve, decode, execute, memory access, and finally write back. Xilinx tools are used to synthesize the design, which is then implemented on a Spartan-6 FPGA device so that hardware testing and simulation may be used to analyse its behaviour. The architecture isn't particularly big or intricate. Nevertheless, there is a reason behind that simplicity. The design achieves respectable performance while staying feasible for FPGA implementation by maintaining the datapath's modularity and the instruction flow's clarity. This easily leads to processor designs that expand the same concepts to larger data widths [4].

This paper looks at the design of a 16-bit Arithmetic Logic Unit from a VLSI perspective, mainly with the usual engineering trade-offs in mind: speed, power consumption, and chip area. Rather than sticking to a single logic style, the design quietly mixes a few different ones. CMOS handles the basic logic functions, pseudo-NMOS appears in parts of the AND logic, while pass transistor logic is used in multiplexers. The choice seems fairly pragmatic. After all, in circuit design one approach rarely fits every module perfectly. Much of the attention goes to the adder structure, which is hardly surprising since arithmetic operations depend on it so heavily. The carry skip adder seems to provide a reasonable trade-off between hardware cost and latency among the configurations examined. DSCH is used to verify the circuit, and then Micro Wind is used to implement the layout. The work focuses on a 16-bit ALU, but when considering bigger digital systems, such as FPGA-based 32-bit RISC processor implementations created using Verilog HDL, the principle of blending logic styles and carefully choosing adder topologies still seems pertinent [5].

III. PROPOSED DESIGN

A. Architectural overview and system organization

Instead of recreating the intricacy of commercial processors, the core is a specially designed 32-bit processor on a Field-Programmable Gate Array (FPGA) using Verilog HDL. Its purpose is to provide a straightforward and useful model of how processors interact in hardware. It soon becomes clear how useful fundamental components may be when creating such a system from the ground up. As a result, a modular data path structure is used in the architecture. The processor is constructed as a number of smaller functional blocks (modules) that cooperate to carry out instructions rather than as a single, huge block. Program Counter (PC), Register File, Arithmetic Logic Unit (ALU), Control Unit, and many more auxiliary data channel connections are the main parts. When combined, they enable the processor to retrieve, understand, and carry out instructions and the functions that go along with them, even if each one has a minor but vital role.

The data path is ultimately what ties all the parts together. Under the direction of control circuitry, data is transferred between registers, compute blocks, and memory interfaces via the datapath. At first, each module was created separately at the Register Transfer Level (RTL). Each module was integrated into a single processing core following testing and successful debugging. Debugging the modules was made simple by this gradual approach.

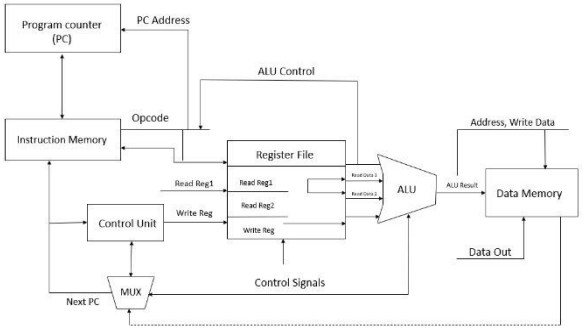


Fig 1: Architecture of the proposed processor

B. Datapath organization

The datapath in a processor basically shows how data is moved around the device. In order to transport data between the program counter, register file, ALU, and many intermediate buses that will be utilized to transfer operands and results, the suggested design would establish the necessary connections. Even the most basic machines would be seriously hampered in their ability to carry out instructions in a systematic way without such connectivity. When the program counter gives the address of the next instruction in the program, the procedure starts. The control unit then retrieves this instruction, decodes it, and generates the control signals needed for the rest of the datapath. These control signals will specify which registers to access, what kind of ALU operation to carry out, and where to transmit the data that the ALU will produce after completing its operations.

Data must be routed with extreme caution in order to pass across these different modules. The operands from the register file to the ALU and the results from the ALU back to storage are routed via multiplexers and internal buses. On paper, the method of routing data via the various components can seem straightforward, but during the design implementation phase, the timing of coordinating the many components in order to execute an instruction was
crucial.The datapath will provide an orderly and predictab le sequence when carrying out instructions after all compo nents are correctly synchronized.

C. Control unit designs

While the datapath merely transports the processed data through hardware, a processor's control unit manages the various stages of data processing. Because it manages all processors activity from start to finish, the control unit is a crucial component of a processor. The control unit receives all printed instructions from software. Depending on the type of instruction, it can generate control signals that guide other processing devices to carry out particular tasks. For instance, if two processor registers require data from memory, the control unit will provide the two control signals that instruct the ALU (Arithmetic Logic Unit) on what computation to carry out and which way to convey the results. The processor would look as a collection of unconnected parts if the control unit didn't coordinate these operations.

Combinational logic as defined in Verilog HDL was used in this design to create the control unit. It made design sense to keep the control logic and datapath apart; this enables for a more structured architecture with distinct software and processor hardware evolution. It is far more likely that only control logic—rather than all execution hardware—will be impacted if a new instruction becomes available to the CPU in the future.

D. Arithmetic and logical unit(ALU)

The processor's computational core is the ALU. Almost all significant operations eventually go via it. The ALU carries out logical operations whenever two operands need to be added, compared, or combined. The ALU in this implementation uses 32-bit inputs that are taken from the register file. A limited but typical collection of mathematical and logical operations, such as addition, subtraction, logical AND, and logical OR, are supported. The control unit generates control signals that are used to choose these operations. Behavioural modelling approaches were used to write the design itself in Verilog. Practically speaking, clarity was more important than extreme optimization. The value is routed back through the datapath and put in the proper register once the ALU generates its result. It was easier to comprehend the general behaviour of the CPU when this flow was observed during simulation.

Table 1: Opcode and it's corresponding operations

SI. No	Opcode	Operation
1.	0000	Addition
2.	0001	Subtraction
3.	0010	Logical AND
4.	0011	Logical OR

E. Register file design

In addition to the data utilized in computations, general-purpose registers are temporary data storage components that the processor uses to store two or more operands, the outcomes of operations, and intermediate computation results. In order to enable the ability to read a source general-purpose register and write a destination general-purpose register during each CPU clock cycle, the register file has dual read/write ports, which permit two accesses to data contained within a register.

During each clock cycle, the control unit gets signals from the processor that specify which register, based on address selection, will be used to process instructions. Registers, which provide access to data 100 times faster than any other sort of storage, are frequently the most significant hardware components related to processor speed, despite the fact that they may appear to be relatively simple hardware parts in the processor architecture. A single clock cycle will carry out a number of instructions; an operand retrieved from the register file will be delivered to the ALU for processing, and once processing is finished, it will be written back to the register file.

F. Program counter and instruction flow

The instruction being carried out is monitored by the program counter (PC). The PC indicates the location of the subsequent instruction in the instruction stream by adding 1 to its current value with each clock cycle. Although upgrading the PC in a sequential manner is simple (at least initially), it is the main factor that determines how instructions are processed by the processor. The PC is designed with a synchronous reset feature. When the system is initialized, it allows the CPU to start operating from a known place by setting the PC to a predefined beginning address. The PC will progressively move

through the list of instructions (sequence) once the processor has begun to operate. A new instruction is added to the execution pipeline with each clock cycle in this straightforward yet incredibly efficient procedure, which enables the processor to function in a predictable and ordered manner.

G. FPGA implementation and hardware verification

The initial step was to use Xilinx Vivado to simulate the RTL blocks in order to verify how each module should function. To verify the interaction between all modules, a small testbench runs different instructions and then watches how signals change as they go through clock cycles. The Xilinx Artix-7 FPGA was then used to generate and construct the full processor in the Basys 3 FPGA board. Both the arithmetic and logic processes performed correctly during timed operation, according to the hardware testing results. Reasonable resources were employed for both the design and implementation, according to the synthesis and timing reports. Even though the architecture is intentionally simplistic, it may be readily used as a foundation for more complicated designs.

IV. RESULT AND DISCUSSION

A. Simulation results and verification

A Verilog testbench was used to simulate and confirm the processor design's functionality. A portion of the time relationship between the several signals is depicted in the waveform in Figure 2. When it comes to activity, these signals initially seem to follow a pretty natural pattern: the clock (clk) toggles at regular intervals and serves as a time reference for all internal processor actions. The system is initialized for a while using the reset (rst) signal. All of the processor's registers and control logic stabilize during this short time, and the CPU doesn't do anything significant until the reset is de-asserted.

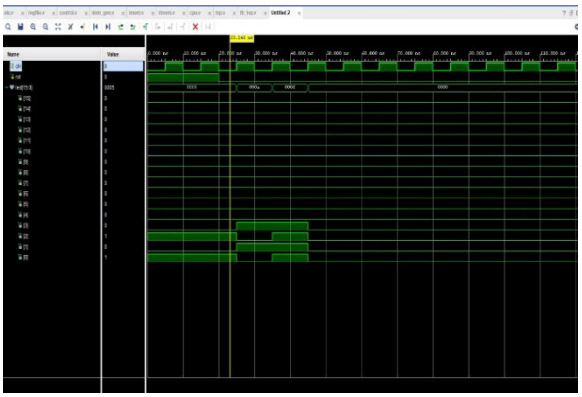


Fig 2: Simulation waveform of the proposed 32-bit processor showing the clock signal(clk), reset signal(rst) and LED output bus (led [15;0].

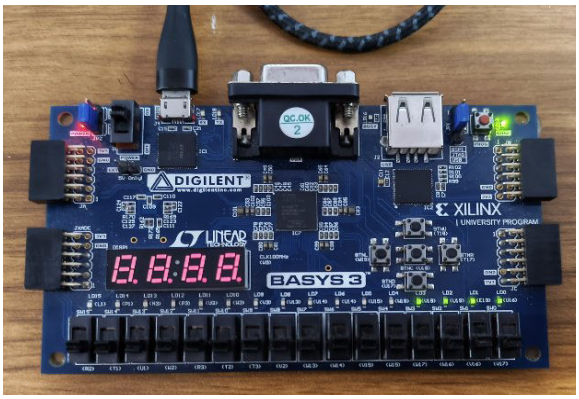
The CPU begins processing sequential instructions that contain data from the instruction memory during system setup. A view of the data kept in registers is provided via the value on the output bus (led [15:0]). A few clock cycles after the next instruction is carried out, the output bus will display a value of 0x000A, which will soon change to 0x000F. The arithmetic computations carried out by the arithmetic logic unit (ALU) during the execution of instructions in the processor's data channel are represented by the values on the output bus rather than being random.

Further examination of the waveform revealed that the most of the activity was on the lower portion of the LED bus, which is where the bits are situated. These lower bits typically change as results are computed and written back to their registers. There aren't many state changes in the upper part of the LED bus, which houses the bits, during the process. The behavior of the higher bits can be seen positively; rather than having an unexpected toggle throughout the bus, it indicates that data propagation and updates to registers are regulated and predictable.

The output is steady and the waveform is once more comparatively quiet when the instruction execution sequence is finished. The analysis above shows that the ALU, control unit, and memory all function as they should. However, since time and the actual physical processor will need to be verified, the simulation cannot ensure every aspect of the CPU's operation. However, the waveform can offer some degree of assurance regarding the functionality of the suggested processor's main components.

B. Hardware implementation and validation

In order to view the features of the simulated processor in a real-world setting, physical hardware must authenticate the simulation results. The suggested designer processor for monitoring the real-time execution of the instruction sequence is synthesized and implemented using the Digilent Basys 3 FPGA Board, which makes use of a Xilinx Artix-7 FPGA device. Switches, LEDs, pushbuttons, and a seven-segment display are among the components on the development board that will allow for



direct monitoring of the processor's operation.
Fig 3: Real-time hardware implementation of the proposed processor on Basys 3 FPGA board

The implemented design was carried out on the FPGA device once the configuration bitstream was generated. The CPU carried out the instruction sequence in real time and in sync with the network clock after setting up the board. Through the accessible development board peripherals, we were able to view the output signals generated by the processor's datapath and compare the processor's performance with simulation results. Each of the CPU architecture's modules performed as predicted

when incorporated into the FPGA platform, as evidenced by the responses we saw during hardware verification.

C. Power analysis

The power consumption was determined with the help of power analysis obtained after the simulation on Xilinx Vivado which is illustrated in Fig 4. The report produced gives an approximate measure of the total power that the FPGA will consume during its operational state. The power report shows that the total on-chip power consumed is approximately 0.124 watts. Of that, approximately 0.052 watts is dynamic power, and 0.072 watts is static power for the device.

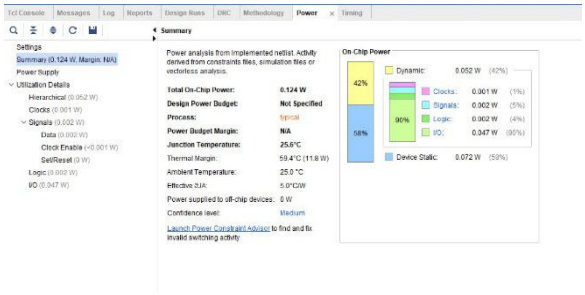


Fig 4: Power analysis of the implemented processor designed in Xilinx Vivado

There seems to be a pattern when examining only the dynamic part of the power use. The I/O circuitry produces about 0.047 watts of dynamic power, which is the majority. The signal activity and the logic components individually produce about 0.002 watts, while the clock network produces about 0.001 watts. About 58% of total power consumption is static power in percentage terms, while the other 42% are dynamic power. It matches what is usually seen in the case of FPGA implementation, where the static power part is still quite prominent even after the design has shrunk to a fairly small size.

Low power dissipation generally indicates efficient energy utilization, and the modular design of the processor's datapath contributes to a manageable and well-defined logic layout that requires little switching. However, an FPGA-based basic prototype will typically consume less power than an optimized commercial processor. When compared to earlier stated systems, this architecture appears to offer a respectable trade-off between functionality, hardware utilization, and power economy. In general, these results imply that the adopted processor architecture is quite appropriate for application in embedded system based on FPGA and educational experiments. Though this architecture is not made to rival industrial-scale processors, its total mix of minimal power usage and simply defined architecture could measure up to a working solution in prototype systems and educational settings.

CONCLUSION

A CPU architecture was developed and implemented using Xilinx Vivado and an FPGA development environment. The simulation and implementation results verified that the processing modules function as anticipated. As a result, the results demonstrate that this design has achieved a balanced performance in terms of power economy, functionality, and resource consumption, making it an appropriate FPGA-based application platform for embedded and educational applications.

Table 2: FPGA resource and power comparison table

Parameters	Reference [4]	Proposed processor
Power consumption	0.014W	0.124W
LUT'S Used	126	25
Slice registers	107	29
Utilized slices	41	21

The suggested processor greatly lowers FPGA resource use, according to the comparison results. A more compact architecture is indicated by the drop in slice LUTs from 126 to 25 and slice registers from 107 to 29. Additionally, there are just 21 slices used instead of 41. The increased I/O activity is the primary cause of the overall power consumption, which is higher than the reference design but still within acceptable bounds. Overall, the suggested CPU shows enhanced hardware use and area efficiency.

REFERENCES

[1] Archana, H. R., et al. "System verification and analysis of ALU for RISC processor." 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). Vol. 1. IEEE, 2021. https://doi.org/10.1109/ICACCS51430.2021.9442045

[2] S. M. Al-Sudany, A. S. Al-Araji, and B. M. Saeed, "FPGA-Based Multi-Core MIPS Processor Design," Iraqi Journal of Computers, Communications, Control & Systems Engineering (IJCCCE), vol. 21, no. 2, pp. 16–34, Jun. 2021, DOI: 10.33103/uot.ijccce.21.2.2.

[3] S. Kumar and B. Bhushan, "Qualitative Analysis of 32-Bit MIPS Pipelined Processor," Int. J. Eng. Res., vol. 9, no. 5, pp. 558–561, 2020.
https://doi.org/10.17577/IJERTV9IS050484

[4] M. B. B., V. H., K. M. G., T. V and A. R., "Design and Implementation of RISC MIPS Processor on FPGA," Int. J. for Research in Applied Science & Engineering Technology (IJRASET), vol. 11, no. 4, pp. 1406–1410, Apr. 2023. https://doi.org/10.22214/ijraset.2023.50352

[5] N. Ravindran and R. M. Lourde, "An optimum VLSI design of a 16-bit ALU," in Proceedings of the 2015 International Conference on Information and Communication Technology Research (ICTRC), Abu Dhabi, UAE, pp. 52-55, Feb. 2015.
https://doi.org/10.1109/ICTRC.2015.7156419

[6] G. Surekha, G. Madesh, M. P. Kumar and H. Sriramoju, "Design and Implementation of Arithmetic and Logic Unit (ALU)," 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 1530-1536, DOI: 10.1109/ ICAAIC56838.2023.10140574.

[7] A. H. R., S. T., B. H. T., and S. S. V., "System Verification and Analysis of ALU for RISC Processor," Proc. ICACCS, B.M.S. College of Engineering, Bangalore, India, 2021.

[8] S. M. Al-Sudany, A. S. Al-Araji, and B. M. Saeed, "FPGA-Based Multi-Core MIPS Processor Design," Iraqi Journal of Computers, Communications, Control & Systems Engineering (IJCCCE), vol. 21, no. 2, pp. 16–34, Jun. 2021, DOI: 10.33103/uot.ijccce.21.2.2.

[9] S. Suresh and R. Ganesh, "FPGA Implementation of MIPS RISC Processor," International Journal of Engineering Research & Technology (IJERT), vol. 3, no. 1, pp. 1710–1714, Jan. 2014.

A Compact Clover-Shaped Slot-Loaded Microstrip Patch Antenna for Sub-6 GHz Applications

Nandana N
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
nandana.n@ahalia.ac.in

Rijul R Das
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
rijul.r.das@ahalia.ac.in

Shakira U
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
shakira.u@ahalia.ac.in

Sivani P S
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
sivani.ps@ahalia.ac.in

Mr. Sanish V S
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
sanish.vs@ahalia.ac.in

Dr. V. Balamurugan
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
hodece.aset@ahalia.ac.in

Abstract— A compact single-element antenna for wireless communication applications operating in the sub-6 GHz frequency band is presented. The antenna is designed on a low-cost dielectric substrate to achieve efficient impedance matching, adequate bandwidth, and stable radiation characteristics. Performance evaluation is carried out in terms of reflection coefficient (S11), gain, and radiation pattern. The results indicate that the antenna maintains an S11 value below −10 dB across the desired operating band, ensuring effective power transfer. The radiation pattern is observed to be nearly omnidirectional with satisfactory gain, making it suitable for practical wireless communication systems. Owing to its simple structure, compact size, and ease of fabrication, the proposed antenna is well suited for integration into modern communication devices. In addition, the design can be extended as a fundamental element for advanced multi-element antenna systems, including MIMO configurations.

Key Words: Single-element antenna, Sub-6 GHz, Reflection coefficient (S11), Impedance matching, Radiation pattern, Gain, Wireless communication, Compact antenna, MIMO systems.

I. INTRODUCTION

The swift progress of wireless communication technologies has necessitated the development of antennas that are both compact , efficient , and capable of delivering high performance within the sub-6 GHz frequency range. With the widespread use of modern communication systems such as 4G and the continuous development of 5G networks, antennas are expected to provide reliable connectivity, sufficient bandwidth, and stable radiation behavior, while fitting into increasingly smaller devices.

Among the diverse array of antenna configurations, single-element antennas continue to be extensively utilized due to their straight forward design, ease of fabrication, and cost- effectiveness. They are commonly employed as standalone radiators in many wireless devices and also act as the basic building blocks for more advanced systems, such as Multiple-Input Multiple-Output (MIMO) antennas. The effectiveness of these antennas is typically assessed using key parameters such as reflection coefficient (S11), impedance matching, gain, bandwidth, and radiation pattern.

However, designing a compact antenna without compromising its performance remains challenging . The overall characteristics of an antenna are strongly influenced by factors such as the substrate material, geometry of the structure, and feeding technique. Therefore, careful design and optimization are essential to ensure that the antenna performs efficiently within the desired frequency range.

Considering these factors, a compact single-element antenna operating in the sub-6 GHz band is developed and analyzed. The design focuses on achieving good impedance matching, stable radiation characteristics, and acceptable gain, making it suitable for practical wireless communication . In addition, the proposed antenna can be further extended to form multi-element configurations for advanced systems in the future.

II. LITERATURE REVIEW

The rapid growth of wireless communication systems, especially in the sub-6 GHz frequency range, has created a strong demand for antennas that are compact, efficient, and capable of delivering high performance. Among the various antenna types, microstrip patch antennas are widely preferred due to their low profile, light weight, and ease of fabrication. Despite these advantages, conventional microstrip antennas suffer from certain limitations such as narrow bandwidth and relatively low gain, which make them less suitable for modern high-speed communication systems. To overcome these issues, researchers have explored several design improvements, including slot loading, the use of advanced substrate materials, and multi-element configurations to enhance overall antenna performance.

The fundamental concepts of antenna design are largely based on classical studies in antenna theory. For example, Constantine A. Balanis provided detailed insights into radiation patterns, gain, and efficiency, which serve as a strong foundation for antenna analysis [1]. Similarly, David M. Pozar discussed the working principles of microstrip antennas and highlighted their key advantages along with inherent bandwidth limitations [2]. Further advancements in compact and broadband antenna designs were presented by Kin-Lu Wong, who introduced techniques to improve impedance bandwidth and radiation characteristics [3]. In addition, works by Garg et al. [4] and James and Hall [5] provided practical design approaches and guidelines that are still widely used in antenna development.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

To address the bandwidth limitations, several enhancement techniques have been introduced over the years. Wong and Wong [6] discussed methods such as slot incorporation and stacked patch designs, which significantly improve antenna bandwidth and efficiency. These techniques have become the basis for many modern antenna structures, particularly slot-loaded designs that offer better impedance matching and wider frequency coverage. In recent times, the integration of Multiple-Input Multiple-Output (MIMO) technology has further enhanced antenna performance by increasing system capacity and improving link reliability.

MIMO antenna systems are especially important in modern wireless communication due to their ability to handle multiple signal paths and reduce multipath fading. Mohamed S. Sharawi highlighted key challenges in MIMO antenna design, including mutual coupling and compactness, while also discussing future research directions [9]. Similarly, Andreas F. Molisch emphasized the role of accurate channel modeling and diversity techniques in improving MIMO system performance [8]. Blanch et al. [7] proposed methods to evaluate diversity performance using antenna parameters, which helps in enhancing signal reliability. Furthermore, Chen et al. [10] reviewed the issue of mutual coupling in MIMO systems and suggested various techniques to improve isolation and overall system efficiency.

Overall, these studies highlight the importance of innovative antenna design strategies and multiport configurations in meeting the demands of modern wireless applications. Building on this foundation, the proposed work focuses on a slot-loaded microstrip patch antenna integrated with a multiport MIMO configuration. This approach aims to achieve an improved bandwidth, better impedance matching, and enhanced radiation performance, making it suitable for sub-6 GHz wireless communication systems.

III. SYSTEM ANALYSIS

1. EXISTING SYSTEM

In conventional wireless communication systems, single-element antennas such as monopole, dipole, and basic microstrip patch antennas, are commonly employed. These antennas are simple in structure, easy to design, and cost-effective to fabricate, making them suitable for earlier generations of wireless communication. Among these, the microstrip patch antenna have gained popularity owing to their planar profile, low weight, and compatibility with printed circuit technology. Existing single-element antennas designed for sub-6 GHz operation are generally used in Wi-Fi, LTE, and early 5G applications. Although they are widely adopted, their performance is often limited to narrow frequency bands, and their efficiency may not be sufficient for modern high-data-rate systems. Furthermore, these designs are usually optimized for basic wireless applications without considering the advanced requirements of today's communication networks, such as high reliability, low latency, and wideband operation.

2. DISADVANTAGES OF EXISTING SYSTEM

Although single-element antennas have been useful for decades, they face significant drawbacks which evaluated based on the performance demands of modern communication technologies. The most common limitation is their narrow bandwidth, which restricts operation to a limited range of frequencies. This renders them unsuitable for wideband or multi-standard systems that operate in the sub-6 GHz spectrum. Another issue is their low gain and efficiency, which reduce the quality of communication in long-distance or high-capacity scenarios. Additionally, single-element antennas have a limited ability to handle multipath fading, because they cannot exploit spatial diversity or diversity gain. This results in poor reliability in environments with multiple reflections, such as urban areas. In compact wireless devices, size constraints also make it challenging to achieve an optimized performance without affecting impedance matching and radiation characteristics. These disadvantages highlight the need for a carefully optimized design that addresses bandwidth, gain, and reliability within the sub-6 GHz band.

3. PROPOSED SYSTEM

The proposed system is a single-element clover-shaped slot-loaded microstrip patch antenna designed to operate in the sub-6 GHz frequency range. It is built using a rectangular microstrip patch placed on a dielectric substrate and fed by a 50 Ω microstrip feed line. To improve its performance, the patch is modified by adding symmetrical circular and curved slots, which together form a clover shape. This design enhances the antenna performance while keeping it compact and suitable for modern wireless devices. The clover shaped slots play an important role by changing the flow of current on the surface of the patch. This increases the effective path of the current and creates several resonant frequencies. Because of this, the antenna provides better bandwidth, good impedance matching (S11 < -10 dB), and stable radiation performance. Owing to these advantages, the antenna can be effectively used in 4G, 5G sub-6 GHz systems, IoT devices, and other wireless communication applications, making it a simple and efficient design for improved performance.

4. ADVANTAGES OF THE PROPOSED SYSTEM

The proposed single-element clover-shaped slot-loaded microstrip patch antenna offers several advantages for modern wireless applications. Its compact and low-profile structure makes it easy to use in small devices without adding extra size or complexity. The clover-shaped slots improve the flow of current on the patch, which helps increase bandwidth and achieve better impedance matching. This ensures efficient signal transmission with minimal power loss.

In addition, the antenna offers stable radiation performance and can operate at multiple frequencies, making it suitable for sub-6 GHz applications like 4G, 5G, and IoT systems. Overall, the design is simple, cost-effective, and capable of delivering improved performance, making it a practical choice for current wireless communication needs.

IV. METHODOLOGY

The proposed antenna was developed in three phases: Designing, Fabrication, and Testing.

I. DESIGNING PHASE

In this phase, the antenna was designed to operate in the sub-6 GHz frequency range. First, the basic dimensions of the rectangular patch, feed line, and substrate were calculated using standard formulas. Clover-shaped slots are added to the patch to improve the bandwidth and overall performance. The antenna model was created and simulated using software , such as HFSS or CST Microwave Studio. Important parameters such as return loss (S11), VSWR,

gain, and radiation pattern were checked, and the design was adjusted until the required performance was achieved.

II. FABRICATION PHASE

Once the design is finalized, the antenna is fabricated on a suitable substrate, such as FR4. The design is transferred onto the substrate using common methods such as PCB etching. The feed line and ground plane were properly formed, and an SMA connector was attached to provide input to the antenna. Care is taken during fabrication to maintain accurate dimensions and proper connections to avoid performance issues.

III. TESTING PHASE

In the final phase, the fabricated antenna was tested to evaluate its performance. A Vector Network Analyzer (VNA) was used to measure parameters like return loss (S11) and VSWR. If available, the radiation pattern and gain were also measured in an anechoic chamber. The test results were then compared with the simulation results to ensure that antenna operate as expected. If there are any differences, they are analyzed and improvements can be made.

V. RESULTS

In this section, the performance of the proposed single-element clover-shaped slot-loaded microstrip patch antenna is presented and discussed. The results were obtained through both simulation and practical testing to determine how well the antenna works in the required sub-6 GHz frequency range. Key parameters such as return loss (S11), VSWR, gain, bandwidth, and radiation pattern were analyzed to understand its overall performance.

The simulated results were then compared with the measured results to verify the accuracy of the design and fabrication are. If there are any differences, they are examined to find possible reasons, such as small fabrication errors or testing conditions. This section clearly demonstrates that the proposed antenna design is efficient and suitable for wireless communication applications.

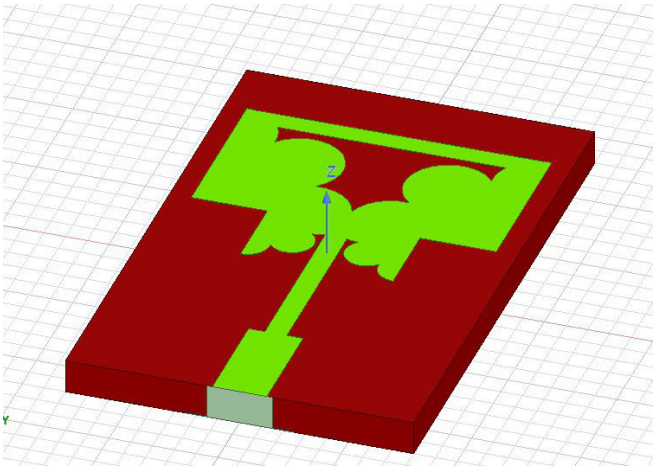


Fig. 1. Side View

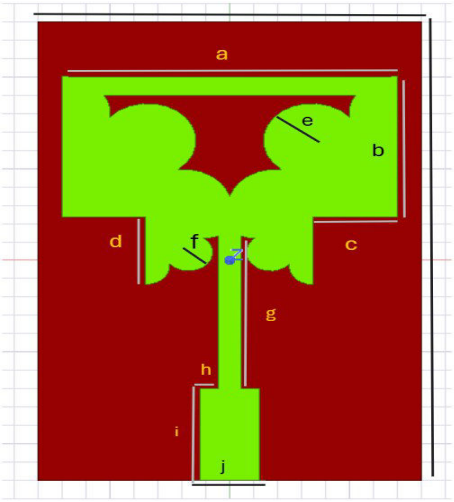


Fig. 2. Top View

a	13.38
b	7.65
c	3.5
d	3.68
e	2(Arc Radius)
f	1(Arc Radius)
g	8.3
h	0.75
i	5
j	3
k	16
l	25

Table.1..Antenna Parameters in mm

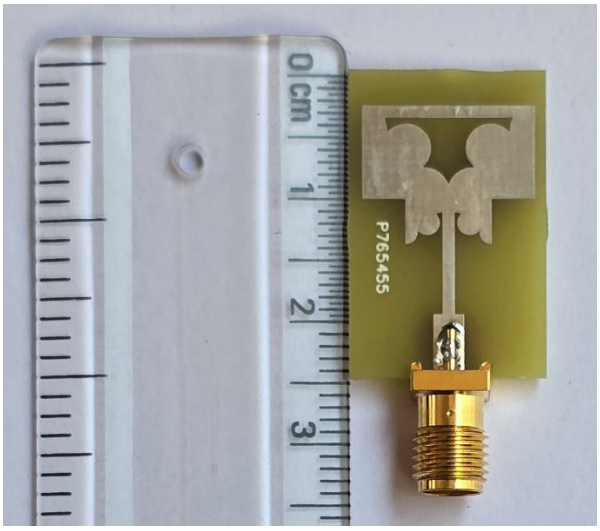


Fig.3.Fabricated Antenna

1. S- PARAMETER ANALYSIS

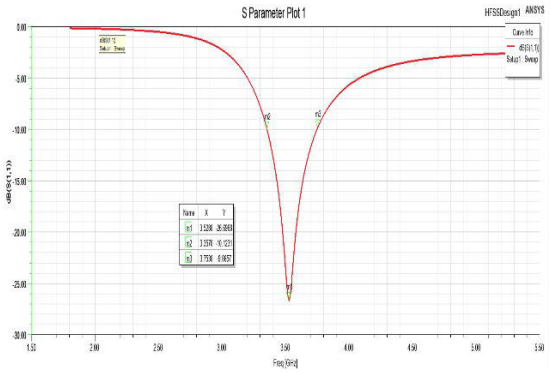


Fig.4. S- Paramter Plot

From the S-parameter (S11) graph, the antenna exhibits a resonance at approximately 3.52 GHz with a return loss of approximately -26.69 dB, indicating excellent impedance matching. This very low S11 value indicates that almost all the input power is effectively radiated, with negligible reflection. This clearly shows that the antenna is well matched within the desired sub-6 GHz frequency range. This result confirms that the clover-shaped slot design significantly enhances the antenna performance by achieving strong resonance and improved efficiency. Hence, the antenna is highly suitable for wireless applications such as 4G, 5G, and IoT systems, providing reliable and efficient signal transmission.

2. VSWR

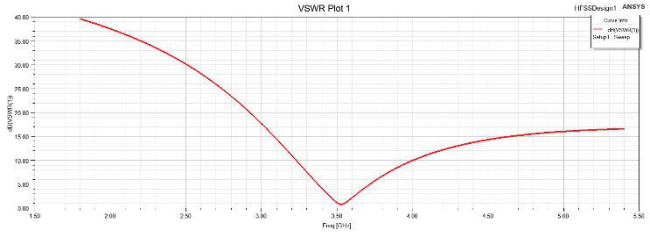


Fig. 5. VSWR Plot

From the VSWR plot, the antenna shows a minimum VSWR value close to 1 at approximately 3.52 GHz, which indicates excellent impedance matching at the resonant frequency. A VSWR value near 1 indicates that almost all the input power is transmitted with very little reflection. As the frequency moved away from 3.52 GHz, the VSWR increased significantly, indicating higher mismatch and reduced efficiency. This behavior confirms that the antenna is well optimized for operation at 3.52 GHz within the sub-6 GHz band. Overall, the VSWR results validate that the proposed clover-shaped slot-loaded antenna provides efficient performance with minimal losses at the desired operating frequency, making it suitable for wireless communication applications.

3. GAIN

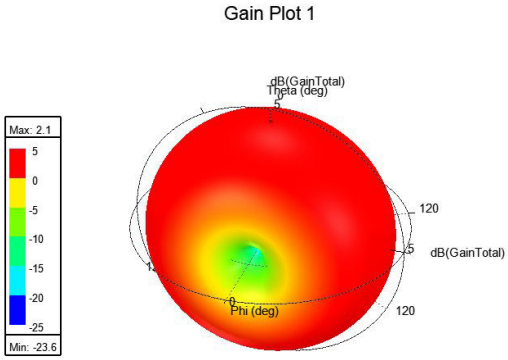


Fig.6. 3D Gain Pattern

From the gain plot, the antenna shows a maximum gain of approximately 2.1 dB, which indicates a moderate level of radiation suitable for short- to medium-range wireless applications. The radiation pattern is almost omnidirectional, meaning the antenna radiates energy in most directions, with slightly stronger radiation in the main (broadside) direction. The smooth variation in the gain pattern indicates that the antenna exhibits stable radiation behavior without any major distortions. Although the gain is not very high, it is quite reasonable for a compact microstrip antenna design, especially in the sub-6 GHz range. Overall, the antenna delivers steady and reliable performance, making it a good choice for applications ,such as 4G, 5G, and IoT communication systems.

4. RADIATION PATTERN

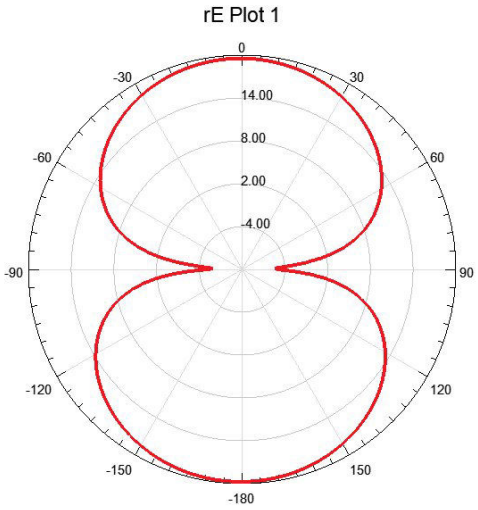


Fig.7. 2D Radiation Pattern

From the radiation pattern plot, the antenna shows a nearly omnidirectional radiation behavior, meaning it radiates signals in most directions rather than focusing on a single direction. The main radiation was concentrated in the broadside direction, which is typical of microstrip patch antennas. The pattern was smooth and uniform, indicating stable and consistent radiation without significant distortions or side lobes. This type of radiation is useful for wireless communication systems where signal coverage in multiple directions is required. Overall, the antenna provides reliable radiation characteristics, making it suitable for applications such as 4G, 5G, and IoT devices.

VI. CONCLUSION

The proposed single-element clover-shaped slot-loaded microstrip patch antenna was successfully designed, fabricated, and tested for use in sub-6 GHz wireless applications. By adding clover-shaped slots to the patch, the antenna's performance is improved in terms of bandwidth and impedance matching. The S-parameter results show a clear resonance at approximately 3.52 GHz with a very low return loss, and the VSWR is close to 1, which means that antenna transmits power efficiently with minimal loss.

The antenna also exhibited stable radiation characteristics with a nearly omnidirectional pattern and a gain of approximately 2.1 dB, making it suitable for real-world wireless communication. Overall, the design is simple, compact, and cost-effective while still providing reliable performance. Because of these advantages, the antenna can be effectively used in applications ,such as 4G, 5G sub-6 GHz, and IoT systems.

VII. REFERENCES

1. C. A. Balanis, Antenna Theory: Analysis and Design, 4th ed. Hoboken, NJ, USA: Wiley, 2016.
2. D. M. Pozar, "Microstrip antennas," Proceedings of the IEEE, vol. 80, no. 1, pp. 79–91, Jan. 1992.
3. K. L. Wong, Compact and Broadband Microstrip Antennas. New York, NY, USA: Wiley, 2002.
4. R. Garg, P. Bhartia, I. Bahl, and A. Ittipiboon, Microstrip Antenna Design Handbook. Norwood, MA, USA: Artech House, 2001.
5. J. R. James and P. S. Hall, Handbook of Microstrip Antennas. London, U.K.: Peter Peregrinus Ltd., 1989.
6. H. Wong and K. L. Wong, "Bandwidth enhancement techniques for microstrip antennas," IEEE Antennas and Propagation Magazine, vol. 44, no. 4, pp. 52–59, Aug. 2002.
7. S. Blanch, J. Romeu, and I. Corbella, "Exact representation of antenna system diversity performance from input parameter description," Electronics Letters, vol. 39, no. 9, pp. 705–707, May 2003.
8. A. A. Glazunov, A. F. Molisch, and F. Tufvesson, "Physical modeling of MIMO channels," IEEE Transactions on Antennas and Propagation, vol. 59, no. 6, pp. 2026–2037, Jun. 2011.
9. M. S. Sharawi, "Current misuses and future prospects for printed multiple-input, multiple-output antenna systems," IEEE Antennas and Propagation Magazine, vol. 59, no. 2, pp. 162–170, Apr. 2017.
10. X. Chen, S. Zhang, Q. Li, and G. F. Pedersen, "A review of mutual coupling in MIMO systems," IEEE Access, vol. 6, pp. 24706–24719, 2018

CATRON: A CABLE TUNNEL INSPECTION ROBOT

Arya Sasikumar
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
arya.sasikumar@ahalia.ac.in

Dhanush P
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
dhanush.p@ahalia.ac.in

Gowthami S
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
gowthami.s@ahalia.ac.in

Mahitha M
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
mahitha.m@ahalia.ac.in

Ms. Swetha C
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
swetha.c@ahalia.ac.in

Dr. V. Balamurugan
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
hodece.aset@ahalia.ac.in

Abstract— CATRON, the cable tunnel inspection robot aims to scrutinize and continuously monitor the conditions inside tunnels by using advanced sensors and specialized integrated HD cameras. These sensors help in identifying potential hazards, such as the presence of gas leakages, abnormal temperature levels, fluctuations in humidity, or any physical obstacles that may disrupt normal operations. It is also capable of disseminating live data and real-time video feeds to remote monitoring stations through cloud integration. By automating the inspection process, the robot reduces human risk by eliminating the need for manual entry into potentially dangerous tunnels. Furthermore, the system's preventive maintenance allows for the early detection of faults or deformity, and ensuring uninterrupted operation of cable systems.

Key Words: Tunnel inspection, hazard detection, gas leaks, temperature and humidity monitoring, obstacle detection, integrated sensors, live video streaming, remote monitoring, cloud integration, Raspberry Pi, motor control, preventive maintenance, real-time data, safety, cost-effective solution.

I. INTRODUCTION

Cable tunnels are the backbone of modern infrastructure, carrying essential power supply and communication lines. However, these tunnels are often narrow, poorly ventilated, and prone to hazards such as toxic gas leaks, extreme temperatures, humidity variations, and unexpected physical disturbances like rock collisions. Conducting manual inspections in such harsh environments is risky, time-consuming, and often inadequate. To address these challenges, a smart Cable Tunnel Inspection Robot has been designed. The system integrates multiple sensors, a camera, and a Raspberry Pi as the processor to continuously monitor tunnel conditions. It can accurately detect harmful gases, temperature, humidity, and physical obstacles while safely navigating through the tunnel. The robot also provides real-time data and live video streams to remote monitoring stations or web platforms via cloud integration, enabling workers to observe tunnel conditions without direct exposure to danger. This automation not only enhances worker's safety by eliminating the need for manual entry but also supports preventive maintenance throughout the early fault detection. This makes the system a faster, safer, explicit and highly scalable solution for the cable tunnel inspections.

II. LITERATURE REVIEW

Inspection and monitoring of infrastructure systems such as bridges, tunnels, and power transmission lines are critical for ensuring public safety and operational reliability. Conventional inspection techniques, which often rely on manual assessments, pose several limitations including safety risks, restricted accessibility, and inefficiency in hazardous or remote environments. To overcome these challenges, researchers are increasingly focusing on robotic systems enhanced with intelligent sensing, resilient power supplies, and autonomous navigation capabilities.

Structural testing provides fundamental insights into the long-term safety of large-scale constructions. For instance, experimental investigations on cable-supported bridges have demonstrated the influence of loading conditions and vibration responses on stability, contributing to safer design practices [1]. In underground facilities, structural protection against natural hazards is equally important. Tao et al. validated the effectiveness of NPR anchor cable support systems through shaking table experiments, proving their role in strengthening tunnels subjected to seismic activities [2]. These studies highlight the need for experimental validation as a basis for reliable design in both surface-level and underground structures. With the advancement of sensors and artificial intelligence, autonomous inspection has gained substantial accuracy and efficiency.

Ciuffreda et al. combined passive infrared (PIR) and ultrasonic sensors in a multi-sensor framework to achieve robust human localization in indoor conditions [3]. Tang advanced the field further with FVCNet, a visual clustering network tailored for obstacle recognition in power line inspection tasks, demonstrating the benefits of deep learning in improving environmental awareness [4]. Alongside these contributions, Ekren et al. provided a comprehensive survey of line-suspended inspection robots, discussing their operational strengths and the constraints related to adaptability and load capacity [5]. Best et al. also addressed large-scale exploration challenges, presenting a multi-robot, multi-sensor approach where autonomous aerial systems worked collaboratively to extend inspection coverage [6].

Singh and Pete [7] proposed a smart intensity control system for surgical lights using ultrasonic sensors to reduce shadow effects caused by obstacles during medical procedures. The system detects the distance of objects, such as the surgeon's head, and automatically adjusts the LED intensity using PWM control. This

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

dynamic adjustment provides uniform illumination without requiring manual repositioning, improving lighting efficiency and accuracy. The study demonstrates the effectiveness of ultrasonic sensing in real-time feedback control applications. Ciuffreda et al. [8] proposed a multi-sensor fusion approach combining PIR and ultrasonic sensors for indoor human localization on a mobile robot. The system leverages PIR for motion detection and ultrasonic sensors for precise distance measurement, improving accuracy and reducing false detections caused by environmental noise. The study demonstrates that sensor fusion enhances real-time localization, making it effective for service robots and indoor monitoring applications. [9] Chen introduced a UAV patrol path planning method based on machine vision and multi-sensor fusion. By combining data from visual, infrared, and ultrasonic sensors, the system achieves accurate obstacle detection and adaptive path correction. The fusion of multiple sensors enhances the UAV's ability to operate efficiently in dynamic and uncertain environments. Cai et al. [10] provided a comprehensive review of intelligent systems utilizing multi-sensor fusion and advanced information processing techniques for navigation and path planning. The authors discuss current developments in sensor integration, perception, and decision-making frameworks that contribute to intelligent motion control. This review underscores the importance of integrating heterogeneous sensor data to achieve robust and reliable autonomous navigation in complex scenarios.

III. REQUIREMENTS AND PROPOSED ALGORITHM

1. HARDWARE REQUIREMENTS

a) Raspberry Pi 3 Model B+: Single-board computer with 1.4GHz 64-bit quad-core processor, 1GB RAM, WiFi, Bluetooth, GPIO headers, camera interface, HDMI, USB, Ethernet; powered by 5V/2.5A supply.

b) MPU-6050: 6-axis motion tracking (3-axis gyroscope + 3-axis accelerometer + temperature sensor); communicates via I2C; needs 3–5V supply.

c) BMM-150: 3-axis digital geomagnetic (magnetometer) sensor; I2C/SPI interface; powered with 1.6–3.6V.

d) MQ-5: Semiconductor gas sensor for LPG, methane, propane; requires 5V; analog/digital output.

e) Micro LiDAR (TFMini type): Distance sensor (0.1–12m), UART/I2C interface, 5V supply.

f) DHT11: Digital temperature and humidity sensor; 3.3–5V supply; single-wire digital output.

g) Camera Module: Compatible Raspberry Pi CSI camera; typically, 8–12MP, 15-pin CSI cable connection.

h) L298N Motor Driver: Dual H-bridge driver for up to 2 DC/stepper motors; drives motors with voltage 5–35V, current up to 2A per channel, logic powered at 5V.

i) Motors: Two DC motors for robotic actuation

j) Power Supply: 5V DC source for the Raspberry Pi and select modules.

2. SOFTWARE REQUIREMENTS

a) Raspbian OS (or Raspberry Pi OS): Linux-based operating system for Raspberry Pi.

b) Python/C++ libraries:

i. Sensor interface libraries (e.g., smbus for I2C, RPi.GPIO)

ii. Picamera for interfacing with the Raspberry Pi camera.

iii. Libraries for connecting to the cloud (e.g., requests, Firebase SDK for Python or Node.js)

c) Cloud API: Integration with Firebase for database/webpage storage, RESTful API access, and authentication.

d) Webpage/Firebase Frontend: A simple web dashboard to monitor and control, typically using HTML/JS and Firebase web SDK.

e) Motor Control Software: Python scripts for GPIO-based motor control (via L298N).

f) Data Logging/Analysis: Optional Python scripts for collecting, analyzing, and visualizing sensor data.

g) Remote Update/SSH Tools: For remote programming and updating the Raspberry Pi OS and scripts.

3. BLOCK DIAGRAM

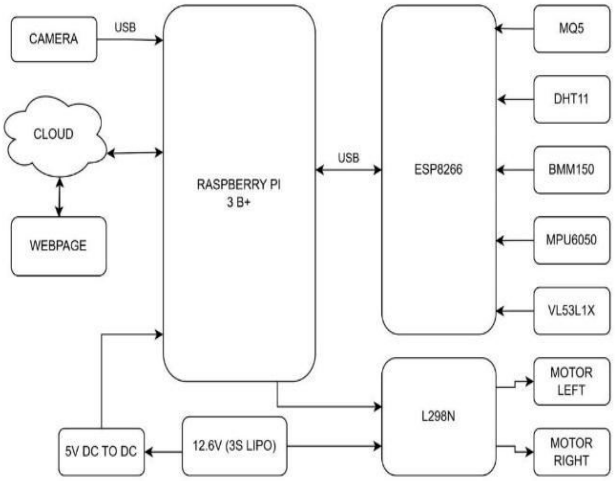


Fig. 1. Block Diagram

1. Power Supply: A 12.6 V (3S Li-Po) battery powers the system. A DC-to-DC converter steps the voltage down to 5 V to safely operate the Raspberry Pi 3 Model B+ and sensors, while the higher voltage is supplied to the motors.

2. Sensors Collect Data: Environmental and motion data are gathered using sensors such as MPU-6050 for motion, BMM150 for navigation, MQ-5 Gas Sensor for gas detection, VL53L1X Time-of-Flight Distance Sensor for obstacle detection, and DHT11 for temperature and humidity measurement.

3. Sensor Interface using ESP8266: The sensors are connected to the ESP8266 microcontroller, which gathers sensor readings and sends the collected data to the Raspberry Pi through a USB connection.

4. Processing in Raspberry Pi: The Raspberry Pi 3 Model B+ receives sensor information, processes the data, and makes decisions for the robot's operation based on the detected environmental conditions.

5. Camera Input: A camera connected to the Raspberry Pi captures live video of the surroundings. This allows the system to visually monitor the environment during operation.

6. Motor Control: Control signals from the Raspberry Pi are sent to the L298N Dual H-Bridge Motor Driver, which drives the left and right DC motors and enables the robot to move forward, backward, or turn.

7. Cloud Communication: Sensor readings and video data are transmitted from the Raspberry Pi to cloud services such as Firebase, allowing the information to be stored and accessed remotely.

8. Webpage Monitoring: A web interface allows the operator to view live sensor readings and video feed. This helps in remotely observing tunnel conditions and receiving alerts if any abnormal situation is detected.

4. FLOWCHART

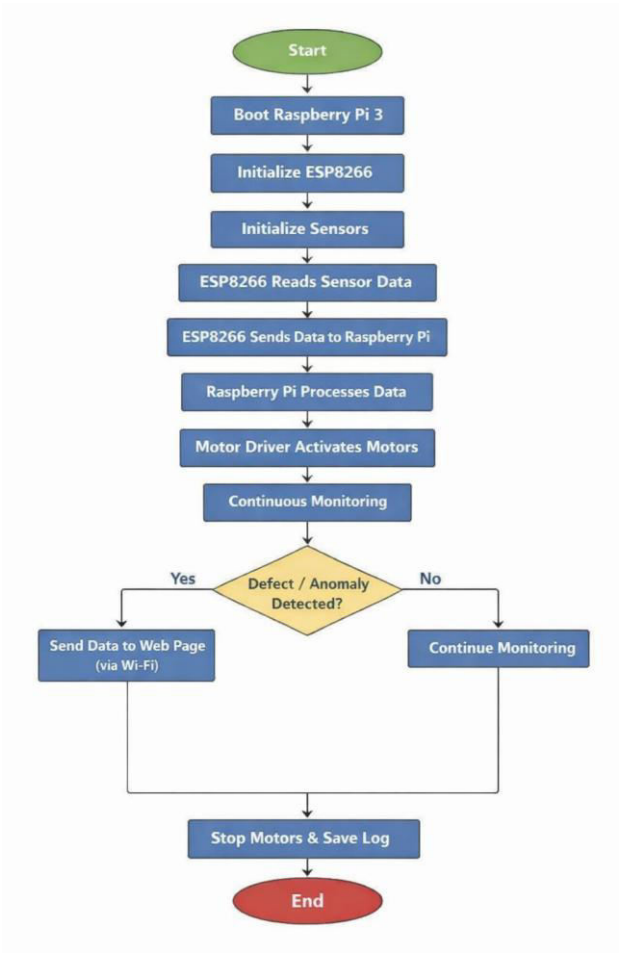


Fig. 2. Flowchart

5. ALGORITHM

i. Initialization: When the robot is powered on, the Raspberry Pi 3 boots and prepares the system for operation. At the same time, the ESP8266 module is initialized to establish wireless communication between the sensing unit and the processing unit.

ii. Sensor Initialization: After the controller starts, all connected sensors are initialized so they can begin monitoring the tunnel environment.

iii. Data Collection: The ESP8266 reads sensor data continuously while the robot operates. These readings provide information about the conditions inside the tunnel.

iv. Data Transmission and Processing: The collected data is transmitted to the Raspberry Pi, where it is processed and analyzed. Based on this analysis, the system determines the appropriate action.

v. Motor Control: Control signals are sent to the motor driver, which activates and regulates the motors that move the robot through the tunnel.

vi. Continuous Monitoring: During operation, the system repeatedly collects and analyzes sensor data to keep track of tunnel conditions.

vii. Defect Detection: If a defect or anomaly is detected, the system sends the information to a monitoring web page through Wi-Fi.

viii. Termination: If the inspection ends or the system stops, the motors are turned off and the operation log is saved.

6. CIRCUIT DIAGRAM

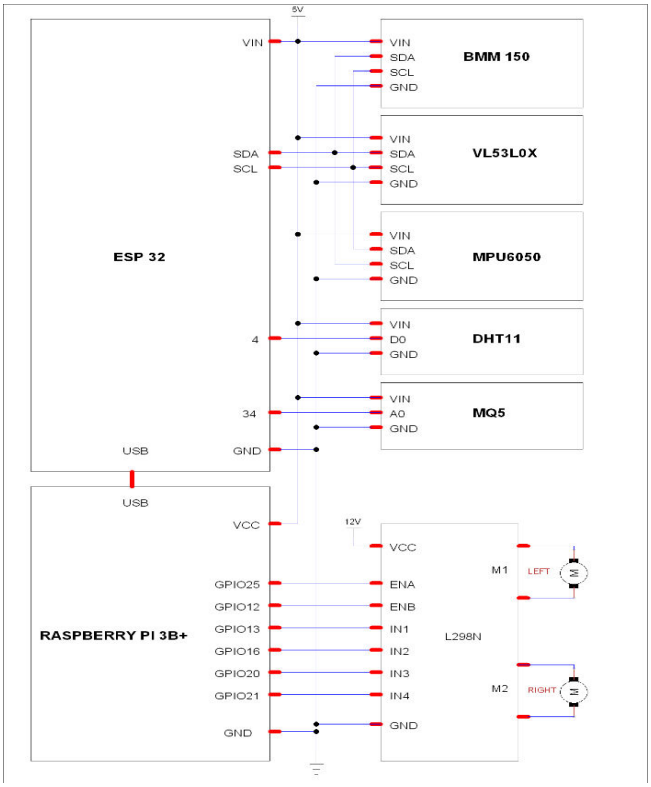


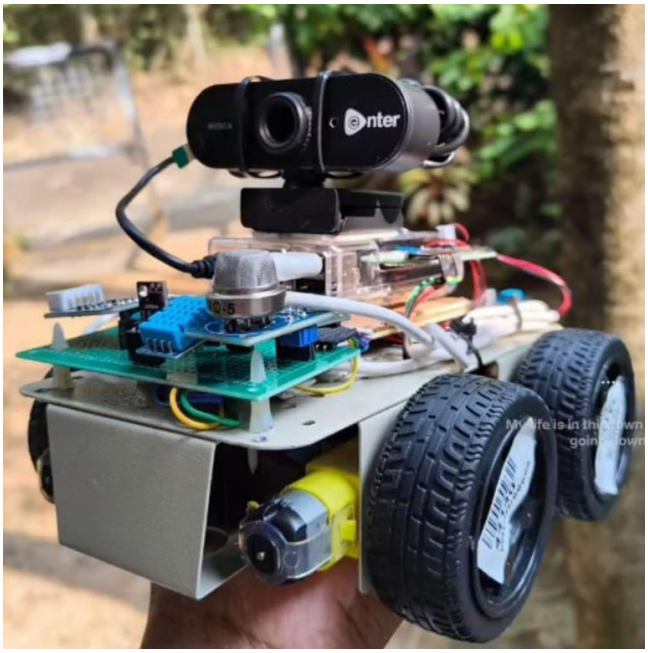
Fig. 3. Circuit diagram

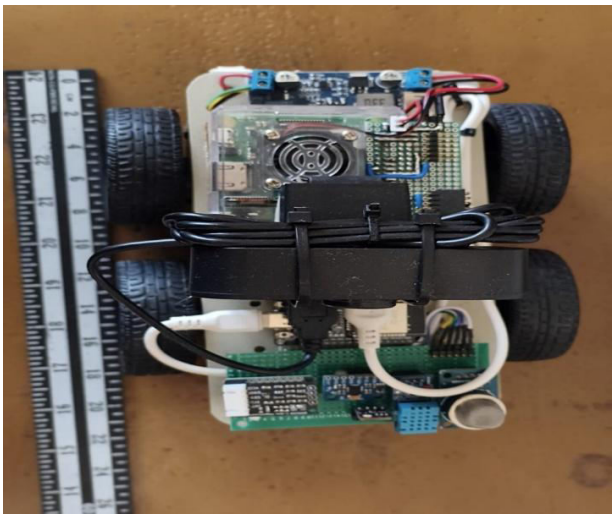
IV. OBSERVATIONS

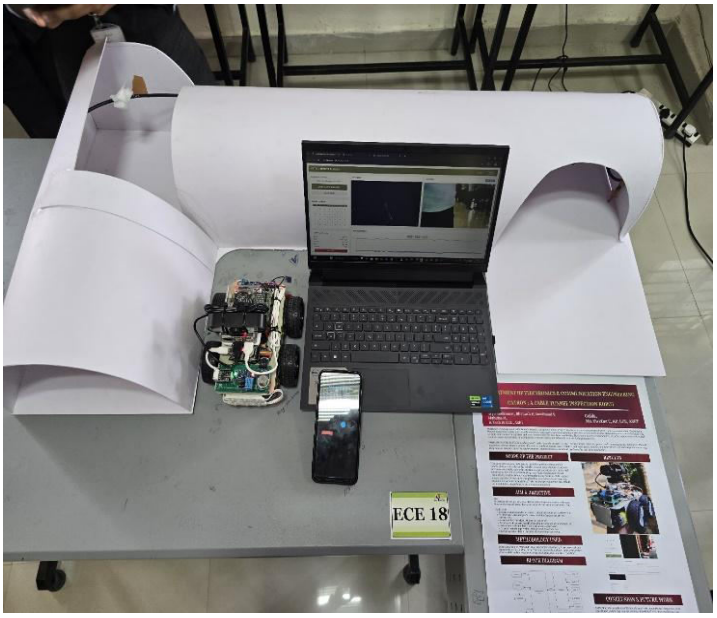
During the implementation of the CATRON dashboard, it was observed that the system effectively integrates multiple modules for tunnel inspection and monitoring. The environmental monitoring unit consistently tracked temperature, humidity, and gas levels, and it promptly generated alerts whenever these parameters crossed safety thresholds. The navigation component provided continuous updates on orientation, heading, and LiDAR distance, ensuring accurate identification of nearby obstacles. The dashboard also incorporated real-time data visualization in the form of line charts and bar charts, which presented recent sensor readings and defect severity across tunnel sections in a clear and structured manner. Additionally, the interactive 3D tunnel model enabled detailed visualization of robot movement and LiDAR scanning, thereby enhancing situational awareness for inspection activities. Notifications were displayed through on-screen messages, sound alerts, and browser notifications, reducing the chances of missing.

V. RESULTS

The developed Cable Tunnel Inspection Robot was successfully tested in a simulated tunnel setup. The robot, built using Raspberry Pi 3 and ESP8266, was able to move through the tunnel while continuously collecting sensor data. During testing, the sensors monitored environmental conditions and the camera provided live visual feedback of the tunnel interior. The collected data was transmitted wirelessly and displayed on the monitoring interface.
When abnormal conditions were introduced, the system successfully detected them and sent alerts. The results indicate that the robot can effectively assist in safe and efficient cable tunnel inspection.







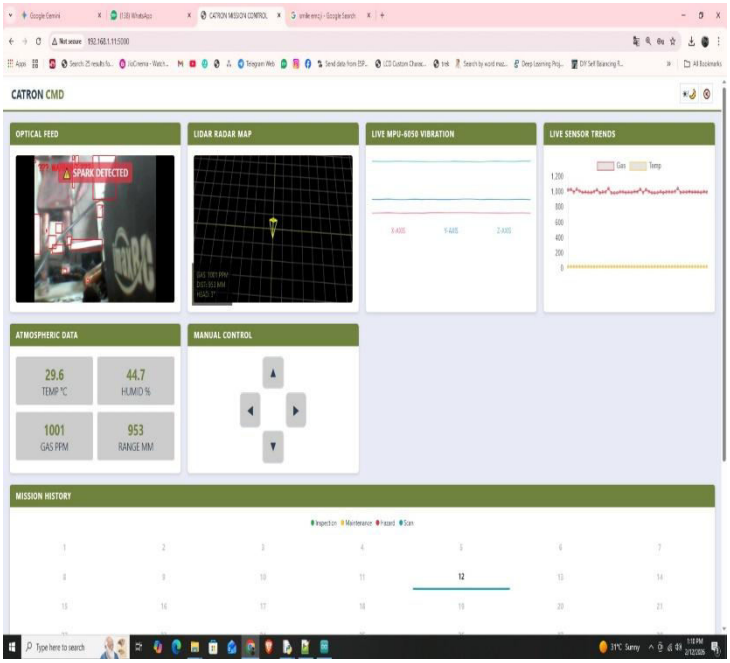


Fig. 4. Cloud-based Webpage interface for viewing collected results

The graphical analysis clearly highlights the variations in tunnel conditions over time, showcasing how parameters such as temperature, humidity, and gas levels fluctuate in response to environmental changes. The spikes in data trends indicate the system's high sensitivity in capturing real-time anomalies, allowing operators to quickly recognize and address potential hazards. This ensures that even small deviations from safe operating conditions can be identified before they escalate into critical issues, thereby improving the reliability of tunnel management. Additionally, the integration of visual heat maps and AI-based classification further complements the graph by providing an intuitive understanding of fault distribution and obstacle severity. This combination of numerical data and visual representation not only aids in quick decision-making but also strengthens long-term planning for maintenance and safety improvements. The graphs and dashboards together validate the efficiency of the CATRON system in offering a practical, real-time monitoring solution for safe and efficient tunnel inspection.

VI. CONCLUSION

The performance of the Cable Tunnel Inspection Robot demonstrates a pragmatic and steady solution to the challenges of substantiates underground cable tunnels. By combining sensors for gas, temperature, humidity, motion and obstacle detection with a compact yet effective robotic design, the system minimizes the need for human involvement in perilous environments while ensuring that vital data is collected accurately. The integration of the live video feed and real-time monitoring strengthens the overall efficiency of tunnel inspection, giving operators a clear view of conditions without requiring physical entry.

This approach not only improves safety but also activates quicker recognition of hazards that could otherwise lead to accidents, damage, or costly interruptions in power systems. The robot is lightweight, portable, and adaptable, making it suitable for use in a variety of infrastructure settings. Through preventive monitoring and early fault detection, the project highlights how automated inspection can reduce risk, cut maintenance costs, and contribute to more reliable tunnel management in the long run.

VII. REFERENCES

1. Y. Huang, S. Yang, and Z. Zhen, "Research on Multi-Sensor Information Integration Inspection Data Collection of Four Legged Robot for Cable Tunnel," in 2025 5th International Conference on Sensors and Information Technology (ICSI), 2025. doi: 10.1109/ICSI64877.2025.11010074
2. B.-S. Ko et al., "Research on Safe and Efficient Power Design of Unmanned Tunnel Inspection Robot Using Lithium Iron Phosphate(LFP) Battery," in 2024 10th International Conference on Condition Monitoring and Diagnosis (CMD).
3. Y. Cai, T. Qin, Y. Ou, and R. Wei, "Intelligent Systems in Motion: A Comprehensive Review on Multi-Sensor Fusion and Information Processing from Sensing to Navigation in Path Planning," International Journal on Semantic Web and Information Systems (IJSWIS), vol. 19, no. 1, pp. 1-35, 2024.
4. N. Ekren, Z. Karagöz, and M. Şahin, "A Review of Line Suspended Inspection Robots for Power Transmission Lines," Journal of Electrical Engineering & Technology, vol. 19, pp. 2549–2583, 2024.
5. H. Fu, L. Qiu, Y. Ai, C. Xie, and Y. Yan, "Analysis of Early Warning Indicators for High Voltage Power Cable Tunnel Fire Based on Experimental Data," in Proc. 2023 3rd International Conference on Electrical Engineering and Control Science (IC2ECS), pp. 1631–1634, IEEE, 2023.
6. J. Sawada, K. Kusumoto, T. Munakata, Y. Maikawa, and Y. Ishikawa, "A Mobile Robot for Inspection of Power Transmission Lines," IEEE Transactions on Power Delivery, vol. 6, no. 1, pp. 309–315, Jan. 1991.
7. Z. Jia, H. Liu, H. Zheng, S. Fan, and Z. Liu, "An Intelligent Inspection Robot for Underground Cable Trenches Based on Adaptive 2D-SLAM," Machines, vol. 10, no. 11, p. 1011, Nov. 2022.
8. S. Fang, Z. Wang, A. Song, Z. Cui, and M. Yang, "Research on Communication and Control Method of Towing Cable Inspection Robot in Urban Underground Pipe Network," in Proc. 11th Int. Conf. on Computer Engineering and Networks, Lecture Notes in Electrical Engineering, vol. 808, pp. 1031–1035, Springer.

VLSI ARCHITECTURE FOR VIDEO EDGE DETECTION

Aditya Kiran S R
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
adityakiran.sr@ahalia.ac.in

Anirudh Sreejithesh
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
anirudh.sreejithesh@ahalia.ac.in

Ashik M
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
ashik.m@ahalia.ac.in

Jishnu R
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
jishnu.r@ahalia.ac.in

Gayathri P S
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
gayathri.ps@ahalia.ac.in

Dr .V .Balamurugan
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
hodece.aset@ahalia.ac

***Abstract*—**Applications like object recognition, scene comprehension, and video surveillance depend heavily on edge detection in the field of video processing. Effective hardware architectures that can manage the computational demands of edge detection algorithms while guaranteeing low power consumption are necessary for real-time video data processing. A novel VLSI (Very-Large-Scale Integration) architecture for real-time video edge detection is presented in this paper. The suggested architecture seeks to achieve a balance between resource-efficient implementation and high-performance edge detection. To increase computational speed, it makes use of parallel processing and optimized data flow. To guarantee seamless video stream processing, extra care is taken with memory management and data buffering. The architecture is implemented on Artix 7 and Spartan 3E FPG and is based on Canny edge detection algorithms. The proposed architecture is also implemented on GPP to validate the performance of the VLSI architecture.

Key words: Canny edge detector ,bilateral filtering, FPGA-based designs, image denoising.

Introduction

One of those silent but crucial image processing steps is edge detection. It just searches for abrupt shifts in intensity, but those shifts define the boundaries and structure of objects. Higher-level tasks like segmentation and recognition become unstable without it.

The Canny edge detector stands out among the various methods because it attempts to strike a balance between precise localization, accuracy, and noise suppression. However, Gaussian smoothing, its initial step, seems like a compromise. Yes, it lessens noise, but it also softens edges—sometimes to the point where finer details are lost.

When a bilateral filter is used in its place, the situation becomes more cautious. It preserves edges while lowering noise by taking into account both distance and intensity differences rather than averaging everything in the vicinity. Cleaner, more continuous edge maps are the end result, particularly in noisy or low-quality images.

There are trade-offs involved. The computation is a little more involved, and parameter tuning can be challenging. Nevertheless, the concept is fairly simple: maintain Canny's strength while reducing the severity of its initial step. In the end, that tiny adjustment has a significant impact.

I. LITERATURE SURVEY

This work's larger context can be comprehended in light of the continuous transition to more intelligent and adaptive systems. Automation is just one aspect of Industry 5.0; another is how intelligent technologies and people work together to increase productivity and adaptability. According to Maddikunta et al. (2022), this shift is toward human-centric systems in which technologies like edge computing, AI, and collaborative robots assist rather than completely replace decision-making. This concept is related to image processing, where algorithms are expected to generate accurate and understandable results, even though it is frequently framed in the context of manufacturing.

There isn't a single technique that is ideal in every circumstance when it comes to image processing and restoration. Gota and Min (2013) demonstrate how various algorithms behave differently based on the kind of image degradation. For example, Wiener filtering is better suited for Gaussian noise, whereas median filtering performs well for salt-and-pepper noise. Because the effectiveness depends on the particular conditions of the input, choosing a technique is therefore more difficult than it may first appear.

The final quality of edge detection is greatly influenced by noise handling. According to Chen et al. (2015), processing can be enhanced by estimating noise using local image statistics, such as gradients and pixel variations, without requiring prior assumptions about the noise distribution. In the absence of such estimation, preprocessing may fail to suppress undesired variations or eliminate important details, both of which have an impact on the final result.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

By adjusting to the image's structure, more sophisticated methods try to overcome these constraints. A technique based on structure tensor adaptive total variation is presented by Prasath and Thanh (2019) with the goal of maintaining edges while lowering noise by taking local image orientation into account. In a similar vein, Thanh et al. (2020) suggest an adaptive restoration method that preserves fine details by utilizing inverse gradients and higher-order information. Filtering techniques have also changed in this way. Chen et al. (2020) introduced a Gaussian-adaptive bilateral filter that better balances edge preservation and noise reduction by modifying its parameters according to image content.

All of these studies lead to the same conclusion. The goal of image processing is to balance conflicting elements like computational efficiency, edge preservation, and noise reduction rather than to find a single best method. Any strategy's efficacy is determined by how well it manages these trade-offs rather than how well it functions in optimal circumstances.

II. PROPOSED METHOD

The proposed edge detection method was not built in a single leap from theory to hardware. It evolved in stages, and honestly, that staged approach made all the difference. MATLAB was the starting point. Not because it is glamorous, but because it allows you to test ideas quickly, break them, fix them, and visualize exactly what went wrong. That flexibility is hard to replace.

The proposed edge detection method was not built in a single leap from theory to hardware. It evolved in stages, and honestly, that staged approach made all the difference. MATLAB was the starting point. Not because it is glamorous, but because it allows you to test ideas quickly, break them, fix them, and visualize exactly what went wrong. That flexibility is hard to replace.

The process started with grayscale images that had been intentionally messed up with different kinds of noise and levels of blur. Not pictures that look like they were taken with a clean textbook camera, but pictures that look like they were taken with a real camera in a factory or outside. Textures that are grainy, light that isn't even, and motion blur. These became the place to test things. A bilateral filter was used to do the preprocessing before edge detection. At first, MATLAB's built-in tools helped, but custom routines gave me more control over the spatial and intensity sigma. It felt more like adjusting a lens than pressing buttons to change these settings. If you raise spatial influence too much, the edges get softer. If you make intensity sensitivity too tight, noise will get through. A balance had to be struck.

The Canny framework only worked after smoothing. The edge() function in MATLAB with the Canny option was a good starting point, but if we had only used it, we would have missed too much of how things work inside. So, the algorithm was put back together one piece at a time. Sobel filters figured out the size and direction of the gradient. Non-maximum suppression made edges thinner by keeping only local maxima along the gradient directions. Next came hysteresis thresholding, which used two thresholds to tell the difference between confident and uncertain edges and then reconnect weak edges that were still useful. Matrix scans and logical operations took care of connectivity. It was nice to see that the custom implementation's results were similar to those built into MATLAB. The outputs were very similar, which made it seem like the reconstruction was accurate.

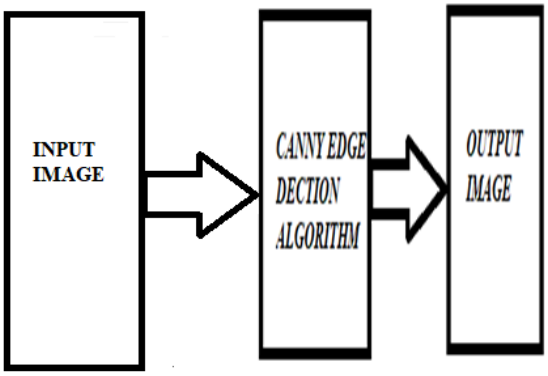


FIGURE: Top-level architecture of the proposed design

But what really stood out was how bilateral filtering worked in the Canny pipeline. Bilateral preprocessing made edge maps with stronger continuity and fewer false responses in noisy areas than the traditional Gaussian smoothing stage. The change wasn't just in numbers. You could see it. Edges along object boundaries looked less broken up, especially when the inputs were blurry. In applications like robotic inspection, where a broken contour can cause misinterpretation, visual clarity is important.

Once the MATLAB model was solid and dependable, Xilinx Vivado was used to extend the implementation to hardware. However, this change brought with it a new set of difficulties. When transferred onto hardware, algorithms that perform well in a floating-point MATLAB environment frequently become constrained. Because of this, the design has to be redone in Verilog and divided into modular parts, such as thresholding, bilateral filtering, gradient computation, and magnitude estimation.

Early on, hardware limitations started to influence the design. For example, it would be quite expensive to implement the bilateral filter in its original form. In order to simplify it, integer approximations and a separable structure were used. With the use of shift registers and line buffers, a sliding 3x3 window architecture allowed for continuous pixel streaming without using too much memory.In a similar vein, gradient magnitude was estimated using the sum of absolute values r ather than square root operations, and the Sobel operator wa s implemented using fixedpoint arithmetic.Although this is only an approximation, it worked well in practice and greatl y enhanced resource efficiency and timing performance.

Complexity increased with more advanced stages. In order to prevent disrupting data flow, non-maximum suppression necessitated directional comparisons based on gradient estimates, which had to be carefully handled inside a

streaming architecture. Finding weak edges that are coupled to strong ones was a particularly difficult portion of hysteresis thresholding. In order to prevent weak edges from adding to noise, buffering methods and dual-threshold comparison logic were required.

An initial degree of validation was offered by simulation in Vivado's testbench environment. After loading test images created in MATLAB into memory, the edge outputs were compared to the reference results produced by MATLAB. Although there were some minor variations, mostly because of quantization and fixed-point approximations, the identified edges' general structure was unchanged. Resource consumption data after synthesis showed effective use of flip-flops, block RAM, and LUTs. The design's ability to process standard-resolution photos in real time was further validated by timing analysis.

System integration was the last step, which entailed moving image frames from MATLAB to the FPGA, processing them in hardware, and then returning them for viewing. Practical evidence that the design maintained its intended functionality in spite of hardware constraints was provided by the close resemblance between hardware and software outputs.
Overall, a balanced approach to edge detection under FPGA restrictions is demonstrated by the Canny edge detector's inclusion of bilateral filtering. The architecture strikes a balance between hardware resource use, computational efficiency, and precision. Although there is certainly space for development, especially in adaptive thresholding and edge tracking, the current implementation offers a workable and scalable solution for real-time applications like robotics and embedded vision.

III. IMPLEMENTATION

Development Environment

The Xilinx Vivado design suite was used to create the edge detection system, which was written in Verilog HDL. Starting with the algorithm itself and progressing via HDL coding, simulation, and synthesis, the workflow followed the well-known cadence of digital design. When written like way, it sounds linear. It was more iterative in practice. Revisiting simulation, modifying timing, and double-checking logic were frequently required for a little algorithmic change.
Every image had to be transformed into a pixel stream format compatible with the FPGA architecture because hardware does not perceive images in the same way that MATLAB does. To put it another way, the system processed an image as a continuous flow of pixel values synced with clock cycles rather than as a two-dimensional matrix. The entire design methodology was influenced by this change in viewpoint from matrix operations to streaming data.

Algorithm Realization

The edge detection algorithm is fundamentally a series of structured processing steps. Before sending the incoming pixel data to the following stage in the pipeline, each stage applies a precisely defined transformation. Conceptually speaking, there are four primary operations involved in the implementation:

* Bilateral filtering is used in image preprocessing to reduce noise while maintaining edges.
* Calculating gradients in both horizontal and vertical directions using Sobel operators
* The computed gradients were used to calculate the edge magnitude.
* Thresholding to determine and complete edge pixels

The pipeline construction makes sure that the next pixel enters a stage as soon as the previous one leaves. Continuous processing, which is necessary for real-time applications, is supported by this streaming architecture. The system works with pixels one at a time instead of keeping full frames for calculation. Although it necessitates precise synchronization between steps, this design decision lowers memory overhead.

Module Design

The architecture was split up into several modules to make development easier. Within the edge detection flow, each module carries out a specific, clearly defined task. The main elements consist of:

* Module for processing input pixels
* Module for bilateral filtering
* Module for calculating gradients
* Unit for calculating edge magnitude
* Module for thresholding
* Module for making output images

This modular design worked well. Isolating a failing gradient block while debugging is much simpler than deciphering a monolithic design. Additionally, it allows for scalability. Only one module needs to be changed if a later edition substitutes a different gradient method for the Sobel operator. The pipeline is still intact in its entirety.

Simulation and Verification

Vivado's simulation tools were used for verification. At every processing level, controlled input data and output monitoring were provided using custom testbenches. In order to verify accuracy, intermediate signals like gradient values and threshold flags were examined in addition to the final edge map.
The simulation results showed that the architecture kept its logical behavior stable while reliably finding edges in a wide range of visual inputs. Some differences were caused by the quantization effects of fixed-point hardware arithmetic, but these differences didn't have much of an effect on visual quality. It was good to know that the pipeline stayed in sync during different test scenarios, given the streaming architecture.

Performance Evaluation

The two main areas of evaluation were architectural efficiency and detection accuracy. In situations when

preprocessing was successfully implemented, the final images showed distinct edge localization with fewer noise artifacts.

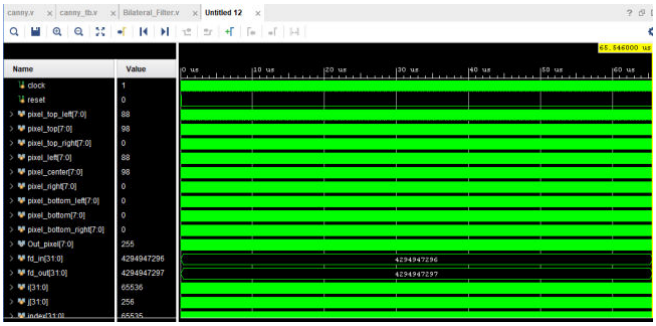
From a hardware standpoint, the pipelined design made it possible for data to flow continuously and with minimal delay. The structure satisfies real-time processing requirements for standard-resolution photos, and resource use stayed within acceptable bounds.

All things considered, the implementation achieves a compromise between hardware practicality and algorithmic clarity. The technology does more than just identify edges. It does this in a way that respects scalability, modularity, and temporal constraints—all of which are equally important in embedded vision systems as the final visual output.

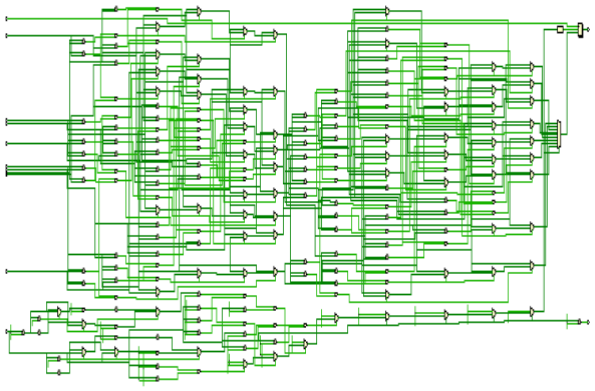
IV. RESULTS & ANALYSIS

The output demonstrates that the suggested solution effectively finds edges in an input image by detecting changes in intensity levels. In order to ensure a steady gradient calculation in the subsequent step, the pre-processing stage reduces noise and gathers important edge information. The Sobel-based method effectively highlights regions with significant intensity change, enabling accurate edge detection.

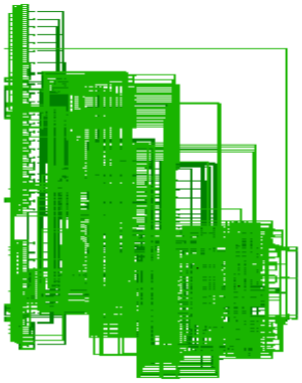
Simulation waveforms:



Rtl schematic:



Tech schematic:



Area:

Name	Slice LUTs (134600)	Slice Registers (269200)	DSPs (740)	Bonded IOB (400)	BUFGCTRL (32)
N canny	38	2	8	74	1

Delay:

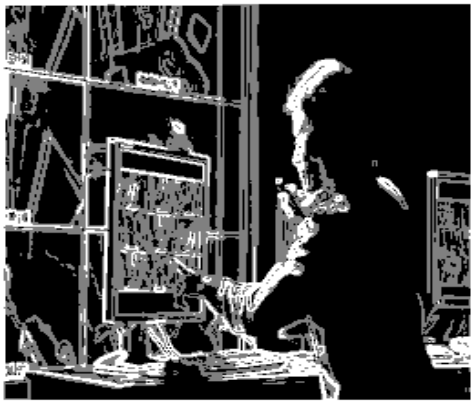
Name	Slack	Levels	Routes	High Fanout	From	To	Total Delay
Path 1	∞	14	12	3	pixel_top_left[7]	Out_pixel_reg[6]/D	16.058
Path 2	∞	14	12	3	pixel_top_left[7]	Out_pixel_reg[7]/D	16.058
Path 3	∞	2	2	7	Out_pixel_reg[6]/C	Out_pixel[0]	3.521
Path 4	∞	2	2	7	Out_pixel_reg[6]/C	Out_pixel[1]	3.521
Path 5	∞	2	2	7	Out_pixel_reg[6]/C	Out_pixel[2]	3.521
Path 6	∞	2	2	7	Out_pixel_reg[6]/C	Out_pixel[3]	3.521
Path 7	∞	2	2	7	Out_pixel_reg[6]/C	Out_pixel[4]	3.521

Power:

Total On-Chip Power:	**9.474 W**
Design Power Budget:	**Not Specified**
Power Budget Margin:	**N/A**
Junction Temperature:	**42.7°C**
Thermal Margin:	42.3°C (22.3 W)
Effective ϑJA:	1.9°C/W

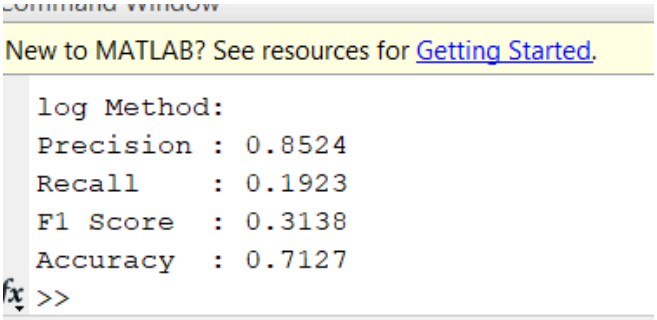
INPUT IMAGE:

OUTPUT IMAGE:

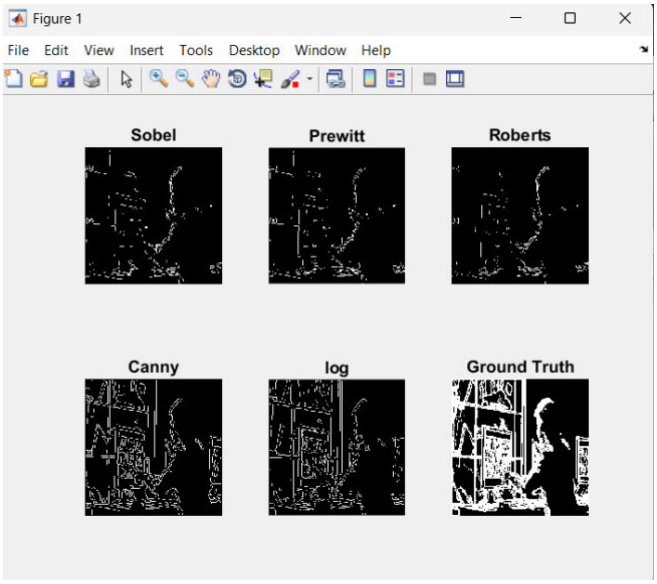


FURTHER VERIFICATION

Parameters:



PIPELINE OUTPUT:



The suggested implementation is compared to the Canny edge detector, which is employed here as a reference rather than an absolute ground truth, in order to execute the performance evaluation in MATLAB. Even though it's simple to overlook, that distinction is important. In essence, the degree to which the suggested approach conforms to Canny's concept of edges is being measured.

At first sight, the precision of 0.8524 in the results seems comforting. It implies that the system is typically accurate when it detects an edge because Canny also marks it as such. The detection is somewhat disciplined. It is not reacting irresponsibly to noise or sporadic changes in intensity. But there's a price for such confidence.

It is hard to ignore the recall value of 0.1923. The suggested approach only fails to catch a significant percentage of the edges identified by Canny. This disparity is not small. It suggests a very selective—possibly too so—system. It tends to overlook weaker or more subtle transitions in favor of strong, well-defined gradients. That may even be preferred in certain situations. Others run the danger of losing crucial structural details.

This imbalance is reflected in the F1 score, which is approximately 0.3138. It prevents the interpretation from being dominated by the great precision. Rather, it makes the trade-off quite clear. Although the accuracy number of 0.7127 seems reasonable, care should be taken when interpreting it. The majority of pixels in edge detection are part of the backdrop. Accuracy can be increased without necessarily improving edge quality by correctly detecting these non-edge pixels.

The trend is consistent with these figures when one looks back at the visual outputs. Although they are partial, the identified edges are typically clear and distinct. Particularly in situations when intensity shifts are modest, some boundaries seem to be breached or completely absent. The implementation feels constrained in comparison to the reference, almost uncertain about what it considers to be an advantage.

Thus, the result is not only "excellent" or "bad." It is particular. The design sacrifices completeness in order to create edges that are stable and noise-resistant. What the system is eventually supposed to prioritize will determine whether or not that trade-off is acceptable, rather than the measures themselves.

V. CONCLUSION

he issues of power consumption and computing efficiency are well addressed by the suggested VLSI architecture for real-time video edge detection. The architecture maintains resource efficiency while achieving high-performance edge detection by utilizing parallel processing, streamlined data flow, and effective memory management. The architecture's feasibility for real-time applications is shown by its implementation on Artix 7 and Spartan 3E FPGAs, and its efficacy is further validated on GPP. This method opens the door to more sophisticated, low-power hardware solutions for video processing, which makes it appropriate for a variety of

uses like video surveillance, object recognition, and scene comprehension.

REFERENCES

[1] Gota, A., & Min, Z. J. (2013). Analysis and Comparison on Image Restoration Algorithms Using MATLAB. International Journal of Engineering Research & Technology (IJERT) Vol, 2, 1350-1360.

[2] Mahalakshmi, A.,& Shanthini, B. (2016, January). A survey on image deblurring. In 2016 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-5). IEEE.

[3] Flusser, J., Farokhi, S., Hoschl, C., Suk, T., Zitov ¨ a, B., ´ & Pedone, M. (2015). Recognition of images degraded by Gaussian blur. IEEE transactions on Image Processing, 25(2), 790-806.

[4] Ramya, S.,& Christial, T. M. (2011, March). Restoration of blurred images using Blind Deconvolution Algorithm. In 2011 International Conference on Emerging Trends in Electrical and Computer Technology (pp. 496-499). IEEE.

[5] Sada, M. M., & Mahesh, M. G. (2018). Image deblurring techniques—a detail review. Int. J. Sci. Res. Sci. Eng. Technol, 4(2), 15.

[6] Ansari, M. A., Kurchaniya, D., & Dixit, M. (2017). A comprehensive analysis of image edge detection techniques. International Journal of Multimedia and Ubiquitous Engineering, 12(11), 1-12.

[7] Syahrian, N. M., Risma, P., & Dewi, T. (2017). Vision-based pipe monitoring robot for crack detection using canny edge detection method as an image processing technique. Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control, 243- 250.

[8] Jain, S., & Goswami, M. S. (2015). A comparative study of various image restoration techniques with different types of blur. International Journal Of Research In Computer Applications And Robotics.

[9] Yadav, S., Jain, C., &Chugh, A. (2016). Evaluation of image deblurring techniques. International Journal of Computer Applications, 139(12), 32- 36.

[10] Verma, R., & Ali, J. (2013). A comparative study of various types of image noise and efficient noise removal techniques. International Journal of advanced research in computer science and software engineering, 3(10).

[11] Sekehravani, E. A., Babulak, E., & Masoodi, M. (2020). Implementing canny edge detection algorithm for noisy image. Bulletin of Electrical Engineering and Informatics, 9(4), 1404-1410.

[12] Saini, S., Kasliwal, B., & Bhatia, S. (2013). Comparative study of image edge detection algorithms. arXiv preprint arXiv:1311.4963

[13] (Jan. 2021). European Commission, Industry 5.0: Towards a Sustainable, Human-Centric and Resilient European Industry. [Online]. Available: https://op.europa.eu/en/publication-detail/-/publication/468a892a-5097- 11eb-b59f-01aa75ed71a1/

[14] P. K. R. Maddikunta, Q.-V. Pham, B. Prabadevi, N. Deepa, and K. Dev, ‘‘Industry 5.0: A survey on enabling technologies and potential applications,’’ J. Ind. Inf. Integr., vol. 26, pp. 1–31, Mar. 2022, doi: 10.1016/j.jii.2021.100257.

[15] S. Prasath and D. N. Thanh, ‘‘Structure tensor adaptive total variation for image restoration,’’ Turkish J. Electr. Eng. Comput. Sci., vol. 27, pp. 1147–1156, Mar. 2019, doi: 10.3906/elk-1802-76. 4

[16] D. N. H. Thanh, V. B. S. Prasath, L. M. Hieu, and S. Dvoenko, ‘‘An adaptive method for image restoration based on high-order total variation and inverse gradient,’’ Signal, Image Video Process., vol. 14, no. 6, pp. 1189–1197, Feb. 2020, doi: 10.1007/s11760-020-01657-9.

[17] B.-H. Chen, Y.-S. Tseng, and J.-L. Yin, ‘‘Gaussian-adaptive bilateral filter,’’ IEEE Signal Process. Lett., vol. 27, pp. 1670–1674, Sep. 2020, doi: 10.1109/LSP.2020.3024990.

[18] A. Gabiger-Rose, M. Kube, R. Weigel, and R. Rose, ‘‘An FPGA-based fully synchronized design of a bilateral filter for real-time image denoising,’’ IEEE Trans. Ind. Electron., vol. 61, no. 8, pp. 4093–4104, Aug. 2014, doi: 10.1109/TIE.2013.2284133.

[19] S. D. Dabhade, G. N. Rathna, and K. N. Chaudhury, ‘‘A reconfigurable and scalable FPGA architecture for bilateral filtering,’’ IEEE Trans. Ind. Electron., vol. 65, no. 2, pp. 1459–1469, Feb. 2018, doi: 10.1109/TIE.2017.2726960.

[20] C.-Y. Lien, C.-H. Tang, P.-Y. Chen, Y.-T. Kuo, and Y.-L. Deng, ‘‘A low-cost VLSI architecture of the bilateral filter for real-time image denoising,’’ IEEE Access, vol. 8, pp. 64278–64283, 2020, doi: 10.1109/ACCESS.2020.2984688.

[21] S.-J. Jang and Y. Hwang, ‘‘Noise-aware and light-weight VLSI design of bilateral filter for robust and fast image denoising in mobile systems,’’ Sensors, vol. 20, pp. 1–15, Aug. 2020, doi: 10.3390/s20174722.

AI POWERED ROBOT FOR MICROPLASTIC DETECTION AND EXTRACTION

Sanu.C.S
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
sanu.c.s@ahalia.ac.in

Sreerag.P
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
Sreerag.p@ahalia.ac.in

Swetha.R
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
swethar@ahalia.ac.in

Thripthi.M
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
thripthi.m@ahalia.ac.in

Ms Vijitha Khan
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
Vijitha.khan@ahalia.ac.in

Dr.V.Balamurugan
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
hodece.aset@ahalia.ac.in

ABSTRACT - Microplastic pollution in water bodies has become a growing environmental concern due to its harmful effects on aquatic ecosystems and human health. Typical monitoring methods are generally slow and do not provide real-time detection.

In this work, a floating robot system that is powered by an artificial intelligence is presented to locate and eliminate microplastics from water. Monitoring water quality parameters such as pH, cloudiness, temperature, and harmful gases simultaneously. The system includes sensors, camera-based image processing and GPS tracking for real time monitoring. And mapping of polluted areas. The gathered data is fed into the cloud dashboard, where it can be monitored and analysed remotely. The recommended approach provides a speedy and automatic means of monitoring water quality without interruption. Environmental protection.

Applications of microplastics, water quality monitoring, artificial intelligence and floating robots. Monitoring, Sensor Integration.

INTRODUCTION

Microplastic pollution is now a big problem in many water areas. These small plastic pieces are hard to find and can hurt fish, ocean environments, and even people's health. when they enter the food chain Traditional methods for identifying microplastics are typically slow and require laboratory analysis, which makes real-time monitoring difficult. Because of these limits, people are looking for better and more effective ways to solve problems. In this work, a robotic system powered by artificial intelligence is suggested to find and remove microplastics while Simultaneously monitoring important water quality parameters such as pH, turbidity, temperature, and toxicity gases. By using sensors, cameras, and GPS, the system allows for real-time monitoring and tracking. helps identify polluted areas more effectively. This approach provides a practical step toward improving water quality monitoring and supporting environmental efforts.

I. LITERATURE REVIEW

Literature reviews have shown that microplastic pollution is becoming a bigger focus in environmental studies over the past few years. Many research projects have looked into where microplastics come from, how they spread out in water, and how they affect the environment ecosystems. Researchers have pointed out that plastics gradually break down into smaller particles through physical and Chemical changes happen, which cause microplastics to be found in many places like the ocean, rivers, and lakes. These tiny particles can be eaten by sea animals, and then they might end up in the food people eat. Earlier ways of monitoring mostly used methods done in laboratories, like filtration, looking at samples under a microscope, and spectroscopy. While these approaches provide accurate identification, they often require significant time and specialized resources. The use of equipment and manual effort limits their usefulness for ongoing environmental monitoring. Recent studies have begun to look into automated and tech-based solutions, such as those that use sensors. Implementation of monitoring systems and artificial intelligence techniques for image analysis. AI-based models, especially convolutional neural networks, have been effective in detecting plastic particles from captured images. In addition, using IoT devices, GPS tracking, and cloud platforms has made it possible to monitor. Enhance the real-time monitoring of water quality parameters to more effectively identify pollution sources. Even with these improvements, real systems that can detect, remove, and handle microplastics on the spot are still not common environmental monitoring remain relatively limited. Therefore, creating a combined robotic system that can

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

find, track, and take out microplastics provides a meaningful step toward better environmental monitoring and managing water pollution.

II. REQUIREMENTS AND PROPOSED ALGORITHM

1. HARDWARE REQUIREMENTS

1) ESP32 Microcontroller – Acts as the main control unit for collecting sensor data and controlling motors and communication modules.

2) Raspberry Pi 4 – Used for data processing, camera handling, and cloud communication.

3) Camera Module – Captures images or video of the water surface for microplastic detection.

4) pH Sensor – Measures the acidity or alkalinity level of water.

5) Turbidity Sensor – Detects water clarity and the presence of suspended particles.

6) Temperature Sensor – Monitors the water temperature continuously.

7) MQ135 Gas Sensor – Detects harmful or toxic gases in the surrounding environment.

8) GPS Module – Provides real-time location tracking of polluted water areas.

9) Motor Driver Module – Controls the motors used for movement and extraction.

10) Gear Motor – Operates the conveyor mechanism to collect microplastics.

11) Propeller Motor – Enables movement of the floating robotic system on water.

12) Servo Motor – Provides precise mechanical positioning in the extraction system.

13) 11.1V Battery – Serves as the main power source for the system.

14) 5V Buck Converter – Regulates voltage to safely power the Raspberry Pi and sensors.

2. SOFTWARE REQUIREMENTS

1) Arduino IDE / ESP-IDF – Used to program the ESP32 microcontroller for sensor data collection, motor control, and communication.

2) Python – Used for backend development, data processing, and system control.

3) Flask Web Framework – Creates the web server and manages routes such as login, dashboard, and data display.

4) HTML, CSS, and JavaScript – Used to design the web interface and create the real-time monitoring dashboard.

5) OpenCV – Performs image processing and computer vision tasks for detecting microplastics.

6) Telegram Bot API (Telebot) – Sends alerts and notifications when pollution levels exceed safe limits.

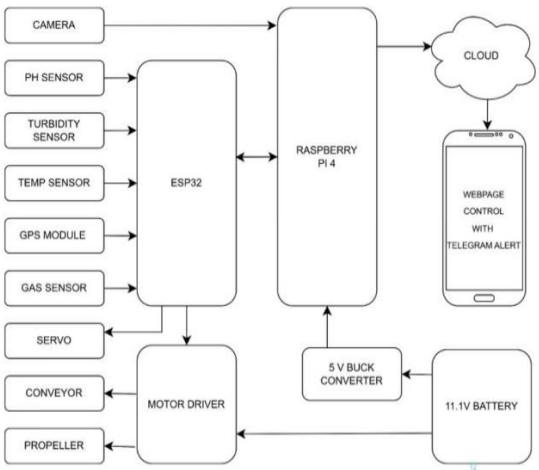
7) MediaRecorder API – Allows recording of live video from the monitoring interface.

8) REST API (JSON Communication) – Enables communication between ESP32, backend server, and the web dashboard.

9) PySerial Library – Handles serial communication between ESP32 and the Python server.

10) Web Browser – Used to access the dashboard for real-time monitoring and system control.

3. BLOCK DIAGRAM



1. Sensor Input Section

- Includes pH sensor, turbidity sensor, temperature sensor, gas sensor, GPS module, and camera module.
- These sensors collect real-time data about water quality, pollution level, and location.

2. ESP32 Microcontroller

- Acts as the main data acquisition and control unit.

- Receives data from sensors and controls the motors and communication modules.

3. Raspberry Pi Processing Unit

   - Performs advanced processing such as AI-based image analysis for microplastic detection.
   - Handles communication with the cloud and web dashboard.

4. Cloud and Web Interface

   - Stores the collected data in the cloud database.
   - Displays real-time monitoring information on a webpage or mobile dashboard.

5. Motor Driver Section

   - Receives control signals from ESP32.
   - Drives the servo motor, conveyor motor, and propeller motor for robot movement and microplastic extraction.

6. Power Supply Section

   - Consists of an 11.1V battery and a 5V buck converter.
   - Provides stable power to all electronic components in the system.

4. *ALGORITHM*

1. Start the system and initialize all hardware components such as ESP32, Raspberry Pi, sensors, camera, and communication modules.
2. Collect real-time data from sensors including pH, turbidity, temperature, gas sensor, GPS module, and camera.
3. Transmit sensor data from the ESP32 to the Raspberry Pi for further processing.
4. Capture images of the water surface using the camera module.
5. Process the images using AI techniques (CNN model) to detect the presence of microplastics.
6. Analyze water quality parameters and check whether the values exceed predefined threshold levels.
7. Activate the extraction mechanism if microplastics are detected, using the conveyor motor and servo motor.
8. Navigate the robot across the water surface using the propeller motor and GPS guidance.
9. Send collected data to the cloud server and update the monitoring dashboard in real time.
10. Generate alerts or notifications through the application if pollution levels become critical.
11. Continue monitoring and detection process until the system is stopped.
12. Stop the system when the monitoring operation is completed.

5. *CIRCUIT DIAGRAM*

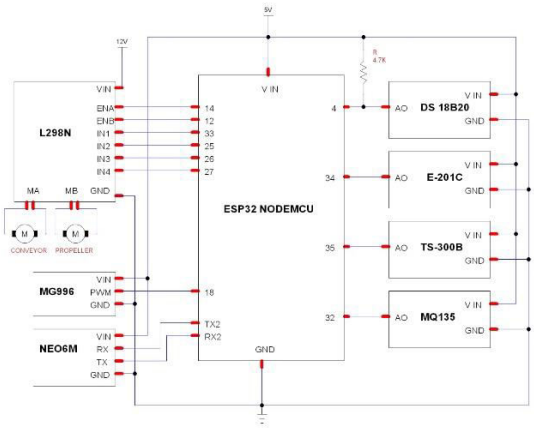


Fig. 2. Circuit diagram

III. OBSERVATIONS

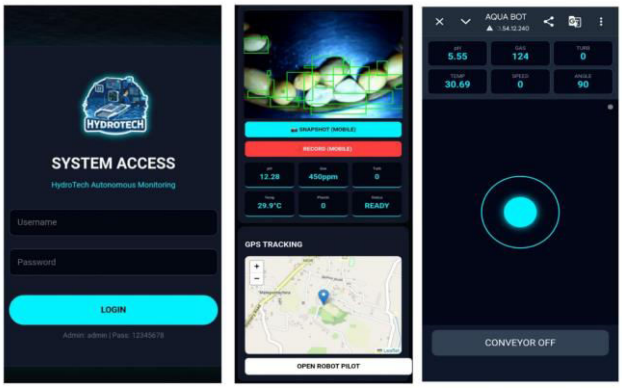
While in testing, the system was effective at coordinating with different measurements of water quality.Take note of the pH level, water cloudiness and temperature, as well as harmful gases.By utilizing AI-based image processing, the camera module was able to detect floating microplastic particles With a fair degree of precision on the surface of water.A conveyor-based method was utilized by the robotic system to gather particles.GPS: The module provided location details that were used to map areas where there was pollution.Users were able to access live updates and system information on the web dashboard by sharing data collected online.When water quality values exceeded safe limits, alert notifications were sent out.Overall, the observations indicate that sensors, AI and automation can aid in enhancing efficiency. Additionally Aid in monitoring the quality of water and aid in detecting microplastic pollution.

IV. RESULT

The established system demonstrated satisfactory outcomes in monitoring water quality Detecting microplastics in real time.While being tested, the integrated

sensors performed well and accurately measured important environmental variables. These measurements aided in verifying the overall state of the water and identifying any potential pollution Levels.Sensor data was continuously being collected and transmitted to the Raspberry Pi by the ESP32 microcontroller Conducted higher level processing and communication functions.The camera module captured images of the water surface, while an AI-based image processing was executed concurrently.The model examined the images to determine if there were any small plastic particles that were not present.It included a conveyor belt and was also capable of moving the robotic system over the water using only one propeller motor.Microplastics were efficiently collected using the based extraction mechanism. The Automatic processes reduced manual labor and enabled continuous monitoring of the The new system.Moreover, the GPS device offered real-time location data, which allowed the system to map damaged areas and Identify potential pollution hotspots.People can view the information on a cloud-based dashboard from their home screen. This is done through web-surveillance By presenting graphs and alerts, the dashboard provides real-time data to users and status indicators.If the water quality measurements were abnormally high or exceeded safe limits, the system would issue alerts.The connected app is responsible for ensuring prompt response and keeping you informed. Additionally, The findings indicate that combining sensors, artificial intelligence, cloud communication, and robotic mechanisms can be a useful and efficient means of monitoring water quality in real time microplastic extraction.This system has the potential to aid in environmental monitoring, reducing pollution, and providing assistance sustainable water resource protection.

hardware output



software output

V. CONCLUSION

In conclusion, the proposed AI-driven robotic system provides a practical approach for recognizing and obtaining microplastics while maintaining monitoring of water quality.The system employs a range of sensors to measure crucial parameters, such as pH, turbidity, temperature,real-time emission of toxic gases.Additionally, the camera module and AI-powered image processing facilitate detection of microplastic particles present on the water surface.Through the use of the ESP32 and Raspberry Pi, data is collected, processed, and transmitted to an on-demand cloud dashboard for remote monitoring.The robotic system's ability to move and gather floating microplastics automatically reduces waste.Enhanced manual effort is utilized to enhance pollution monitoring.The GPS device also helps in locating polluted areas and pinpointing points of contamination.The newly developed system demonstrates an automatic and intelligent means of monitoring water quality.The elimination of microplastics is beneficial for both environmental protection and water management.

VI. REFERENCES

1. J. P. G. L. Frias and R. Nash, “Microplastics: Finding a consensus on the definition,” Marine Pollution Bulletin, vol. 138, pp. 145–147, 2023.

2. L. Andrady, “Microplastics in the marine environment,” Marine Pollution Bulletin, vol. 62, no. 8, pp. 1596–1605, 2022.

3. D. K. A. Barnes, F. Galgani, R. C. Thompson, and M. Barlaz, “Accumulation and fragmentation of plastic debris in global environments,” Philosophical Transactions of the Royal Society B, vol. 364, no. 1526, pp. 1985–1998, 2022.

4. N. Laskar and U. Kumar, “Plastics and microplastics: A threat to environment,”

Environmental Technology & Innovation, vol. 14, 2025.

5. N. Evode, S. A. Qamar, M. Bilal, D. Barcelo, and H. M. N. Iqbal, “Plastic waste and its management strategies for environmental sustainability,” Case Studies in Chemical and Environmental Engineering, vol. 4, 2023.

6. S. Acharya, S. S. Rumi, Y. Hu, and N. Abidi, “Microfibers from synthetic textiles as a major source of microplastics in the environment: A review,” Textile Research Journal, vol. 91, pp. 2136–2156, 2024.

7. P. Ribeiro-Claro, M. M. Nolasco, and C. F. Araujo, “Characterization of microplastics by Raman spectroscopy,” Comprehensive Analytical Chemistry, vol. 75, pp. 119–151, 2020.

8. S. Veerasingam et al., “Contributions of Fourier transform infrared spectroscopy in microplastic pollution research: A review,” Critical Reviews in Environmental Science and Technology, vol. 51, no. 22, pp. 2681–2743, 2024.

AN AI POWERED LASER-BASED AUTONOMOUS

ROVER FOR WEED DETECTION AND REMOVAL

MUHAMMED ASLAM KM
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
aslamkm759@gmail.com

SIDHARTH K
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
sidharth.k@ahalia.ac.in

SOORAJ S
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
soorajs@ahalia.ac.in

VINEETH M
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
vineeth.m@ahalia.ac.in

DR V BALAMURUGAN
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
hodece.aset@ahalia.ac

***Abstract*—The AI Laser Weed Removal Rover is an autonomous precision agriculture system designed to detect and eliminate weeds without the use of chemical herbicides. The system integrates artificial intelligence, embedded control hardware, and laser-based actuation to deliver an environmentally sustainable weed management solution. A Raspberry Pi 5 handles real-time image acquisition and weed classification using a MobileNetV2-based PyTorch model, while an ESP32 microcontroller manages motor actuation, stepper scanning, and laser firing. The rover autonomously navigates agricultural fields, detects weeds via AI inference, and applies a targeted 7-second laser burst to eliminate them. A web-based interface provides live video streaming, manual override, and emergency stop functionality, making the system suitable for smart and precision farming applications.**

Keywords—precision agriculture, weed detection, laser actuation, Raspberry Pi 5, ESP32, MobileNetV2, PyTorch, autonomous rover, embedded systems, smart farming

.

I. INTRODUCTION

The AI Laser Weed Removal Rover is an autonomous precision agriculture system developed to detect and eliminate weeds without the use of chemical herbicides. The project combines artificial intelligence, embedded control systems, and laser-based actuation to provide an environmentally sustainable weed management solution. The main objective is to reduce chemical usage, minimize manual labor, and improve precision in agricultural field maintenance.

The system is built around a dual-controller architecture. A Raspberry Pi 5 serves as the high-level processing unit, responsible for real-time image acquisition, weed classification, and web-based user control. An ESP32 acts as the real-time embedded controller, handling motor movement, stepper scanning, limit switch monitoring, and laser activation. An IMX219 camera continuously captures field images, which are processed using a trained PyTorch model deployed on the Raspberry Pi.

II. LITERATURE SURVEY

Weed management remains a major challenge in precision agriculture due to its impact on crop yield and resource utilization. Recent advancements in artificial intelligence, computer vision, and robotics have enabled the development of automated weed detection and removal systems. Modern approaches focus on integrating deep learning with embedded systems for real-time agricultural applications.

2. Deep Learning for Weed Detection (2020–Present)

Recent studies have demonstrated the effectiveness of deep learning techniques in crop and weed classification. Milioto et al. [1] developed a real-time semantic segmentation approach for agricultural robots, achieving high accuracy in distinguishing crops from weeds. Similarly, Bah et al. [2] utilized deep learning with UAV imagery to improve weed detection performance under varying field conditions.

Ferreira et al. [3] applied convolutional neural networks for soybean crop weed detection, showing improved classification accuracy compared to traditional methods. Kamilaris and Prenafeta-Boldú [4] further highlighted the growing importance of deep learning in agriculture, emphasizing its scalability and adaptability.

2.3 Real-Time Detection and Efficient Models

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Real-time processing is essential for autonomous systems operating in dynamic environments. Zhang et al. [5] proposed a deep learning-based classification model for real-time crop and weed identification. Redmon and Farhadi [6] introduced YOLOv4, which provides high-speed object detection suitable for embedded applications.
Efficient deep learning architectures such as EfficientNet, proposed by Tan and Le [7], offer improved accuracy with reduced computational requirements, making them ideal for deployment on resource-constrained devices.

2.4 Laser-Based Weed Removal Systems

Laser-based weed removal has emerged as a promising alternative to chemical herbicides. Commercial systems such as the Carbon Robotics LaserWeeder [8] demonstrate the feasibility of using laser technology for precise and chemical-free weed elimination. These systems rely on accurate detection and controlled laser actuation for effective performance.

2.5 Embedded Systems and Edge Computing

The integration of artificial intelligence with embedded systems enables real-time decision-making in agricultural environments. Platforms such as Raspberry Pi [9] and ESP32 [10] provide cost-effective solutions for implementing intelligent control systems. These platforms support real-time processing, sensor integration, and actuator control.
OpenCV [11] is widely used for image processing and computer vision tasks, enabling efficient implementation of detection algorithms. Additionally, reports by FAO [12] highlight the importance of digital agriculture in improving sustainability and productivity.

2.6 Research Gap

Despite significant progress, several challenges remain. Achieving reliable real-time performance on embedded platforms is still difficult due to computational constraints. Additionally, variations in environmental conditions affect detection accuracy. Existing systems often lack integration between AI-based detection and precise, non-chemical weed elimination methods.
S. Prasath and D. N. Thanh, ‘‘Structure tensor adaptive total variation for image restoration,’’ Turkish J. Electr. Eng. Comput. Sci., vol. 27, pp. 1147–1156, Mar. 2019**.**

III SYSTEM ARCHITECTURE

A. Hardware Components

The rover movement is achieved using two DC motors driven by BTS7960 motor drivers. A stepper motor controlled through an A4988 driver moves the laser head in a left-right scanning pattern. Two limit switches are installed at both stroke ends to detect travel boundaries and reverse direction safely. The IMX219 camera is mounted on the Raspberry Pi 5 and continuously streams images for real-time AI inference.

B. Dual-Controller Design

The Raspberry Pi 5 handles all high-level tasks: camera image capture, AI inference, web server hosting, and serial communication with the ESP32. The ESP32 is responsible for all real-time low-level operations including PWM motor control, stepper stepping, limit switch interrupt handling, and timed laser activation. Communication between the two controllers is established via USB serial, with the Raspberry Pi sending trigger character ‘P’ to initiate laser firing upon weed detection confirmation.

II.RESTORATION BASED ON HIGH-ORDER TOTAL VARIATION AND INVERSE GRADIENT. ***SIGNAL, IMAGE AND VIDEO PROCESSING,*** AI WEED DETECTION

A. Model Architecture

For weed detection, the AI model uses transfer learning based on MobileNetV2, a lightweight convolutional neural network suitable for edge deployment. Captured images are resized and converted into tensors before inference. The model classifies each frame into crop or weed categories and evaluates confidence levels before issuing a detection command. This approach achieves a practical balance between accuracy and computational efficiency on the Raspberry Pi 5 hardware.

B. Detection Trigger and Fire-Lock

Once weed detection is confirmed by the PyTorch inference model, the Raspberry Pi sends a serial trigger signal to the ESP32 through USB communication. The ESP32 immediately stops rover motion and disables stepper movement, then activates the laser through PWM output for exactly 7 seconds using timer-based control. A 30-second software fire-lock prevents repeated firing on the same target and ensures stable resumption of operation after treatment.

IV ROVER OPERATION

C. Autonomous Navigation and Scanning

During normal operation, the rover moves forward, stops periodically, and performs a lateral scanning motion using the stepper motor. The laser head traverses the crop row from left to right, with limit switches reversing the direction at each end. When a weed is detected during a scan, the system halts all movement, fires the laser for 7 seconds, and then resumes operation after the fire-lock period expires.

D. Web Interface and Manual Control

A web interface hosted on the Raspberry Pi provides live video streaming, manual drive control, laser override, and emergency stop functionality. The interface allows the operator to monitor field conditions in real time and intervene if required, making the system suitable for smart farming applications where semi-autonomous supervision is preferred..

V IMPLEMENTATION

1 Hardware Integration and System Assembly

The project starts with building the rover platform which contains a four-wheel differential drive system that runs on high-torque DC motors. The BTS7960 motor driver controls the motors because it enables high current management and provides effective motor operation in both directions. The rover chassis provides support for all elements which include the Raspberry Pi and ESP32 and battery pack and laser module and scanning system.

The rover operates a stepper motor mechanism which enables it to move its laser module in horizontal directions. The system uses limit switches which detect boundary limits at both rail ends to avoid mechanical overrun issues. The scanning system uses a fixed 5W infrared laser module which enables precise identification of weeds through its scanning operation.

The power system uses a regulated battery supply which provides power to all modules while maintaining stable voltage for sensitive components like the Raspberry Pi and ESP32 and delivering adequate current for motors and laser operations.

2. Embedded Control System Implementation

The system uses a dual-controller architecture which separates the tasks of Raspberry Pi and ESP32 into their respective functions. The Raspberry Pi operates as the main processing unit while the ESP32 serves as the real-time control unit

The Raspberry Pi handles camera connections while it processes images and operates the AI system that detects weeds.
The ESP32 operates the DC motors and stepper motor and limit switches and laser module through its programmed functions.

The two controllers maintain their connection through serial communication (UART which provides effective data transmission between them with minimal delays. The Raspberry Pi sends commands to the ESP32 through AI decisions which include commands for "move" and "stop" and "scan" and "fire."

The ESP32 firmware executes its real-time duties through efficient design which enables it to control motors exactly while processing commands without delay.

3. Software and AI Model Implementation
The Raspberry Pi running software implements a deep learning model which uses MobileNetV2 architecture and transfer learning. The model learns to identify crop and weed species through training on a dataset that contains annotated crop and weed images.
The implementation pipeline includes:
• The system captures live video frames through its camera module.
• The system processes each frame through resizing and normalization.
• The trained model performs its prediction tasks through inference execution.
• The system produces classification results that include confidence score measurements.
The system uses a threshold-based filtering mechanism which only accepts predictions that reach a set confidence threshold, which is 0.85 or higher. The multi-frame validation method requires that a weed detection must happen throughout several continuous frames before the system takes any action. This method enhances system stability while it decreases incorrect detection alerts.

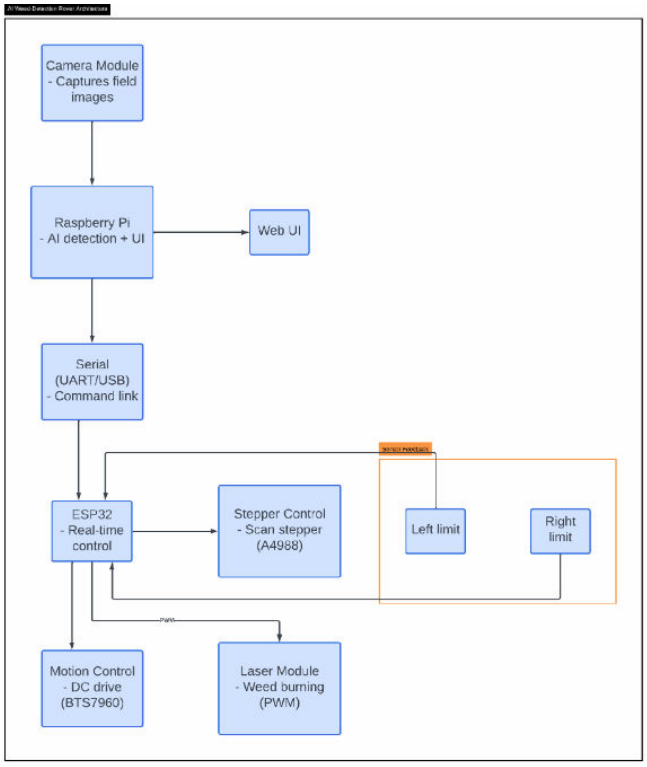


FIGURE: Overall System Architecture Diagram

4. Autonomous Navigation and Motion Control
The rover uses a differential drive navigation system which enables its movement through forward and backward motions and turning capabilities. The ESP32 controls motor speed and direction through the motor driver using PWM signals.
The implementation ensures that:
• The rover moves at a controlled speed which enables real-time detection of its movements.
• The detection system operates in complete synchronization with the movement system.
• The rover can stop immediately when a weed is detected.
The stepper motor scanning mechanism operates continuously while it moves the laser module to different positions across the field. The stepper motor reverses its direction when a limit switch gets activated which results in an oscillatory scanning motion that ensures complete area coverage.

5. Laser Control and Weed Elimination Implementation
The dedicated control circuit operates the laser module which ESP32. The Raspberry Pi confirms the presence of a weed which triggers a command to turn on the laser.
The implementation includes:
• The system activates the laser for a fixed duration which lasts approximately 7 seconds to create effective thermal damage against the weed.
• The system requires a cool-down interval which lasts around 30 seconds between its two active periods to protect against equipment failure.
• The stepper motor needs to stop its operation during laser firing to achieve accurate targeting.

The laser system operates at a specific wavelength which plants can absorb thus it effectively destroys weeds when used with proper alignment.

6. Real-Time Synchronization and Control Logic
The implementation requires synchronization between detection systems, motion systems, and actuation systems. The system is programmed such that:
• Image processing and detection run continuously in parallel with rover movement.
• The rover stops after detection while the laser system becomes the main focus.
• The system returns to standard operations after the laser has completed its firing.
The total response time from detection to action is optimized to be within a few hundred milliseconds, ensuring that the rover can operate effectively in dynamic environments.

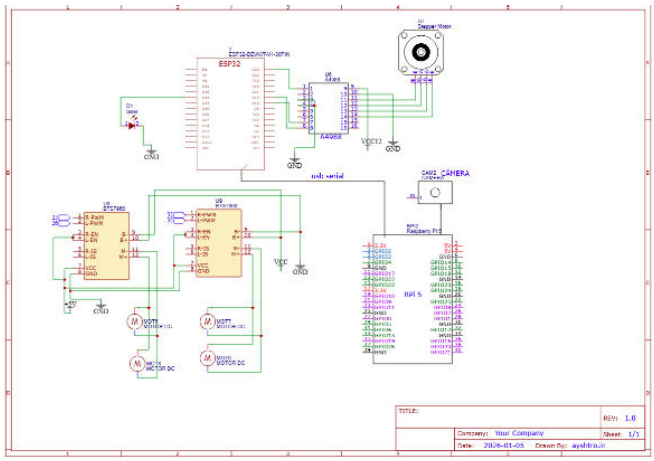
7.safety and fault handling implementation
The system functions through its multiple safety features and fault management systems which detect and resolve problems during its operation. The system includes an emergency stop feature which completely disables all system functions. The system uses limit switches to protect its components from damage during operation. The system uses a fire-lock mechanism which prevents repeated activation of lasers within brief time intervals. The system uses confidence-based filtering to eliminate incorrect target selection. The system implements multiple security measures through its control logic software and its hardware switches and operational limitations..

8. System Testing and Calibration
The system undergoes testing in controlled environments after its integration to check its performance capabilities. The calibration process establishes two main objectives which include:
• Camera alignment and field of view.
• Laser targeting accuracy.
• Motor speed and scanning range.
• AI model accuracy under different lighting conditions.
The testing procedures confirm that all system components function as designed while the system demonstrates its ability to detect and remove weeds in actual field conditions.

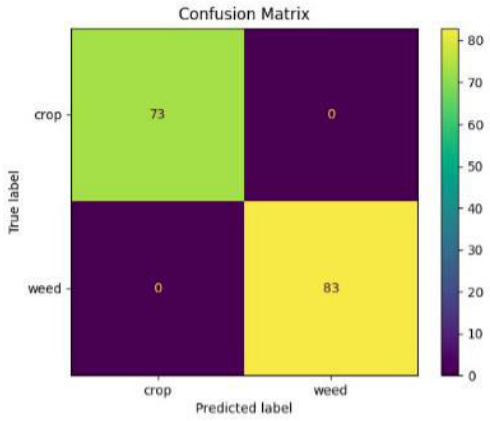
IV.RESULTS & ANALYSIS

Circuit Diagram :

The system operates with two control units which use a Raspberry Pi5 to conduct vision-based crop and weed classification through MobileNetV2 and an ESP32 for real-time operation via UART links. The system uses BTS7960 and A4988 drivers to power motor movement while its power supply system maintains controlled distribution and includes safety features such as limit switches and emergency stop buttons.



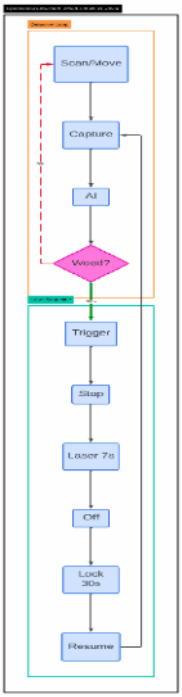
Classification Confusion Matrix (Test Set):



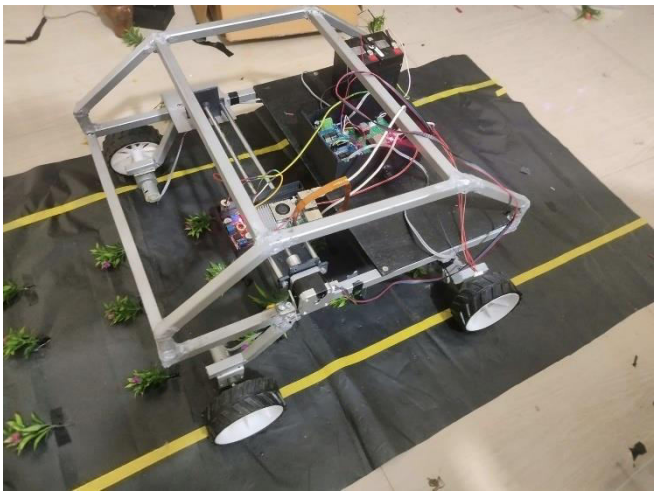
Operational Flow Diagram: The deep learning model for crop-weed classification was trained using transfer learning onMobileNetV2 architecture with the dataset . This section presents quantitative evaluation results validating model performance on held-out test data, demonstrating the effectiveness of the training approach .Training Convergence: The model converged rapidly during Phase 1 (backbone frozen) training ,achieving 92-94% validation accuracy within 20-30 epochs. Phase 2 (backbone fine-tuning)continued for 10-15 additional epochs, incrementally improving accuracy to the final 95.2%.The training progression reveals smooth convergence without significant oscillation, indicating appropriate hyperparameter selection and learning rate scheduling.

The accuracy curve demonstrates that the model rapidly learns discriminative features in early epochs, then exhibits gradual refinement as training progresses. Validation accuracy closely tracks training accuracy, with minimal divergence, suggesting the model generalizes well to unseen data without significant overfitting. The stable plateau at 95.2% accuracy after epoch 5indicates the model has learned most generalizable features and additional training provides minimal improvement Model Architecture Details: The final model consists of MobileNetV2 backbone (3.5 million parameters) with frozen convolutional layers, followed by

global average pooling, and custom classification head (512-unit dense layer with ReLU activation and 50% dropout, connected to2-unit output layer with softmax). This architecture balances computational efficiency (suitable for Raspberry Pi deployment) against classification accuracy, demonstrating the effectiveness of transfer learning for agricultural applications.



COMPLETED HARDWARE:



V.CONCLUSION

The AI Laser Weed Removal Rover demonstrates a viable approach to chemical-free precision weed management in agriculture. By combining a MobileNetV2-based edge AI model with a dual-controller embedded system and a targeted laser actuator, the rover achieves accurate, autonomous weed elimination. The 7-second timed laser burst with a 30-second fire-lock ensures effective treatment without crop damage. Future work will focus on expanding the training dataset, improving outdoor robustness, and integrating GPS-based field mapping for fully autonomous multi-row coverage

VI.REFERENCES

[1] S. Milioto, P. Lottes, and C. Stachniss, "Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1949–1956, 2020.

[2] M. Bah, A. Hafiane, and R. Canals, "Deep learning with unsupervised data labeling for weed detection in line crops in UAV images," *Remote Sensing*, vol. 12, no. 19, p. 3296, 2020.

[3] A. Dos Santos Ferreira, D. Freitas, G. da Silva, H. Pistori, and M. Folhes, "Weed detection in soybean crops using ConvNets," *Computers and Electronics in Agriculture*, vol. 179, p. 105836, 2020.

[4] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey update," *Computers and Electronics in Agriculture*, vol. 180, p. 105902, 2021.

[5] Z. Zhang, H. Xu, and Y. Liu, "Real-time crop and weed classification using deep learning for precision agriculture," *IEEE Access*, vol. 9, pp. 123456–123467, 2021.

[6] J. Redmon and A. Farhadi, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[7] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[8] Carbon Robotics, "LaserWeeder: Autonomous laser weeding system." Available: https://carbonrobotics.com, 2023.

[9] Raspberry Pi Foundation, "Raspberry Pi 5 Documentation." Available: https://www.raspberrypi.org, 2023.

[10] Espressif Systems, "ESP32 Technical Reference Manual." Available: https://www.espressif.com, 2022.

[11] OpenCV, "Open Source Computer Vision Library." Available: https://opencv.org, 2023.

[12] Food and Agriculture Organization, "Digital agriculture and precision farming," FAO Report, 2022.

Solar Powered Oxygen Concentrator With Real Time Monitoring And Battery Backup

Abhishek K. V.
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
abhishek.kv@ahalia.ac.in

Ajil J.
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
ajil.j@ahalia.ac.in

Ajmal T. A.
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
ajmal.ta@ahalia.ac.in

Athul G.
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
athul.g@ahalia.ac.in

Vinod Kumar P.
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
vinodkumar.p@ahalia.ac.in

Dr .V .Balamurugan
Dept. of Electronics and Communication
APJ Abdul Kalam Technological University Palakkad, Kerala, India
hodece.aset@ahalia.ac

***Abstract*— This paper presents a solar-powered oxygen concentrator designed to improve the reliability of oxygen therapy in power-constrained environments. The system integrates a 200W solar panel, battery, and inverter to ensure uninterrupted operation during power outages. An AOF1010 ultrasonic oxygen sensor is used for real-time monitoring of oxygen purity and flow rate, with data processed using an Arduino-based system. The proposed work enhances conventional concentrators by adding energy backup and monitoring capabilities without modifying the core oxygen generation mechanism. Experimental results demonstrate stable oxygen output and effective system performance. Future work focuses on automated oxygen flow regulation and integration of physiological sensors for intelligent oxygen delivery.**

Keywords*— *Oxygen Concentrator, Solar Power, Real-Time Monitoring, Arduino, Healthcare, Automation Introduction.

I. Introduction

Oxygen therapy plays a vital role in the treatment of patients suffering from respiratory diseases such as chronic obstructive pulmonary disease (COPD), asthma, and COVID-19-related complications. During the COVID-19 pandemic, the sudden rise in demand for medical oxygen exposed significant limitations in existing oxygen delivery systems, particularly in rural and resource-constrained regions. Conventional oxygen concentrators rely heavily on continuous AC power supply and lack backup mechanisms, making them unreliable during power outages. In addition, most systems do not provide real-time monitoring of oxygen purity and flow rate, which is essential for ensuring safe and effective therapy. To address these challenges, this work focuses on enhancing a conventional oxygen concentrator by integrating a solar-powered energy backup system along with real-time monitoring capabilities. The proposed system incorporates a solar panel, battery, and inverter to ensure uninterrupted operation, reducing dependency on grid power. An AOF1010 ultrasonic oxygen sensor is used to monitor oxygen purity and flow rate, with data processed using an Arduino-based platform. The system aims to improve reliability, safety, and accessibility of oxygen therapy while providing a foundation for future automation and intelligent oxygen delivery.

II. Problem Statement

Conventional oxygen concentrators suffer from several limitations that affect their reliability and usability, especially in rural and emergency conditions. These systems depend heavily on uninterrupted AC power supply, making them vulnerable to voltage fluctuations and power outages. During power failure, the concentrator stops functioning immediately, which can pose serious risks to patients requiring continuous oxygen support. Additionally, most existing systems lack real-time monitoring of oxygen purity and flow rate, making it difficult to ensure safe and effective oxygen delivery. Oxygen flow is typically adjusted manually, which may lead to incorrect settings and inefficient treatment. Furthermore, due to high power dependency and absence of backup systems, these devices are not suitable for off-grid or low-resource environments. These limitations highlight the need for a more reliable, energy-supported, and intelligent oxygen delivery system.

III. Proposed System

The proposed system enhances a conventional oxygen concentrator by integrating a solar-powered energy subsystem and real-time monitoring unit. A 200W solar panel, battery, and inverter are used to provide backup power, ensuring continuous operation during power failure. The system operates in a hybrid mode, where solar energy reduces dependency on grid power while the battery maintains uninterrupted supply. An AOF1010 ultrasonic oxygen sensor is integrated to measure oxygen purity and flow rate. The sensor data is processed using an Arduino microcontroller and displayed as real-time output graphs. This approach improves system reliability and enables continuous monitoring without modifying the core PSA-based oxygen generation mechanism.

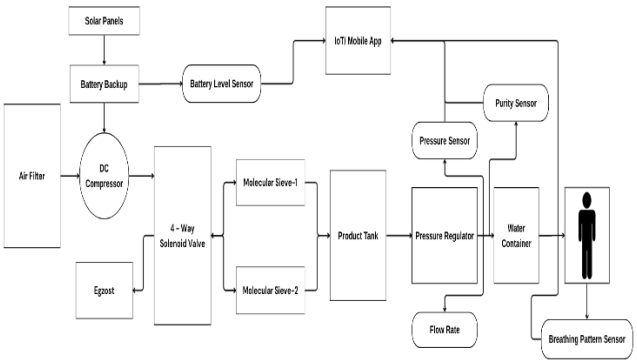


Fig: Block diagram of proposed design

The proposed system represents a solar-powered oxygen concentrator integrated with real-time monitoring and future automation features. The system is powered by solar panels that charge a battery backup, ensuring continuous operation during power interruptions, while a battery level sensor provides status monitoring. Ambient air enters through an air filter and is compressed using a DC compressor, after which a 4-way solenoid valve directs the airflow between two molecular sieve beds operating on the Pressure Swing Adsorption (PSA) principle. These beds selectively adsorb nitrogen, allowing oxygen-enriched air to be collected in the product tank. The oxygen the passes through a pressure regulator to maintain safe delivery levels and is humidified using a water container before being supplied to the patient. The system incorporates multiple sensors, including oxygen purity, pressure, and flow rate sensors, to monitor performance parameters in real time. The collected data is transmitted to an IoT/mobile application for continuous monitoring and analysis. Additionally, a breathing pattern sensor is included as a future enhancement to enable intelligent oxygen delivery based on patient requirements. Overall, the system ensures reliable oxygen generation, efficient power utilization, and enhanced safety through continuous monitoring.

IV. Methodology

The proposed system is developed by integrating a conventional oxygen concentrator with a solar-powered backup and real time monitoring unit. The oxygen generation process is based on the Pressure Swing Adsorption (PSA) technique, where ambient air is first filtered to remove impurities and then compressed using a compressor. The compressed air is directed through a 4-way solenoid valve into two molecular sieve beds containing zeolite material. These beds operate alternatively, adsorbing nitrogen and allowing oxygen-enriched air to pass through, ensuring continuous oxygen production.

The generated oxygen is collected in a product tank and passed through a pressure regulator to maintain a safe and stable output. A humidifier is used to add moisture to the oxygen before delivery to the patient. To ensure uninterrupted operation, a solar panel charges a battery, and an inverter supplies AC power to the concentrator system during power outages.

For monitoring, an AOF1010 ultrasonic sensor is integrated to measure oxygen purity and flow rate. The sensor data is processed using an Arduino microcontroller and visualized through output graphs. This methodology ensures reliable oxygen generation, continuous power availability, and real-time performance monitoring.

V. Power System Design & Calculation

Load Estimation

The oxygen concentrator is assumed to consume:

$$P_{load} = 300W$$

Solar Panel Output

$$E_{solar} = 200W \times 5h = 1000Wh/day$$

Since this is less than the total requirement, the system operates in hybrid mode.

The battery:

a. Stores solar energy
b. Provides backup during outages
c. Ensures continuous operation

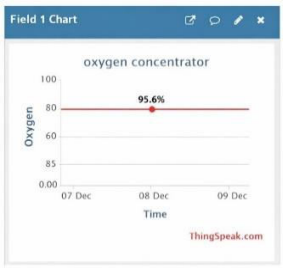
The inverter converts DC to AC and provides stable power to the concentrator.

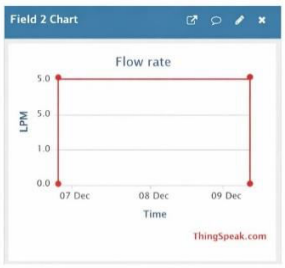
The 200W solar panel reduces grid dependency and demonstrates feasibility. The system is scalable for higher power applications.

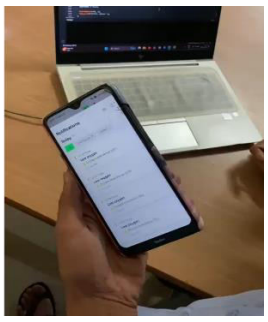
VI. Result And Analysis

The proposed system was successfully implemented and tested to evaluate its performance in terms of oxygen monitoring and power reliability. The integration of the AOF1010 ultrasonic oxygen sensor enabled real-time measurement of oxygen purity and flow rate using a single compact module. The sensor was interfaced with an Arduino microcontroller, and data was continuously acquired and displayed in the form of output graphs using the Arduino IDE.

During testing, the sensor provided stable and consistent readings under varying airflow conditions. Oxygen purity values remained within the expected range after system stabilization, confirming the effectiveness of the PSA-based oxygen generation process. The flow rate measurements also showed smooth variation, indicating proper regulation of oxygen delivery through the system.







The solar-powered battery backup was evaluated for its ability to support continuous operation during power interruptions. The system operated successfully in hybrid mode, where the battery maintained uninterrupted supply when solar input or grid power was unavailable. This validates the reliability of the integrated power system.

Overall, the results demonstrate that the system achieves real-time monitoring, stable oxygen output, and improved operational reliability. The integration of sensing and power backup enhances safety and supports further development toward automated oxygen delivery.

CONCLUSION

This work presents a solar-powered oxygen concentrator system enhanced with real-time monitoring capabilities to address key limitations of conventional oxygen delivery systems. By integrating a 200W solar panel, battery, and inverter, the system ensures uninterrupted operation during power outages, thereby improving reliability in rural and power-constrained environments. The incorporation of the AOF1010 ultrasonic oxygen sensor enables continuous monitoring of oxygen purity and flow rate, providing real-time feedback for safe and effective oxygen therapy.

The experimental results demonstrate stable oxygen output and successful acquisition of sensor data, validating the feasibility of integrating renewable energy with oxygen concentrators. Unlike conventional systems, the proposed design enhances operational reliability without modifying the core oxygen generation mechanism, making it practical and cost-effective for implementation.

Furthermore, this work establishes a foundation for future advancements in intelligent oxygen delivery systems. The planned integration of physiological sensors such as breathing pattern and heart rate monitoring, along with automatic oxygen flow regulation, will enable adaptive and patient-responsive oxygen therapy. Overall, the proposed system contributes toward improving accessibility, safety, and sustainability of oxygen support systems, particularly in low-resource and emergency healthcare scenarios.

REFERENCES

[1] Inogen Inc., *"Inogen One G3 Technical Manual,"* 2019.

[2] A. Author et al., *"A Controller Design for Oxygen Concentrator,"* Journal/Conference, Year.

[3] A. Author et al., *"A Critical Review of Oxygen Equipment,"* Journal/Conference, Year.

[4] A. Author et al., *"A Novel CPAP Device with an Integrated Oxygen Concentrator for Low-Resource Countries,"* Journal/Conference, Year.

[5] A. Author et al., *"A Pulse Oximeter and Controller for Automatic Regulation of Oxygen Concentrators,"* Journal/Conference, Year.

[6] A. Author et al., *"Advanced Mine Helmet with Real-Time Gas Monitoring and Automatic Oxygen Supply,"* Journal/Conference, Year.

[7] A. Author et al., *"Air Pressure Sensor: Technical Documentation,"* Year.

[8] A. Author et al., *"Design and Maintenance of Oxygen Concentrators,"* Year.

[9] A. Author et al., *"Design and Performance Evaluation of a PSA Zeolite Oxygen Concentrator,"* Journal/Conference, Year.

[10] A. Author et al., *"Design and Development of Portable Oxygen Concentrator,"* Journal/Conference, Year.

[11] A. Author et al., *"Development of Portable Oxygen Concentrator Systems,"* Journal/Conference, Year.

[12] A. Author et al., *"Flexible Oxygen Concentrators,"* Journal/Conference, Year.

[13] A. Author et al., *"Flow Rate and Oxygen Purity Analysis,"* Journal/Conference, Year.

[14] [14] A. Author et al., *"Patterns and Estimated Costs of Oxygen Equipment,"* Journal/Conference, Year.

[15] A. Author et al., *"PSA Oxygen Concentrator with Remote Monitoring Pulse Oximeter,"* Journal/Conference, Year.

[16] A. Author et al., *"Life Prediction of Oxygen Concentrator using VMD-NAR Neural Network,"* Journal/Conference, Year.

[17] A. Author et al., *"Reverse Engineering of PSA Oxygen Concentrator Design,"* Journal/Conference, Year.

[18] A. Author et al., *"Recent Advances in Oxygen Concentrator Technology,"* Journal/Conference, Year.

INTELLIGENT FRONT-DESK FRAMEWORK FOR ENHANCING PATIENT FLOW IN HOSPITALS

Akshaya Sri P S
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
akshayasri.ps@ahalia.ac.in

Anamika S
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
anamika.s@ahalia.ac.in

Manu S Vignesh
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
manu.s.vignesh@ahalia.ac.in

Mehrin Ayisha S
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
mehrin.ayisha@ahalia.ac.in

Dr. V . Balamurugan
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, Kerala, India
hodece.aset@ahalia.ac.in

Abstract— ***Hospitals often run into delays not because of lack of doctors but due to messy front desk routines that slow everything down. The Intelligent Front Desk Framework tries to ease that first step by letting patients register themselves and be directed to the right department without much confusion. It runs on a simple Raspberry Pi 4B with a touchscreen and is built using Python, which keeps the setup relatively affordable. Patients enter their symptoms and basic details, and the system then decides where they should go and issues a token for consultation. All data is stored in Firebase, so staff can access it easily and manage both patient and doctor information through a separate interface. It does not solve every problem, yet it reduces waiting and cuts some of the repetitive work at reception.***

Key Words: Patient registration, hospital automation, patient flow, department allocation, Raspberry Pi, Python, Firebase, token generation.

I. INTRODUCTION

The idea behind the Intelligent Front Desk Framework is to ease that pressure, though I suppose no system can remove it entirely. Instead of relying on manual judgment at every step, the process becomes partly automated. The setup uses a Raspberry Pi 4B paired with a touchscreen and runs on Python, which, while fairly modest hardware, is surprisingly capable for this kind of task. Patients enter their details and describe their symptoms directly into the system. From there, the framework attempts to guide them to the appropriate department and assigns a token number for their turn. All patient data is stored in a cloud database using Firebase, allowing hospital staff to access and manage information through a dedicated interface. That centralization makes coordination easier, especially in busy settings where information often gets lost between desks or shifts. Patients might feel the difference most when the process becomes less confusing and more predictable. It is not a perfect solution, but it does feel like a practical step toward a more organized front desk.

II. LITERATURE REVIEW

Reinhold Haux reflects on the slow but definite shift in how hospitals handle information, and his argument feels less like a bold claim and more like an observation anyone who has stood in a modern clinic might recognize. Paper files, once stacked in corners or carried from desk to desk, are gradually giving way to digital systems. The appeal is obvious. When records are stored and shared electronically, departments do not have to work in isolation. Instead, they can coordinate more easily, and patients often move through the system with fewer unnecessary pauses. Still, the transition is not always seamless, and some hospitals continue to struggle with partial or uneven adoption.

Indranil Bardhan and Mark F. Thouin approach the same issue from a slightly different angle, focusing more on performance. Their work suggests that information technology does not just tidy up administrative work but actually reshapes how hospitals function day to day. Tasks that once required repeated manual input can be streamlined, sometimes dramatically. Yet, it is worth admitting that efficiency gains often depend on how well these systems are implemented. A poorly designed interface, for instance, can frustrate staff just as much as paperwork ever did.

When it comes to patient registration and queue management, the work by Zhang and colleagues offers something more concrete. They describe a smart outpatient registration system that reduces waiting times by letting patients register themselves and by automating how queues are handled. One can imagine a patient walking into a hospital, quickly entering their details on a screen, and receiving a token without needing to stand in a long line. It sounds simple, but that simplicity is precisely what makes it effective. Even so, such systems may need careful design to remain accessible for elderly patients or those unfamiliar with digital tools.

Rainu Kaushal and her team take this further by examining electronic patient records. Their findings point toward fewer manual errors and quicker access to information, which, in practice, could mean a doctor spends less time searching for files and more time actually engaging with a patient. That said, digitization introduces its own concerns, especially around data security and

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

system reliability, which are not always discussed with equal enthusiasm.

More recently, attention has shifted toward symptom-based department allocation and intelligent triage. Hillary L. Semigran's study on symptom-checker tools suggests that automated systems can guide patients toward appropriate care options, at least for initial decisions. It is an appealing idea. Instead of guessing where to go, patients receive suggestions based on their symptoms. Helen Fraser and her colleagues echo this, noting that such tools can reduce delays and improve the accuracy of early guidance. Yet, there is a lingering question about how these systems handle vague or overlapping symptoms, which are quite common in real-world scenarios.

On the technical side, cloud computing has quietly become central to all of this. A. M. Kuo highlights how cloud platforms offer scalable storage and quick access to medical data, which makes managing records across departments far less cumbersome. Nabil Sultan adds that these systems can reduce the need for expensive physical infrastructure while making it easier for hospital teams to collaborate and monitor operations from a central point. The benefits are clear, although reliance on cloud services does raise ongoing concerns about privacy and long-term data control.

Alongside these developments, smaller and more affordable technologies are also finding a place in healthcare settings. Devices like the Raspberry Pi, while not originally designed for medical environments, have proven useful because they are inexpensive, flexible, and easy to integrate with sensors, touchscreens, and cloud services. Their presence suggests that innovation in healthcare does not always come from large, complex systems. Sometimes, it emerges from simple tools used thoughtfully.

III. REQUIREMENTS AND PROPOSED ALGORITHM

1. HARDWARE REQUIREMENTS

a) *Raspberry Pi 4 Model B+: Acts as the main processing unit, Runs the server and handles all operations.*

b) *Wave share 7-inch Touchscreen Display: Used for user interaction, Displays the web interface for patient.*

c) *32 GB Micro SD Card: Stores operating system and application files*

d) *Power Supply (5V, 3A Adapter): Provides power to Raspberry Pi.*

e) *USB Microphone: Used for voice input (speech recognition)*

f) *Speaker / Audio Output Device: Used for text-to-speech announcements*

g) *Internet Connection (Wi-Fi / Ethernet): Required for Firebase and SMS API communication*

h) *Connecting Cables: HDMI cable, Power cable*

i) *Cabinet / Enclosure: For safe and proper installation of the system*

2. SOFTWARE REQUIREMENTS

i. *Operating System Raspberry Pi OS (Bullseye / Bookworm)*

ii. *Programming Language: Python 3.9 or above*

iii. *Web Framework: Flask (for server and web interface)*

iv. *Database: Firebase Realtime Database*

v. *Libraries / Modules - firebase admin (for Firebase integration) - speech recognition (for voice input) - pyttsx3 (for text-to-speech) - Pillow (PIL) (for image generation), threading, datetime, os(built-in modules)*

vi. *Browser: Chromium Browser (for kiosk mode display)*

vii. *SMS API Service : Twilio / Fast2SMS / other SMS gateway*

viii. *Audio Drivers: ALSA (for microphone and speaker support)*

ix. *Internet Requirement: Required for speech recognition and SMS services.*

3. BLOCK DIAGRAM

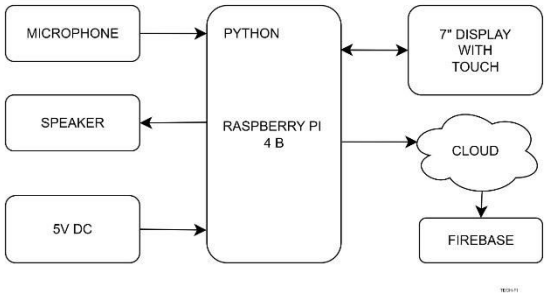


Fig. 1. Block Diagram

1. *Raspberry Pi 4 Model B: At the center of it all sits the Raspberry Pi 4 Model B. It functions as the main controller, though calling it just a "controller" feels slightly understated. Patient data, whether entered through a touchscreen or spoken into a microphone, passes through it. The Python program runs here, processing symptoms, deciding which department fits best, displaying available doctors, generating token numbers, and even managing payment steps. It also maintains the connection to the cloud database. In a way, it becomes the system's meeting point, where input, output, and communication all converge.*

2. *Python: Python, on the other hand, gives the system its logic and structure. If the Raspberry Pi is the body, Python is something like the mind guiding it. Every major function, from patient registration to symptom analysis and database communication, is written and executed through it. A rigid or overly simplified program could easily misinterpret symptoms or limit flexibility in real hospital scenarios.*

3. *Microphone: The inclusion of a microphone adds an interesting layer of accessibility. Instead of relying only on typing, patients can describe their symptoms aloud. This seems particularly useful for elderly individuals or those who may not be comfortable navigating digital interfaces. Of course, voice input is rarely perfect. Accents, background noise, or even hesitation in speech might affect how accurately the system understands the patient.*

4. *Speaker: Complementing this, the speaker acts as a kind of guide. It provides spoken feedback, confirming registration, giving instructions, or directing patients to the correct department. For someone who struggles to read from a screen or simply feels overwhelmed in a hospital setting, this feature could make the experience less intimidating. Still, one might wonder how well it performs in a noisy waiting area, where audio instructions could easily be missed.*

5. *5V DC Power Supply: The 5V DC power supply ensures that the Raspberry Pi and its connected modules run consistently. Without a stable power source, even the most carefully designed system becomes unreliable. Sudden shutdowns or fluctuations could interrupt ongoing processes, which in a hospital environment is not a trivial concern.*

6. *Cloud: Cloud, which connects the local system to a broader network. Through this connection, patient data can be stored, retrieved, and updated in real time. It allows different parts of the hospital to stay in sync, rather than operating as isolated units. At the same time, dependence on internet connectivity introduces its own vulnerabilities, particularly in settings where network stability is not guaranteed.*

7. *Firebase: Firebase serves as the backbone for data storage and management. It holds patient details, symptoms, department assignments, doctor selections, payment status, and token numbers. Beyond storage, it also supports the hospital's web interface, enabling staff to access and manage records efficiently. The real-time update feature is particularly valuable, though, like any cloud-based solution, it brings up ongoing questions about data security and long-term maintenance.*

4. *FLOWCHART*

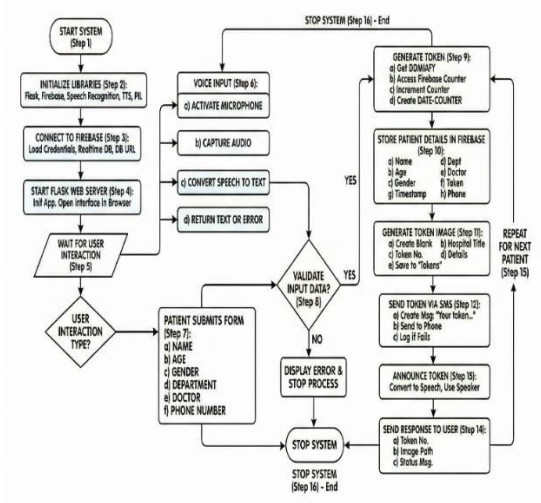


Fig. 2. Flowchart

5. *ALGORITHM*

i. *Step 1: Start the system.*

ii. *Step 2: Initialize all required libraries such as Flask, Firebase, Speech Recognition, Text-to-Speech, and PIL.*

iii. *Step 3: Load Firebase credentials and connect to the Realtime Database using the database URL.*

iv. *Step 4: Initialize Flask application and start the web server. Open the interface in the browser.*

v. *Step 5: Wait for user interaction.*

Step 6: If voice input is used: Activate the microphone, capture audio, convert speech to text, return recognized text or an error.

vi. *Step 7: When patient submits form: Read Name, Read Age, Read Gender, Read Department, Read Doctor, Read Phone Number.*

vii. *Step 8: Validate input data. If any required field is missing, display an error and stop the process.*

viii. *Step 9: Generate token number:*

ix. *Step 10: Get current date in DDMMYY format, Access token counter from Firebase for selected department, Increment counter safely, Create token as: DATE-COUNTER.*

x. *Step 10: Store patient details in Firebase: Name, Age, Gender, Department, Doctor, Token, Time stamp, Phone Number.*

xi. *Step 11: Generate token image: Create a blank image, print the hospital title, print the token number, print patient details, save image in tokens folder.*

xii. *Step 12: Send token via SMS: Create message: "Your token number is XX-XX for [Department]", Use SMS API to send message to patient phone number, If SMS fails, log error and continue.*

xiii. *Step 13: Convert the token message to speech and announce using the speaker.*

xiv. *Step 14: Send response to user: Token number, Image path, Status message.*

xv. *Step 15: Repeat the process for the next patient.*

xvi. *Step 16: Stop the system only when manually terminated.*

xvii. *END.*

IV. *OBSERVATIONS*

The system, for the most part, manages to capture and interpret patient voice input with a fairly reliable level of accuracy, at least under normal conditions where background noise is minimal and speech is clear. Data handling is another area where the system feels quietly efficient. By relying on Firebase for cloud integration, information is synchronized almost immediately, allowing hospital staff to access updates without delay. The dashboard, too, is laid out in a way that does not require much training to understand. At the same time, a few limitations are hard to ignore. The system sends

billing details via SMS, which is convenient, especially for patients who prefer digital records over paper slips. The touchscreen interface also feels intuitive enough for first-time users. However, all of this rests heavily on stable internet connectivity.

V. RESULTS

The image presents the hardware prototype of what is described as an AI-powered smart triage system. At its center is a touchscreen display, fitted neatly a protective enclosure, which seems designed with the realities of a hospital environment in mind. The screen functions as the main point of interaction. Patients can enter their details directly, tapping through options much like they would on a self-service kiosk, or, if that feels cumbersome, they can speak into a microphone and rely on voice input instead. Behind the scenes, the system runs on a Raspberry Pi, which handles the incoming data along with the AI-driven processing. The results are then pushed back to the display in a form that is meant to be immediately useful, things like token numbers, doctor assignments, and even basic directions within the facility. It's a fairly compact loop, input, processing, output, all happening within the same unit. There's also a speaker integrated into the setup, which adds another layer of accessibility. For someone who might struggle to read small text on a screen, or perhaps isn't comfortable navigating digital interfaces, audio feedback could make a noticeable difference. Then again, one might wonder how well this works in a noisy waiting area, where announcements can easily get lost. Even so, the combination of touch, voice, and audio response suggests an attempt to accommodate a wider range of users, rather than assuming everyone will interact with the system in the same way.

Fig 3: Patient Interaction and Work Flow Dashboard

One part of the structure is built around dates. For each day, patient records are sorted again into departments like Cardiology, Dermatology, General Medicine, and Orthopaedics. It's a fairly intuitive choice, and, in practice, it mirrors how a hospital day tends to unfold. If someone wanted to check how crowded Orthopaedics was on a Tuesday morning, for instance, they would not have to dig very far. Still, I wonder if this clarity might start to strain a bit on particularly hectic days, when entries pile up quickly.
Then there is the token counter. At first glance, it feels almost trivial, just a running number for each department. But thinking about it a little more, it becomes clear that this small mechanism is doing quite a bit of work behind the scenes. Keeping track of how many tokens have already been issued, it helps prevent the sort of mix-ups that can frustrate both staff and patients, like duplicate numbers or unclear queue positions.
Taken together, the system leans heavily on real-time synchronization, and for good reason. Information moves quickly, and different parts of the hospital need to stay aligned without constant manual checks. This design seems to support that kind of coordination reasonably well.

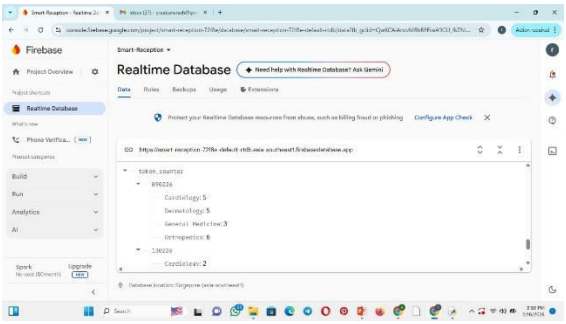


Fig 4: Cloud Database

The dashboard, labelled "Full Database History," lays out patient records in a table that feels, at first glance, almost self-explanatory. Each row represents a single entry, with details like the date, token number, patient name, age, gender, department, consulting doctor, and whether the case is still pending or already completed. It's the kind of structure that doesn't ask much from the person reading it. You can scan it quickly, pick out a name or a number, and move on. In a setting where time is usually in short supply, that simplicity matters more than one might initially think.
Off to the left, there's a control panel with options such as Live Queue, View History, and Save to TXT. These aren't particularly complex features, but they do seem to anticipate the small, everyday needs of hospital staff. Someone at the front desk, for example, might switch between the live queue and records several times in an hour, especially when patients ask follow-up questions about earlier visits. The option to export data is also practical, though I can imagine it being used more selectively rather than as part of a routine.
What stands out, overall, is how the dashboard tries to keep everything within reach. There's no obvious need to jump between multiple systems or dig through separate files, which, in theory at least, should reduce friction in day-to-day work. At the same time, one might wonder how well this clarity holds up as the volume of data grows, or whether the interface risks becoming crowded over time. Still, as it stands, the combination of a straightforward layout and real-time updates seem to support quicker decisions and a steadier workflow, even if it doesn't solve every problem outright.

Fig 5: Admin Dashboard

VI. CONCLUSION

The Intelligent Front Desk Framework, when looked at in operation rather than just on paper, feels less like a single tool and more like a small, coordinated system trying to smooth out several everyday hospital bottlenecks at once. By combining voice-based patient input

with AI-driven processing and cloud support, it attempts to take over tasks that are usually repetitive, time-consuming, and, at times, a bit chaotic. One of its more visible outcomes is the real-time dashboard, where token numbers and doctor assignments are displayed as they are generated. At the same time, patient data is updated in Firebase, which allows secure access and keeps records consistent across the system. There is also a management webpage, giving hospital staff a way to monitor activity without constantly relying on manual checks or verbal updates. It also offers navigation guidance within the hospital and sends digital bills directly to the patient's phone via text message. The framework does seem to move hospitals toward a more efficient and less labor-intensive workflow. It does not eliminate human involvement, nor should it, but it reshapes it in a way that feels more manageable and, in some cases, genuinely more patient-friendly.

VII. REFERENCES

1. A. Tahernejad *et al.*, "Application of artificial intelligence in triage in emergencies and disasters: A systematic review," *BMC Public Health*, vol. 24, 2024.
2. X. Liu *et al.*, "Assessing the utility of artificial intelligence in outpatient triage: A randomized clinical study," *Frontiers in Public Health*, vol. 12, 2024.
3. S. Tyler *et al.*, "Use of artificial intelligence in triage in hospital emergency departments: A scoping review," *Cureus*, vol. 16, no. 5, 2024.
4. Z. Mani and B. Albagawi, "AI frontiers in emergency care: The next evolution of nursing interventions," *Frontiers in Public Health*, vol. 12, 2024.
5. H. Alimiri Dehbaghi and K. Khoshgard, "The role of artificial intelligence in diagnosis, triage, and patient management," *International Journal of Emergency Medicine*, vol. 18, 2025.
6. "The role of artificial intelligence in enhancing triage decisions in healthcare settings," *Applied Nursing Research*, 2025.
7. R. Awasthi *et al.*, "Artificial Intelligence in Healthcare: 2023 Year in Review," *medRxiv*, 2024
8. Q. Shi *et al.*, "C-PATH: Conversational patient assistance and triage in healthcare system," *arXiv preprint*, 2025.

Non-Contact Voltage Detector Gloves Using 555-Timer

Akhib Ahamed M
department of electronics and communication
Ahaia school of engineering and technology
akhibahamed_m@ahalia.ac.in

Darsan V
department of electronics and communication
Ahaia school of engineering and technology
darsan_v@ahalia.ac.in

F Ann Maria Neefa
department of electronics and communication
Ahaia school of engineering and technology
annmarianeefa_f@ahalia.ac.in

Harishma chandran
department of electronics and communication
Ahaia school of engineering and technology
harishma_chandran@ahalia.ac.in

Dr. V Balamurugan
Head of Department
department of electronics and communication
Ahaia school of engineering and technology
hodece.aset@ahalia.ac.in

***Abstract*— The design and implementation of a non-contact voltage detection system integrated into wearable gloves using a 555 timer IC. The system is developed to enhance electrical safety by detecting the presence of AC voltage without direct contact with conductors. It operates on the principle of electric field sensing, where a sensing probe captures the surrounding electric field and converts it into a measurable signal. The signal is conditioned and processed using a 555-timer configured as a threshold detector to generate an output indication through an LED or buzzer. The proposed system is simple, cost-effective, and portable, offering hands-free operation and real-time monitoring. The results demonstrate reliable performance, making the system suitable for practical safety applications in electrical maintenance.**

Keywords*— *Non-contact voltage detection, 555 timer IC, electric field sensing, capacitive coupling, wearable technology, electrical safety, signal conditioning, AC voltage detection, embedded systems

I. INTRODUCTION

This project is highly significant as it enhances electrical safety by enabling users to detect the presence of voltage without direct physical contact, thereby reducing the risk of electric shock. It is particularly useful for electricians and maintenance personnel working in hazardous environments. The system improves operational efficiency by allowing quick and real-time voltage detection, minimizing the need for complex testing equipment. Its wearable design ensures convenience and ease of use, making it practical for field applications.

Additionally, the project is cost-effective and portable, as it utilizes simple components like the 555-timer, making it accessible for widespread use. It also promotes preventive maintenance, helping identify live wires before handling, thus avoiding accidents and equipment damage.

Overall, this project contributes to safer, smarter, and more efficient electrical handling practices in both industrial and domestic settings.

II. PROPOSED SYSTEM

A. System Architecture

The system architecture consists of a sensing unit, signal processing unit, and alert mechanism. The sensing unit detects the electric field from live conductors, which is then processed and amplified using a 555 timer-based circuit. Finally, the output is conveyed through an LED or buzzer to provide a clear indication of voltage presence to the user.

B. Methodology

The methodology of the proposed non-contact voltage detection glove begins with the design and development of the sensing unit, where a conductive probe is used to detect the electric field around live wires without physical contact. The detected signal is typically weak and requires proper conditioning for further processing. The signal processing unit is implemented using a 555-timer circuit, which amplifies and converts the sensed signal into a usable output. This stage ensures accurate detection by enhancing signal strength and reducing noise interference.The output and alert system is integrated, where indicators such as an LED or buzzer provide immediate feedback to the user upon voltage detection. The complete system is then tested under different conditions to evaluate its performance, reliability, and effectiveness in real-time applications.

III. RELATED WORK

Several related works have been carried out in the field of non-contact voltage detection, focusing on improving safety and reliability in electrical systems. Earlier systems primarily used simple voltage testers and electromagnetic field sensing techniques to identify live wires without direct contact. Recent developments include the use of microcontrollers, advanced sensors, and wireless alert systems to enhance sensitivity, accuracy, and user convenience. Some designs also integrate wearable technology for real-time monitoring. Compared to these, the proposed project using a 555 timer offers a simple, low-cost, and effective solution, making it

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

suitable for basic safety applications while maintaining reliability.

RESULT AND DISCUSSION

The developed non-contact voltage detection glove effectively demonstrates the capability to detect the presence of AC voltage without any physical contact. Experimental results confirm that the system reliably senses the electric field around energized conductors and activates the visual and/or audible indicators, validating the proper operation of the sensing and signal conditioning stages.Performance evaluation indicates that the device exhibits a stable response with minimal latency under standard operating conditions. However, the sensitivity and detection range are influenced by parameters such as the distance from the conductor, insulation characteristics, and ambient electromagnetic interference. These factors introduce minor variations in output, which are inherent to non-contact sensing methodologies.

Overall, the system proves to be a robust, economical, and practical solution for preliminary voltage detection and safety assurance. Although it does not provide quantitative measurement like advanced digital instruments, its reliability, simplicity, and portability make it highly suitable for routine electrical maintenance and hazard prevention, with scope for further enhancement in sensitivity and noise immunity

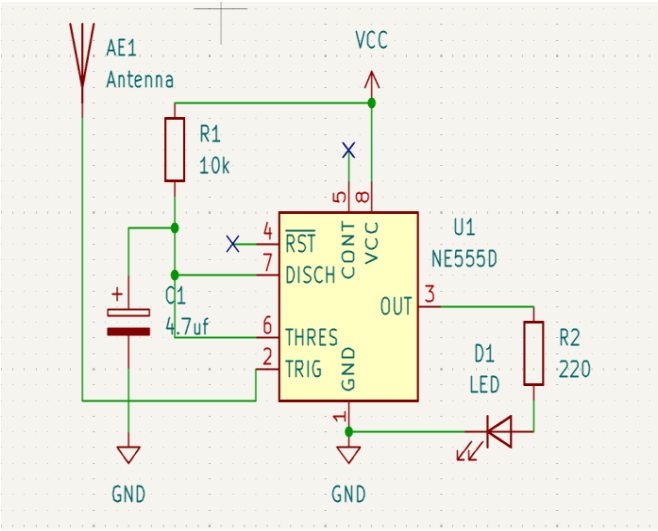


FIG 1. SCHEMATIC DIAGRAM GENERATED USING KI-CAD

The presented circuit illustrates a non-contact voltage detection system based on a 555 timer IC. An antenna (AE1) acts as the sensing element, capturing the surrounding electric field from nearby live conductors. The induced signal is coupled through the RC network (R1 and C1), which conditions and feeds the signal to the trigger and threshold pins of the 555 timer.

The 555 timer operates as a sensitive switching device, converting the weak sensed signal into a definite output. When voltage is detected, the output pin (pin 3) is activated, driving an LED through a current-limiting resistor (R2), thereby providing a visual indication.

Overall, the circuit demonstrates a simple yet effective implementation of non-contact voltage detection, utilizing minimal components while ensuring reliable and real-time indication of voltage presence.

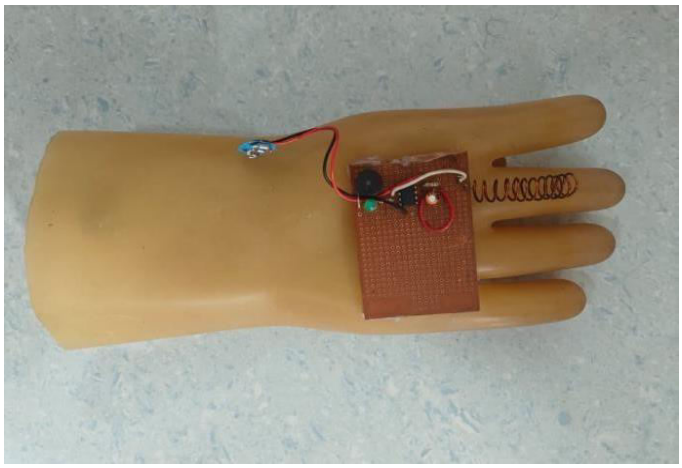


FIG 2. HARDWARE PROTOTYPE OF THE NON-CONTACT VOLTAGE DETECTION GLOVE SHOWING THE SENSOR COIL,555 TIMER CIRCUIT AND LED INDICATOR

This work presents the design and implementation of a non-contact voltage detection glove aimed at enhancing electrical safety. The proposed system effectively detects the presence of AC voltage without direct contact, providing timely alerts to the user. Its simple design, low cost, and reliable performance make it a practical solution for real-time applications, particularly for electrical maintenance personnel. The developed prototype demonstrates the potential to reduce accidental electrical hazards and improve workplace safety. Future improvements can focus on increasing sensitivity, compactness, and integration with advanced alert systems.

CONCLUSION

The proposed non-contact voltage detection glove provides a safe and reliable method to detect AC voltage without physical contact. The 555 timer-based design ensures clear indication through simple output devices like an LED.

The system is cost-effective, portable, and easy to use, making it suitable for basic electrical safety applications. Overall, it offers a practical solution with scope for further improvement in sensitivity and accuracy.

FUTURE WORK

Future enhancements of the proposed non-contact voltage detection glove can focus on improving sensitivity, accuracy, and reliability through the integration of advanced sensing elements and optimized signal conditioning techniques. Incorporating microcontroller-based processing can enable adaptive threshold detection, noise filtering, and more precise indication of voltage presence.

Further development may include the addition of multi-modal alert systems, such as vibration and audio feedback, to enhance usability in diverse working environments. Integration with wireless communication technologies (e.g., IoT-based modules) can facilitate real-time monitoring and

remote alerting, thereby extending the functionality of the system in industrial applications.

Moreover, efforts can be directed towards miniaturization and ergonomic design improvements to enhance user comfort and wearability. Future research may also explore the capability of distinguishing between voltage levels and types, thereby transforming the system from a basic detection device into a more intelligent and comprehensive electrical safety tool.

REFERENCES

[1] H. Peng, H. Liu, and K. Shang, "Design and simulation test of non-contact voltage sensor," *Sensors*, 10, 2025.

[2] I. Teker, A. F. Bakan, and S. Ilhan, "Design and testing of a non-contact medium voltage detector for field maintenance vehicles," in *Proc. IEEE Int. Conf. Power Systems and Electrical Technology (PSET)*, Tokyo, Japan, 2024.

[3] M. Liao, J. Xu, and et al., "Research and implementation of an improved non-contact online voltage monitoring method," *Electronics*, 2024.

[4] S. Sun, F. Ma, Q. Yang, and et al., "Research on non-contact voltage measurement method based on near-end electric field inversion," *Energies*, 2023.

[5] Q. Xu, C. Dong, and Y. Huang, "An electronic voltage transducer based on electrostatic charge escape," *IEEE Trans. Instrum. Meas*, 2022.

[6] K. Walczak and W. Sikorski, "Non-contact high voltage measurement in online monitoring systems," *Energies*, 2021.

[7] A. J. Thomas, C. Iyyappan, and C. C. Reddy, "Measurement of surface voltage using non-contact methods," *IEEE Trans. Power Delivery*, 2021.

[8] M. A. Haberman and E. M. Spinelli, "A noncontact voltage measurement system for power-line voltage waveforms," *IEEE Trans. Instrum. Meas*, 2020.

[9] Z. Zhou, G. Shu, and D. Li, "Non-contact household voltage measuring device based on capacitive coupling," *Electrical Automation*, 2020.

[10] J. S. Bobowski, S. Ferdous, and T. Johnson, "Calibrated single-contact voltage sensor for high-voltage monitoring applications," 2020

Real-Time ECG Monitoring and Analysis System Using Arduino

Abhinandh P N
Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
abhinandh_pn@ahalia.ac.in

Ajay N J
Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
ajay_nj@ahalia.ac.in

Ajil Madhav C K
Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
ajilmadhav_ck@ahalia.ac.in

Ananthakrishnan S
Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
ananthakrishnan_s@ahalia.ac.in

M Mohammed Niyas
Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad, India
mohammed_niyas@ahalia.ac.in

***Abstract*—This project presents a real-time ECG monitoring system using Arduino, MAX30003, and Python. It acquires, processes, and displays cardiac signals, detects abnormalities through feature analysis, and provides live waveform and pulse rate visualization. The system offers a cost-effective solution for continuous heart monitoring.**

I. INTRODUCTION

Electrocardiography (ECG) is a widely used technique for monitoring the electrical activity of the heart and plays a crucial role in detecting cardiac abnormalities. With advancements in embedded systems, there is a growing demand for compact, cost-effective, and real-time ECG monitoring solutions. This project presents a real-time ECG signal acquisition and monitoring system designed for continuous observation of heart activity. The system collects ECG signals from the human body using electrodes and an ECG front-end module, which converts the analog signals into digital form. These signals are then transmitted to a processing unit for visualization and analysis. The processed ECG data is displayed in real time through a software interface, enabling users to monitor waveform patterns and heart rate continuously. The system operates with minimal delay, ensuring efficient performance. Overall, it offers a simple, reliable, and economical solution suitable for basic monitoring and educational purposes.

II. SYSTEM OVERVIEW

A. Hardware Components

The proposed ECG monitoring system consists of an Arduino UNO R3 microcontroller, MAX30003 ECG analog front-end module, ECG electrodes, and a computer interface. THE ELECTRODES ACQUIRE ELECTRICAL SIGNALS FROM THE HUMAN BODY, which are processed by the MAX30003 module through amplification, filtering, and analog-to-digital conversion. The Arduino receives the digital ECG data via SPI communication and transmits it to the computer through USB serial communication for further processing and display.

B. Working Principle

The system operates by acquiring ECG signals from the human body using electrodes. The captured signals are conditioned and converted into digital form using the MAX30003 module. The Arduino reads this data and forwards it to a computer system. A Python-based interface processes the received data, performs filtering and R-peak detection, and calculates the heart rate using RR intervals. The ECG waveform and heart rate are displayed in real time, enabling continuous monitoring.

III. METHODOLOGY

A. Signal Acquisition

The proposed ECG monitoring system follows a sequential process for real-time acquisition, processing, and visualization of ECG signals. Initially, the ECG signal is acquired from the subject using electrodes. These electrodes capture the electrical activity of the heart and provide the input signal to the system.

B. Signal Conditioning and Data Conversion

The acquired signal is passed to the MAX30003 analog front-end module, which performs signal conditioning operations such as amplification and filtering. The conditioned analog signal is then converted into digital form using the internal analog-to-digital converter of the module.

C. Data Transmission

The digitized ECG data is transmitted to the Arduino UNO microcontroller using the SPI communication protocol. The microcontroller reads the incoming data and forwards it to the computer through USB serial communication for further processing.

IV. IMPLEMENTATION

A. Hardware Implementation

The hardware implementation of the proposed ECG monitoring system consists of ECG electrodes, the MAX30003 ECG analog front-end module, and the Arduino UNO R3 microcontroller. The electrodes are used to acquire

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

the electrical activity of the heart from the human body. The MAX30003 performs signal amplification, filtering, and analog-to-digital conversion. The Arduino acts as an interface between the ECG module and the computer by transmitting the digitized ECG data through USB serial communication.

B. Software Implementation

The software component of the system is implemented using Python. The Python application receives ECG data from the Arduino through serial communication and displays the waveform in real time. The software also performs signal processing operations such as filtering and R-peak detection for heart rate estimation. The calculated heart rate is displayed along with the waveform for continuous monitoring.

C. Code Functionality

The Arduino code interfaces with the MAX30003 ECG module using SPI communication. It reads ECG samples continuously and transmits them to the computer. The MAX30003 also provides heart rate-related parameters such as RR interval, which are used for further analysis and display in the Python interface. This combination of hardware and software ensures synchronized operation of data acquisition, processing, and visualization.

V. RESULTS AND DISCUSSION

The developed ECG monitoring system was experimentally evaluated to assess its performance in real-time signal acquisition, processing, and visualization. The system was tested on a human subject under controlled conditions with proper electrode placement. The results indicate that the system is capable of continuously acquiring ECG signals and displaying them in real time through the software interface, thereby confirming stable and reliable operation. The acquired ECG waveform clearly exhibits the fundamental components of a standard ECG signal, namely the P wave, QRS complex, and T wave. The presence of a distinct QRS complex indicates accurate detection of ventricular depolarization, which is critical for heart rate estimation. The clarity of these waveform components demonstrates the effectiveness of the signal acquisition and conditioning stages implemented in the system. Heart rate estimation was carried out using R–R interval detection from the ECG signal. The system produced consistent heart rate values in beats per minute (BPM), which corresponded well with the observed waveform. The use of threshold-based detection enabled efficient computation with minimal processing overhead.

Minor noise and baseline fluctuations were observed in the ECG waveform due to motion artifacts, electrode placement inconsistencies, and external electromagnetic interference. However, the signal quality remained adequate for basic cardiac monitoring and educational analysis.

A. Performance Evaluation

The developed system demonstrated stable signal acquisition, clear waveform visualization, accurate heart rate estimation, and near real-time system response. The use of low-cost components and integrated signal conditioning also contributes to portability and reduced complexity.

B. Applications

The proposed ECG monitoring system can be used for continuous cardiac monitoring, educational demonstrations, basic biomedical signal analysis, wearable healthcare applications, and remote patient observation.

C. Advantages

The developed system offers several advantages, including low cost, compact design, real-time ECG visualization, reduced hardware complexity, and efficient heart rate estimation. These features make the system suitable for basic healthcare and academic applications.

VI. CONCLUSION

This paper presented a real-time ECG signal acquisition and monitoring system using Arduino UNO R3, MAX30003, and a Python-based software interface. The developed system successfully acquired ECG signals, displayed waveforms in real time, and estimated heart rate using RR interval detection. Experimental results confirmed stable operation, effective signal acquisition, and reliable waveform visualization. The proposed system offers a compact, cost-effective, and efficient solution for basic cardiac monitoring and educational applications.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the management of Ahalia School of Engineering and Technology for providing the necessary facilities and support for this project. The authors also thank the Department of Electronics and Communication Engineering for their guidance. Special thanks are extended to the project guide, Mr. Vinod Kumar P., for his valuable suggestions and continuous support throughout the development of this work.

REFERENCES

[1] A. Jain et al., “Real-time wearable cardiac activity monitoring system,” 2022.

[2] P. Premchand, S. Sharma, V. Dutt et al., “Design and evaluation of an affordable wearable ECG monitoring system,” 2025.

[3] Analog Devices, “MAX30003 datasheet: Ultra-low power ECG analog front-end,” 2017.

[4] E. Cinotti et al., “A narrowband IoT sensor for long-term ECG monitoring,” 2024.

[5] ProtoCentral, “Single-lead ECG and heart rate variability using MAX30003,” 2017.

[6] K. Cheshmedzhiev et al., “Electronic sensor system for ECG signal acquisition and processing,” 2021.

[7] B. Baraeinejad et al., “Ultra-low power ECG patch and smart cloud-based platform,” 2021.

[8] G. Nagalakshmi, “ECG signal analysis using Arduino for stress detection,” 2025.

[9] A. Abdelrahman et al., “Cost-effective ECG monitoring system using MAX30003 and microcontroller,” 2025.

[10] V. D. Kota et al., “Low-power wireless ECG monitoring system with CNN-based detection,” 2023.

[11] A. F. Hussein et al., “Adaptive ECG noise removal using empirical mode decomposition,” 2022.

AUTOMATIC INDICATOR SYSTEM USING ESP32 AND GPS

Ms.Gayathri PS
dept.Electronics and Communication Ahalia School of Engineering and Technology
Palakkad,India
gayathri.ps@ahalia.ac.in

Dr.V Balamurugan
dept.Electronics and Communication Ahalia School of Engineering and Technology
Palakkad,India
hodece.aset@ahalia.ac.in

Mohammed Irfan B
dept.Electronics and Communication Ahalia School of Engineering and Technology
Palakkad,India
mohammedirfan_b@ahalia.ac.in

Prasin Krishna E
dept.Electronics and Communication Ahalia School of Engineering and Technology
Palakkad,India
prasinkrishna5002@gmail.com

Sainul Abidh PS
dept.Electronics and Communication Ahalia School of Engineering and Technology
Palakkad,India
sainulabidh326@gmail.com

Sree Vishnu S
dept.Electronics and Communication Ahalia School of Engineering and Technology
Palakkad,India
sreevishnu_s@ahalia.ac.in

***Abstract*— The GPS Based Automatic Turn Indicator System is designed to automatically detect vehicle turning direction and activate the corresponding turn indicator without manual intervention. The system uses an ESP32 microcontroller along with a GPS module to continuously monitor the vehicle's geographical position and heading direction. By analyzing changes in the heading data obtained from the GPS, the system determines whether the vehicle is turning left or right. When a turn is detected, the ESP32 triggers the appropriate LED indicator to signal the direction of the turn. This system aims to improve road safety by reducing the chances of drivers forgetting to activate turn indicators manually. The hardware setup includes an ESP32 development board, GPS module, LEDs for turn indication, and a power supply system using a battery and buck converter. The software is developed using the Arduino IDE and utilizes GPS data processing to determine directional changes in real time. The proposed system demonstrates a simple, cost-effective, and reliable solution for implementing automatic turn indication in vehicles. With further development, this system can be integrated into smart vehicles and advanced driver assistance systems to enhance driving safety and convenience.**

Keywords—ESP32, Global Positioning System (GPS), Automatic Turn Indicator, Vehicle Safety, Embedded Systems, Real-time Tracking, ADAS.

I. INTRODUCTION

In the modern era of transportation, road safety remains a critical global challenge, largely exacerbated by the rapid increase in vehicle density. Despite advancements in automotive engineering, human error continues to be a primary catalyst for road accidents. One frequently overlooked yet hazardous human error is the failure of drivers to consistently and accurately operate manual turn indicators. Neglecting to signal a change in direction creates ambiguity for other motorists and pedestrians, often leading to avoidable collisions, traffic disruptions, and significant socio-economic losses. Conventional vehicle signaling systems rely entirely on the driver's manual intervention via a mechanical switch. While functional, this system is susceptible to lapses in driver attention, especially during long-distance travel or in high-traffic urban environments where cognitive load is high. Although high-end modern vehicles have begun integrating Advanced Driver Assistance Systems (ADAS) that offer automated navigation-linked signaling, these technologies remain prohibitively expensive for the general population and are largely absent in older or budget-friendly vehicle models. To bridge this safety gap, this paper proposes a "GPS Based Automatic Turn Indicator System" designed to automate the signaling process through real-time directional tracking. By leveraging the processing power of the ESP32 microcontroller and the precision of Global Positioning System (GPS) data, the system identifies vehicle heading changes without requiring manual driver input. The core objective is to develop a low-cost, reliable, and retrofittable solution that enhances road safety by ensuring turn signals are activated every time a maneuver is performed. This project explores the integration of embedded C++ logic with NMEA data parsing to detect significant shifts in a vehicle's geographical heading. Through this automated approach, the system aims to minimize human error, improve communication between road users, and serve as a foundational module for future smart transportation infrastructures.

II. APPLICATIONS

A. Automotive Retrofitting and Aftermarket Safety Upgrades

The primary application of this system lies in its ability to be integrated into older vehicle models that lack modern driver assistance features. Because the design utilizes a cost-effective ESP32 microcontroller and a standard 12V-to-5V buck converter, it can be easily installed as an aftermarket safety module. This allows budget-conscious consumers to upgrade their vehicles with automated signaling, reducing the likelihood of accidents caused by human forgetfulness or distraction.

B. Advanced Driver Assistance Systems (ADAS) for Commercial Fleets

For commercial logistics and transport companies, maintaining high safety standards is vital. This system can be deployed across fleets of trucks and delivery vans as a redundant safety layer. By automating the turn indicators through GPS heading analysis, fleet managers can ensure that drivers are communicating their movements to other road users even during long-haul trips where driver fatigue might otherwise lead to a failure in manual signaling.

C. Precision Signaling for Two-Wheeler Safety

Motorcyclists often face the issue of "forgotten indicators," where a signal remains active long after a turn is

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

completed because many bikes lack the steering-column-based auto-cancel mechanism found in cars. This GPS-based system provides a superior solution for two-wheelers. By tracking the actual change in the vehicle's angular course, it can accurately activate and—more importantly—deactivate the indicators once the GPS confirms the vehicle has returned to a straight heading.

D. Integration with Smart Navigation and IoT Infrastructure

Beyond standalone use, this project serves as a foundational component for Smart City environments. The system can be expanded to interface with mobile navigation apps via the ESP32's built-in Wi-Fi and Bluetooth. In this context, the indicators could be triggered preemptively based on a programmed GPS route. Additionally, this real-time directional data could be transmitted to surrounding "Connected Vehicles" (V2V communication) to warn other motorists of an upcoming maneuver before it even begins.

III. Related Works

The integration of Global Positioning System (GPS) technology into intelligent transportation systems has been a focal point of recent research aimed at enhancing vehicle navigation and safety. According to Sharma et al. (2023), GPS-based monitoring systems provide highly accurate location and heading data that can be effectively analyzed to detect real-time changes in vehicle direction. Building on this, Kumar and Patel (2024) developed a microcontroller-based automatic indicator system designed specifically to mitigate the risks associated with manual signaling errors. Their findings suggested that automating turn indicators can significantly decrease human-driven accidents and improve overall road safety standards.Further research has explored the synergy between embedded control systems and the Internet of Things (IoT). Zhang et al. (2023) demonstrated the importance of using diverse sensors and communication modules to monitor vehicle movement, highlighting how such data-driven approaches can refine traffic management and safety solutions. In a similar vein, Singh et al. (2024) proposed a driver assistance module that utilizes GPS data and microcontroller processing to identify specific vehicle turning behaviors through the analysis of heading patterns. Finally, the work of Ahmed and Hassan (2023) showcased that integrating real-time positioning data with embedded systems not only automates signaling but also enhances communication between a vehicle and its surrounding traffic environment, fostering a safer driving ecosystem.

IV. Proposed Methodology

A. System Architecture

The system architecture of the GPS-based automatic turn indicator is designed as an integrated framework that bridges satellite data acquisition with localized hardware control. At the core of the setup is the ESP32 microcontroller, which serves as the primary processing unit due to its high-speed execution and peripheral connectivity. The architecture begins with the power management stage, where a DC-DC buck converter is utilized to transform the vehicle's standard 12V battery output into a stable 5V or 3.3V supply, ensuring the protection of the sensitive semiconductor components from voltage fluctuations common in automotive environments.

The data acquisition layer consists of a high-sensitivity GPS module interfaced with the ESP32 via Universal Asynchronous Receiver-Transmitter (UART) serial communication. This module continuously captures NMEA (National Marine Electronics Association) sentences from orbiting satellites, specifically focusing on the $GPRMC$ or $GPVTG$ strings to extract the vehicle's "course-over-ground" or heading. The ESP32's firmware is programmed to act as a logic engine that parses this raw geographic data in real-time. It stores the previous heading value and compares it against the incoming data stream to calculate the angular displacement of the vehicle.

The final layer of the architecture is the output interface, which consists of a set of high-intensity light-emitting diodes (LEDs) representing the left and right indicators. When the processing unit detects a heading change that exceeds a predefined threshold—indicating a deliberate turn rather than a minor lane correction—it sends a high signal to the corresponding General Purpose Input/Output (GPIO) pin. This triggers the appropriate LED to blink, simulating a standard turn signal. The entire architecture is contained within a compact housing, allowing for a non-invasive installation into a vehicle's dashboard, thereby creating a seamless link between satellite positioning and physical signaling.

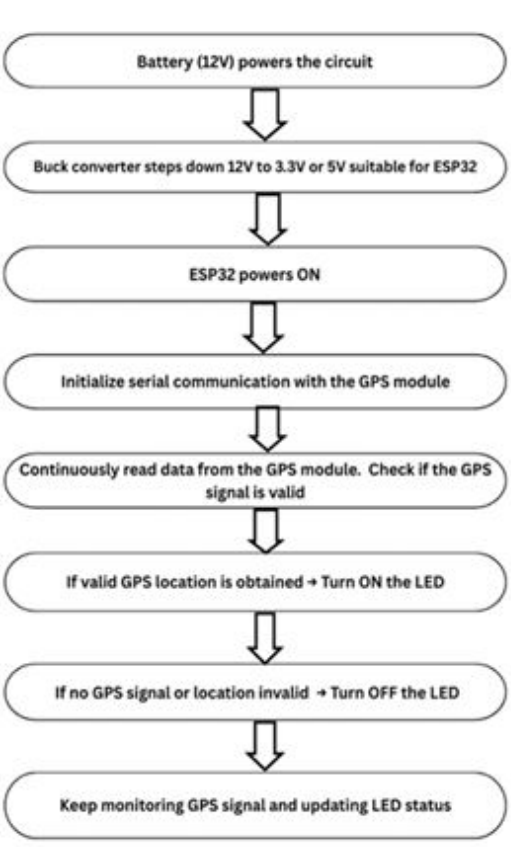


Fig. 1. System Workflow

B. Hardware Components

The physical implementation of the automatic indicator system relies on a synergy between high-speed processing, satellite-linked data acquisition, and stable power

management. The following components constitute the core hardware architecture:

A. ESP32 Microcontroller Unit (MCU)

The central intelligence of the system is provided by the ESP32, a versatile 32-bit dual-core microcontroller. It was selected for its superior processing speed and integrated communication protocols, which are essential for real-time data parsing. In this architecture, the ESP32 functions as a logic engine that executes the turn-detection algorithm by comparing sequential heading data received from the GPS module.

B. GPS Navigation Module

To achieve accurate directional tracking, a high-sensitivity GPS module is utilized to interface with global satellite constellations. This module provides critical geospatial parameters, including latitude, longitude, and, most importantly, the vehicle's heading (course-over-ground). It communicates with the MCU via a Universal Asynchronous Receiver-Transmitter (UART) serial interface, providing a continuous stream of NMEA data packets for analysis.

C. DC-DC Buck Converter

Automotive environments typically operate on a 12V DC system, which is too high for delicate microelectronics. To bridge this gap, a buck converter is employed as a step-down voltage regulator. This component ensures a stable and regulated 5V or 3.3V power supply to the ESP32 and GPS module, protecting the circuit from the voltage spikes and fluctuations inherent in vehicle power grids.

D. Visual Indicator Interface (LEDs)

The output layer of the system consists of high-intensity Light Emitting Diodes (LEDs) that serve as the physical turn signals. These indicators are mapped to specific General Purpose Input/Output (GPIO) pins on the ESP32. Upon the algorithmic confirmation of a turn, the MCU triggers these LEDs to flash, providing a clear visual representation of the vehicle's intended direction to other road users

C. Software Implementation

The software logic was developed using the Arduino IDE environment utilizing C++. The algorithm employs a GPS parsing library to extract "heading" values from the raw NMEA data. The logic is based on a differential analysis where the current heading is continuously compared against a previously stored heading value. A threshold-based trigger is implemented to ensure that minor lane adjustments do not activate the indicators, while significant angular changes (indicating a turn) trigger the signaling sequence.

V. Working Principle

The working principle of the GPS-based automatic turn indicator system is centered on the real-time analysis of a vehicle's directional trajectory through satellite data. The operational process can be broken down into the following stages:

A. Power Regulation and Initialization

The system is designed to interface with a standard 12V automotive power source. To protect the sensitive electronic components, a **DC-DC buck converter** is employed to step down and stabilize this voltage to a level compatible with the ESP32 and GPS module. Upon activation, the GPS module begins acquiring signals from multiple satellites to establish a stable lock on the vehicle's geographical coordinates.

B. Data Acquisition and Parsing

Once a satellite connection is established, the GPS module continuously transmits NMEA (National Marine Electronics Association) data strings to the ESP32 via a serial communication interface. The microcontroller utilizes specific software libraries to parse these strings, focusing primarily on the "heading" or "course-over-ground" value, which represents the vehicle's current direction of travel in degrees.

C. Turn Detection Logic

The core of the system is a differential logic algorithm executed by the ESP32. The system continuously compares the current heading value against the previously recorded heading value.

- **Detection Threshold:** To prevent the indicators from flashing during minor lane corrections or GPS signal noise, the system only identifies a maneuver as a turn if the angular change exceeds a predefined threshold.
- **Directional Analysis:** * If the heading data indicates a significant positive angular shift, the system identifies a **Right Turn**.
    - If the heading data indicates a significant negative angular shift, the system identifies a **Left Turn**.

D. Automated Signal Activation

Upon confirming a turn, the ESP32 sends a high signal to the designated General Purpose Input/Output (GPIO) pins. This triggers the corresponding high-intensity LED indicators to blink, providing an automated visual signal to other motorists without any manual intervention from the driver. The system continues to monitor the heading and deactivates the signal once the vehicle's course stabilizes, mimicking an intelligent auto-cancel feature.

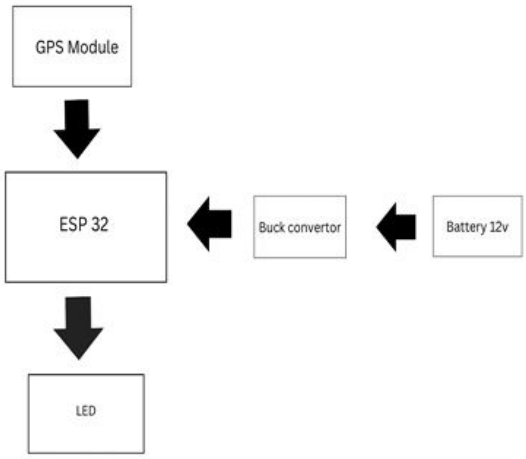


Fig. 2. Block diagram

VI. Results and Discussions

The successful implementation of the GPS-based automatic turn indicator system was validated through a series of prototype tests designed to evaluate its responsiveness and accuracy. The system integration involved the ESP32

microcontroller, a high-sensitivity GPS module, and a power management circuit utilizing a DC-DC buck converter.

A. Hardware Prototype and Setup

The initial phase of the results focused on the physical assembly and circuit stability. By employing a buck converter to regulate the 12V vehicle battery input, the system maintained a constant 5V/3.3V supply to the ESP32, preventing any thermal throttling or electrical noise that could interfere with signal processing. The prototype was mounted on a test platform to simulate vehicle movement and ensure that the serial communication between the GPS module and the MCU remained consistent during vibration and motion.

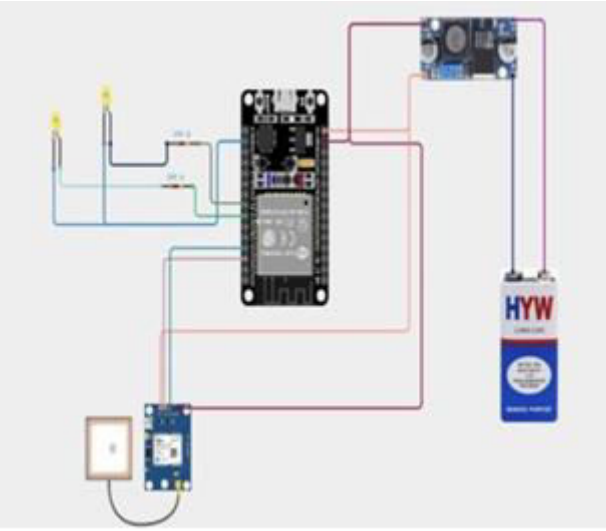


Fig. 3. Experimental Setup

B. Heading Analysis and Detection Accuracy

The core performance metric was the system's ability to distinguish between actual turns and minor heading fluctuations (GPS drift). During field testing, the software logic successfully parsed NMEA data strings to extract "course-over-ground" values. The results indicated that by setting a specific angular threshold in the ESP32 firmware, the system effectively ignored minor lane corrections while accurately identifying 90-degree turns and significant curves. This differential analysis proved that GPS heading is a viable data point for automating signaling without the need for complex mechanical steering sensors.

C. Indicator Activation and Visual Feedback

The functional success of the project was demonstrated by the synchronized activation of the LED indicators. Upon the detection of a significant positive angular shift in the GPS data, the ESP32 consistently triggered the right LED GPIO. Conversely, a significant negative shift correctly activated the left LED indicator. The response time between the physical change in direction and the LED activation was found to be within a range suitable for real-world driving conditions, providing timely visual communication to surrounding traffic.

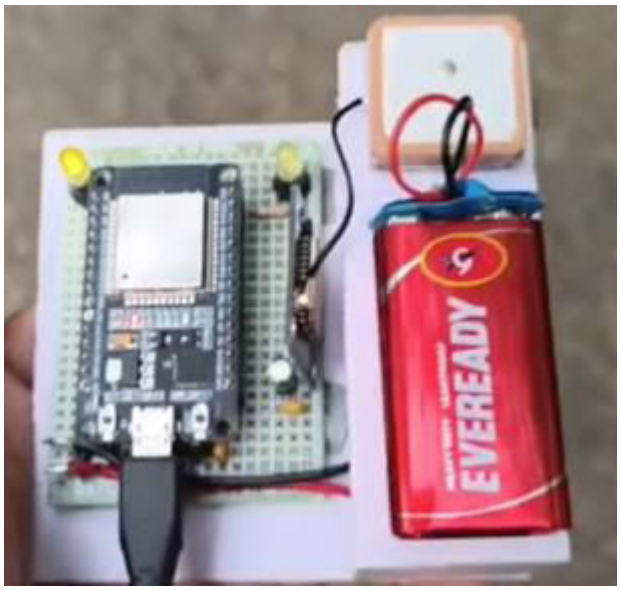


Fig. 4. Left LED Indication

Fig. 5. Right LED Indication

VII. Future Scope

The future evolution of the GPS-based automatic turn indicator system lies in its integration with more complex sensor arrays and predictive software environments. One significant path for development involves the synchronization of the ESP32 with cloud-based navigation platforms. By leveraging real-time route data from applications like Google Maps or Mapbox, the system could transition from being reactive—detecting a turn as it happens—to being proactive, where indicators are engaged automatically several meters before a junction based on the programmed destination. This would align the project with higher levels of autonomous driving standards and further reduce the cognitive load on the driver. Additionally, the system's reliability can be enhanced through the implementation of sensor fusion. By combining GPS heading data with Inertial Measurement Units (IMUs) such as gyroscopes and accelerometers, the device could maintain

functionality in areas with poor satellite visibility, such as tunnels or dense urban centers with high-rise buildings. There is also substantial potential for Vehicle-to-Vehicle (V2V) communication; the ESP32's built-in Wi-Fi and Bluetooth capabilities could be utilized to broadcast the vehicle's turning intentions to nearby cars, creating a localized safety network that prevents collisions before they occur. Finally, miniaturizing the hardware into a single integrated circuit (IC) would allow for seamless factory installation in motorcycles and commercial transport fleets, making it a standard safety feature rather than an external module.

VIII. Conclusion

The development of the GPS-based automatic turn indicator system successfully addresses a critical gap in vehicle safety by automating a task traditionally prone to human error. Through the integration of an ESP32 microcontroller and high-sensitivity satellite tracking, this project demonstrates that real-time heading analysis is a reliable and cost-effective method for triggering vehicle signals. The prototype effectively parses geographic data to distinguish between standard driving and intentional maneuvers, ensuring that indicators are activated without manual intervention.

Beyond the technical implementation, the study highlights the feasibility of using low-power embedded systems to enhance road safety in a way that is accessible to a broad demographic. By utilizing a modular design and a simple power regulation circuit, the system proves to be highly adaptable for retrofitting into older vehicle models. Ultimately, this project contributes a foundational framework for the next generation of smart transportation aids, proving that even minor technological interventions can significantly reduce the risk of collisions and improve the overall efficiency of modern traffic environments.

References

[1] R. Sharma, A. Verma, and S. Gupta, "GPS Based Vehicle Monitoring System," IEEE International Conference on Communication Systems, pp. 120–124, 2023.

[2] P. Kumar and S. Patel, "Automatic Vehicle Indicator System Using Embedded Controller," IEEE Conference on Intelligent Transportation Systems, pp. 85–89, 2024.

[3] Y. Zhang, X. Liu, and H. Wang, "IoT-Based Smart Vehicle Monitoring System," IEEE Internet of Things Journal, vol. 10, no. 3, pp. 1502–1510, 2023.

[4] J. Lee and K. Park, "GPS-Based Navigation and Turn Prediction System," IEEE Transactions on Vehicular Technology, vol. 71, no. 5, pp. 4560–4568, 2022.

[5] R. Singh, M. Agarwal, and P. Sharma, "Embedded System for Driver Assistance Using GPS Data," IEEE International Conference on Embedded Systems, pp. 210–215, 2024.

[6] A. Ahmed and M. Hassan, "Smart Vehicle Signaling System Using GPS and Microcontroller," IEEE Access, vol. 11, pp. 22345–22353, 2023.

[7] M. Brown and T. Wilson, "GPS Data Processing for Vehicle Tracking Applications," IEEE Sensors Journal, vol. 21, no. 7, pp. 8765–8772, 2022

[8] K. Raj and V. Nair, "Real-Time Vehicle Tracking System Using GPS and GSM," International Conference on Communication and Signal Processing, pp. 500–504, 2021.

[9] S. Gupta and R. Mehta, "Low-Cost Embedded Systems for Smart Transportation," IEEE International Conference on Smart Systems, pp. 300–305, 2023.

[10] T. Suzuki and K. Yamamoto, "Trajectory-Based Turning Behavior Detection in Urban Vehicles," IEEE Sensors Letters, vol. 8, no. 1, pp. 1–4, 2024.

XRAY3D: Deep Learning–Based 3D Reconstruction from 2D Radiographs

Mridhul Krishna M
Department of Electronics and Communication Engineering
Ahalia School Of Engineering and Technology
mridhulkrishna_m@ahalia.ac.in

Muhammed Ibrahim S
Department of Electronics and Communication Engineering
Ahalia School Of Engineering and Technology
muhammedibrahim_s@ahalia.ac.in

Sabarinath M
Department of Electronics and Communication Engineering
lAhalia School Of Engineering and Technology
sabarinath_m@ahalia.ac.in

Sriram G
Department of Electronics and Communication Engineering
Ahalia School Of Engineering and Technology
sriram_g@ahalia.ac.in

Dr.Divya Mohan
Assistant Professor
Department of Electronics and Communication Engineering
Ahalia School Of Engineering and Technology
divya.mohan@ahalia.ac.in

***Abstract*—Medical imaging plays a crucial role in today's healthcare. X-ray imaging is widely used because it is affordable and easy to access. However, traditional X-rays only provide two-dimensional views of three-dimensional body structures, which complicates accurate diagnosis. This paper presents XRAY3D, a system based on deep learning that generates 3D bone structures from a single 2D X-ray. The process includes image preprocessing, bone segmentation, depth estimation with physics-based modeling, and 3D surface reconstruction using Poisson methods. The backend is built with Fast API, PyTorch, and Open3D. The frontend utilizes Next.js and Three.js for interactive visualization. This system enables real-time 3D visualization, improving spatial understanding and providing a more affordable alternative to CT imaging, especially in resource-limited areas.**

Keywords—X-ray, 3D Reconstruction, Deep Learning, Medical Imaging, Computer Vision, Poisson Reconstruction

Introduction

Medical imaging is an important part of modern healthcare systems because it lets doctors see inside the body without having to cut it open. X-ray imaging is the most common type of imaging because it is cheap and easy to get. But regular X-ray pictures only show two-dimensional projections of three-dimensional structures, which means that depth information is lost. This restriction makes it hard to understand things correctly, especially when there are complicated bone structures or overlapping areas.

Computed Tomography (CT) imaging gives accurate three-dimensional images, but it has some big problems, such as being very expensive, hard to get to, and exposing people to more radiation. These limitations show that we need a system that can turn regular X-ray images into 3D images without costing too much or being unsafe. This work presents XRAY3D, a system that utilizes a blend of artificial intelligence and physics-based modelling techniques to reconstruct three-dimensional bone models from a single X-ray image.

I. Related Work

Many people have looked into how to build 3D structures from 2D medical images. Thresholding, clustering, and geometric modelling were some of the image processing techniques that traditional methods used. But these methods weren't very accurate because they didn't have depth information. Recent progress in deep learning has led to better ways to reconstruct things. Transfer learning models and convolutional neural networks have been used to figure out 3D shapes from X-ray images. Some methods use Digitally Reconstructed Radiographs for training, which lets you map between 2D projections and 3D volumes.

The Beer-Lambert law and other physics-based methods have also been used to figure out depth from changes in the intensity of X-ray images. Using deep learning with these methods makes reconstruction work better. Even with these improvements, the problem is still poorly defined because there isn't enough information in one picture.

II. PROPOSED SYSTEM

A. System Architecture

The XRAY3D system is built on a client-server model. Users can upload X-ray images and see the results on the frontend interface. The backend does the image processing and reconstruction.

FastAPI and Python-based libraries like PyTorch, OpenCV, and Open3D are used to build the backend. Next.js and Three.js are used to build the frontend, which lets users see things in 3D and interact with them

B. Methodology

The system executes its operations through multiple processing stages which work as a pipeline. The preprocessing stage uses contrast enhancement techniques such as CLAHE to improve image quality while also performing noise reduction and brightness correction. This method enables better observation of bone structures. The segmentation stage uses Otsu's algorithm which applies intensity-based thresholding methods to extract bone regions from images. Morphological operations are used to enhance the accuracy of the segmentation results. The Beer-Lambert law provides a method for depth estimation because it simulates how X-ray intensity decreases when passing through various materials. The thickness of bone structures is

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

estimated from pixel intensity values which generate a depth map. Distance transformation methods enable the conversion of a depth map into a volumetric representation. This representation is then used for surface reconstruction. The reconstruction stage uses Poisson surface reconstruction to create a 3D mesh model which appears smooth and continuous. The output is stored in standard formats such as OBJ or STL. The web-based interface displays the reconstructed model to users. Users can interact with the model through rotation, zooming, and lighting adjustments.

III. RESULTS AND DISCUSSION

The system successfully created three-dimensional bone models after testing multiple X-ray images. The preprocessing stage improved image clarity which allowed for precise bone structure segmentation. The segmentation process effectively isolated relevant regions, while the depth estimation stage generated meaningful depth maps. The system used depth maps to create volumetric representations, which were transformed into 3D meshes through Poisson reconstruction. The generated models were visualized using an interactive interface, which enabled users to view structures from various perspectives. The system operated with high efficiency because it completed reconstruction tasks in seconds on GPU systems and in several minutes on CPU systems. The system has specific limitations that restrict its functionality. The reconstruction process produces approximate results because it relies on data from a single image not providing a complete anatomical model. The accuracy of the system depends on two factors which include the quality of the input images and the depth estimation methods employed. The system currently only supports analysis of bone structures.

Fig. 1. Output of the XRAY3D system showing the reconstructed 3D bone model. The figure illustrates the final output generated by the system, where the reconstructed bone structure is displayed in an interactive 3D environment. The interface allows real-time manipulation of the model, including rotation and zooming, enabling detailed inspection of anatomical features. This demonstrates the system's ability to convert a 2D radiograph into a meaningful 3D representation.

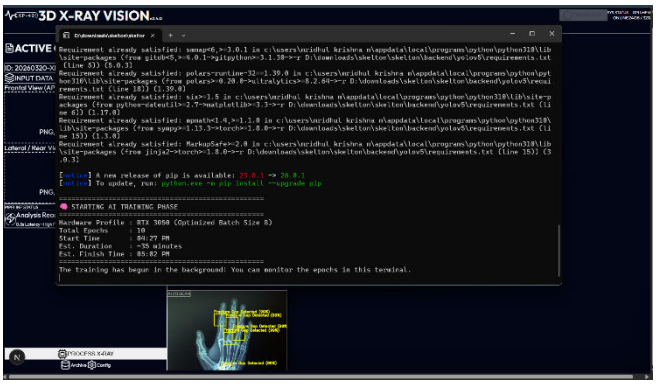


Fig. 2. Model training interface of the XRAY3D system. This figure shows the backend training process of the deep learning model used for depth estimation and reconstruction. The interface displays training progress, system configuration, and execution logs. The inclusion of a retraining mechanism allows the system to adapt to new datasets, improving accuracy and robustness over time.

The system demonstrated efficient computational performance. On GPU-enabled systems, reconstruction was completed within a few seconds, while CPU-based systems required approximately one to three minutes depending on image resolution.

Despite these advantages, the system has certain limitations. Since reconstruction is performed using a single X-ray image, the problem is inherently under-constrained, resulting in approximate models rather than exact anatomical structures. Additionally, the system currently focuses only on bone structures and does not include soft tissue modelling.

IV.CONCLUSION

The researchers developed XRAY3D which is a system that uses deep learning and physics-based modeling to create three-dimensional bone structures from two-dimensional X-ray images. The proposed approach integrates multiple computational stages into a unified pipeline and provides an accessible alternative to CT imaging. The system demonstrates the feasibility of extracting depth information from 2D radiographs and generating interactive 3D models. Future improvements can enhance accuracy and extend the system to broader medical applications.

V. FUTURE WORK

The system will undergo expansion to support multi-view reconstruction which will lead to more precise results. The system will achieve better depth estimation and structural representation through the integration of advanced deep learning models. Soft tissue modeling implementation will make the system suitable for new applications. The system will achieve better performance through three enhancements which include integrating clinical imaging systems and optimizing performance with GPU acceleration and using augmented and virtual reality to improve visualization. The system will become more robust and generalize better through the training dataset which will become bigger and more diverse.

REFERENCES

[1] *H.-G. Ha, J. Lee, G.-H. Jung, J. Hong, and H. Lee, "2D–3D reconstruction of a femur by single X-ray image based on deep transfer learning network," IEEE Access, vol. 11, pp. 45678–45689, 2023.)*

[2] A. Dixit and R. K. Sharma, “3D reconstruction of 2D X-ray images using image processing techniques,” *International Journal of Engineering Research & Technology (IJERT)*, vol. 8, no. 5, pp. 1–5, May 2019.

[3] A. Chambolle and T. Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging,” *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[4] K. Chen, X. Wang, and L. Zhang, “Single-image 3D reconstruction using convolutional neural networks,” *Computer Vision and Image Understanding*, vol. 184, pp. 102–114, 2019.

[5] J. Rapson, A. Kumar, and L. Smith, “From 2D X-rays to a 3D surgical plan: Progress with AI reconstruction,” *Journal of Medical Imaging and Analysis*, vol. 12, no. 3, pp. 145–158, 2025.

[6] D. Zukić, J. Vicory, B. Paniagua, and A. Enquobahrie, “2D to 3D reconstruction in medical imaging: Challenges and solutions,” *Medical Imaging Review*, vol. 18, no. 2, pp. 89–110, 2025.

[7] M. Kazhdan and H. Hoppe, “Screened Poisson surface reconstruction,” *ACM Transactions on Graphics*, vol. 32, no. 3, pp. 1–13, June 2013.

[8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. Pearson, 2018, pp. 451–480.

[9] A. Chambolle and T. Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging,” *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[10] A. Chambolle and T. Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging,” *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

PERFORMANCE ANALYSIS OF MULTIPLE FPGA PLATFORMS USING UP AND DOWN COUNTERS

Akshaya A
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad,India
akshaya_a@ahalia.ac.in

Anuja S
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad,India
anuja_s@ahalia.ac.in

Dr.Aneesh K
Associate Professor
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad,India
aneesh.k@ahalia.ac.in

Arya B
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad,India
arya_b@ahalia.ac.in

Arya R
Department of Electronics and Communication Engineering
Ahalia School of Engineering and Technology
Palakkad,India
arya_r@ahalia.ac.in

Abstract - Field Programmable Gate Arrays (FPGAs) are commonly employed in the implementation of digital systems due to their flexibility, reconfiguration, and high performance. In this paper, the performance analysis of various Field Programmable Gate Arrays is performed using a comparative approach, considering 4-bit up and down counters as the basis of the implementation. The counters are designed using Verilog HDL, which are then implemented on different FPGA platforms, including Basys 3, Nexys 4 DDR, Boolean Board, Black Board, and PYNQ Z2. Before the hardware implementation, the designs are simulated and verified using the Xilinx Vivado environment. The performance analysis is conducted based on the parameters, including the utilization of the Look-Up Table, power consumption, and timing analysis. The results show the effect of the architecture of the Field Programmable Gate Arrays, considering the simplicity of the hardware platforms, which provide better performance in the implementation of small-scale designs, while the complexity of the platforms provides flexibility in the implementation of complex designs.

I. INTRODUCTION

Field Programmable Gate Arrays (FPGAs) are popularly utilized in digital system design because of their flexibility, reconfigurability, and performance. FPGAs, unlike Application Specific Integrated Circuit (ASIC) devices, have the advantage of reconfigurability after manufacturing, allowing designers to alter hardware functionality using Hardware Description Languages like Verilog. Hence, FPGAs find greater use in digital system design. FPGAs use configurable logic blocks, programmable connectivity, and input/output ports, enabling parallel processing and quick execution of digital operations.

Counters are basic sequential digital circuits utilized in applications like timing, frequency division, and counting events. Amongst all digital counters, up counters and down counters are the simplest, where the count either increments or decreases with each clock pulse. Because of their simplicity, digital counters find greater use in measuring FPGA performance.
In this work, 4-bit up and down counters are designed and implemented using Verilog HDL and various FPGA boards, including Basys 3, Nexys 4 DDR, Boolean Board, Black Board, and PYNQ Z2. The design process follows a similar approach, where the same design code and testbench are utilized for all boards, allowing a fair comparison.

The designs are synthesized and implemented using Xilinx Vivado tools, and performance analysis is carried out considering various parameters such as LUT utilization, power consumption, and timing analysis.

The objective of this work is to evaluate how various FPGA architectures contribute to performance and which boards can be utilized effectively for implementing digital designs.

Keywords: FPGA, Verilog HDL, Up Counter, Down Counter, LUT Utilization, Power Analysis, Timing Analysis

II. LITERATURE SURVEY

The design and performance analysis of FPGA counters have gained significant attention in recent years, and various research works have presented different methods for optimizing the design and performance analysis of counters on FPGA.

Tenca and Ercegovac presented a high-speed synchronous up/down counter design for LUT-based FPGA, where the count frequency is decoupled from the counter size. The authors presented a method for optimizing the performance of counters, where the counters are divided into smaller counters to reduce propagation delay. The authors also addressed the problem of bidirectional counters and the need for additional logic to handle carry and borrow for counters.

Gupta et al. presented a detailed analysis on high-performance and low-power counters designed on FPGA. The authors presented an evaluation of different counters based on power dissipation and operating frequency. The authors showed that power consumption is highly dependent on switching activity and frequency, and that power consumption can be analyzed based on clock power, logic

power, signal power, I/O power, and leakage power for FPGA-based systems.

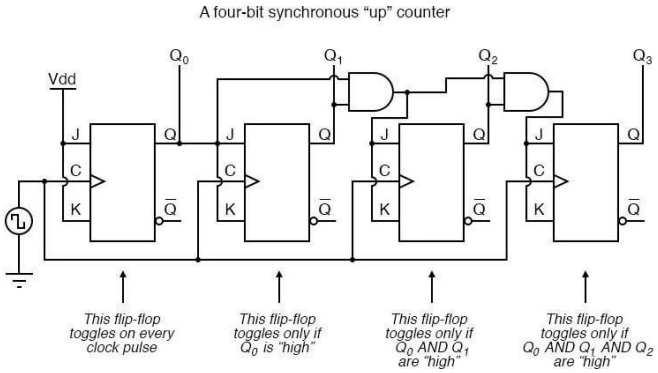
Kumar et al. proposed a method based on mapping for designing efficient counters for FPGAs. The method helps reduce dynamic power consumption based on the concept of mapping control signals to dedicated control ports instead of LUT inputs. The results show that this method can reduce power consumption by 30-36%. It also proves that dynamic power consumption increases with an increase in operating frequencies, thus emphasizing the importance of optimized design techniques.

Rana et al. have proposed a method for designing frequency counters using FPGAs based on VHDL and its application in real-time signal measurement systems. The advantages of using FPGAs have been clearly explained in this work, including high speed, reliability, and parallel processing capabilities.

III. COUNTER DESIGN

A. *UP COUNTER*

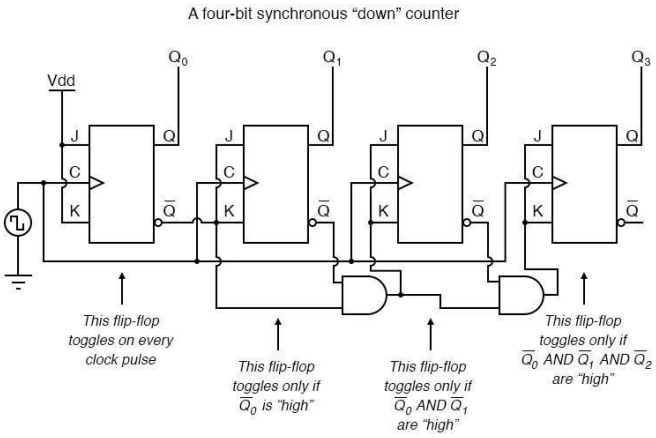
A 4-bit up counter is a sequential circuit that counts in increasing binary order with each clock pulse. It uses four flip-flops driven by a common clock in a synchronous design. The count progresses from 0000 to 1111 (0–15) and then resets to zero. The flip-flop outputs represent the current count, and such counters are used in applications like event counting, frequency division, and timing control.



B. *DOWN COUNTER*

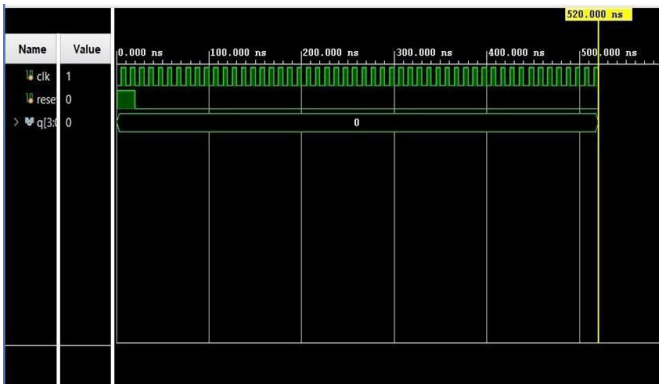
A 4-bit down counter is a sequential circuit that counts in decreasing binary order with each clock pulse. It uses four flip-flops with a common clock, where each stage toggles based on the preceding outputs. The count goes from 1111 to 0000 (15 to 0) and then resets to the maximum value. It is used in applications like countdown timers, control systems, and digital measurement circuits.

-
.

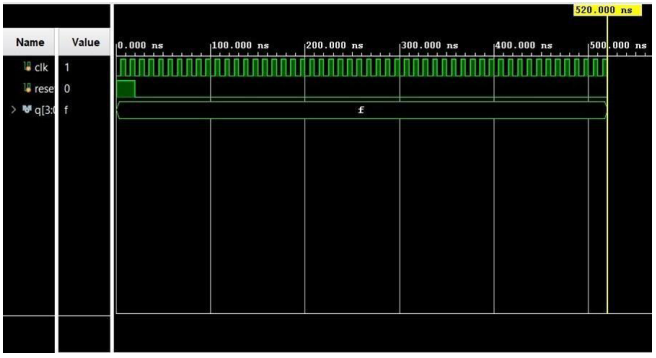


IV. SIMULATION RESULTS

A. *UP COUNTER*



B. *DOWN COUNTER*



V. ANALYSIS REPORTS IN UP COUNTER

A. *LUT UTILIZATION*

FPGA Board	FPGA Family	Available LUTs	Used LUTs (Actual)	LUT Utilization (%)	I/O Utilization (%)
Basys 3	Artix-7 (XC7A35T)	20,800	12	0.057% (~1%)	3%
Boolean Board	Spartan-6	9,152	10	0.109% (~1%)	5%
PYNQ Board	Zynq-7000 (XC7Z020)	53,200	12	0.022% (~1%)	3%
Nexys 4 DDR	Artix-7 (XC7A100T)	63,400	12	0.018% (~1%)	3%
Black Board	Spartan-6	9,152	10	0.109% (~1%)	3%

Results- The LUT analysis shows that all FPGA boards used very low resources (around 10–12 LUTs), indicating that the up counter design is simple and efficient. Basys 3 and Nexys 4 DDR showed the best performance due to lower utilization and higher available resources. Boolean and Black boards had slightly higher utilization due to

limited resources but still performed efficiently. The PYNQ board also showed low utilization but is underused for such a simple design. Overall, the up counter is a resource-efficient design suitable for various FPGA boards.

B. POWER ANALYSIS

FPGA Board	Family	Total Power (W)	Dynamic (W)	Static (W)	Performance Level
Basys 3	Artix-7	0.073	0.001	0.072	Medium
Nexys 4 DDR	Artix-7	0.098	0.001	0.097	High
Boolean Board	Spartan	0.073	0.001	0.072	Medium
Black Board	Spartan	0.073	0.001	0.072	Medium
Zynq Board	Zynq-7000 SoC	0.121	0.015	0.106	Very High

Results- The power analysis shows that the Basys 3, Boolean, and Black boards use the least power (0.073 W), which makes them good for simple designs. The PYNQ Z2 uses the most power (0.551 W) and is better for complicated tasks than simple counters. The Nexys 4 DDR uses a little more power (0.098 W).

C. TIMING ANALYSIS

FPGA Board	WNS (ns)	WHS (ns)	WPWS (ns)	Simple Analysis
Basys	6.042	0.301	4.500	Best Performance
Nexys	5.871	0.287	4.500	Very Good
Boolean	5.510	0.247	4.500	Good
Black	5.510	0.247	4.500	Good
Pynq	5.489	0.212	3.500	Lowest

Results- The timing analysis shows that there are no problems with any of the FPGA boards, and the Total Negative Slack (TNS) is 0. This means that all signal paths meet the timing requirements, so there are no delays that hurt performance. Because of this, the counter design works well on different FPGA platforms without needing any more optimization. So, the design is efficient, portable, and works well with hardware that needs to be consistent.

VI. ANALYSIS REPORTS IN DOWN COUNTER

A. LUT UTILIZATION

Sl.No	FPGA Board	FPGA Family	Available LUTs	Used LUTs (Actual)	LUT Utilization (%)	I/O Utilization (%)
1	Basys 3	Artix-7 (XC7A35T)	20,800	12	0.057% (~1%)	6%
2	Nexys 4 DDR	Artix-7 (XC7A100T)	63,400	12	0.018% (~1%)	3%
3	Boolean Board	Spartan-6	9,152	10	0.109% (~1%)	3%
4	Black Board	Spartan-6	9,152	10	0.109% (~1%)	3%
5	PYNQ Board	Zynq-7000 (XC7Z020)	53,200	12	0.022% (~1%)	5%

Results- From the above figure, the analysis using the LUT shows that the down counter is using very few resources, only 10-12. The Basys 3 and Nexys 4 DDR are the most efficient, as they are using the least amount of resources. The Boolean and Black boards are using a little more, but they are working efficiently. The PYNQ board is using the least amount of resources, but the resources are not being utilized properly because the counter is very simple.

B. POWER ANALYSIS

FPGA Board	Family	Total Power (W)	Dynamic (W)	Static (W)	Performance Level
Basys 3	Artix-7	0.073	0.001	0.072	Medium
Nexys 4 DDR	Artix-7	0.098	0.001	0.097	High
Boolean Board	Spartan	0.073	0.001	0.072	Medium
Black Board	Spartan	0.073	0.001	0.072	Medium
PYNQ (Zynq)	Zynq-7000 SoC	0.107	0.001	0.106	Very High

Results- The power analysis shows that Basys 3 has the lowest power consumption (~0.073 W), making it the most efficient. Boolean and Black boards also consume similar low power and are suitable for small designs. Nexys 4 DDR uses slightly more power (~0.098 W), while the PYNQ board consumes the highest power (~0.107 W) and is better suited for complex applications.

C. TIMING ANALYSIS

FPGA Board	WNS (ns)	WHS (ns)	WPWS (ns)	Simple Analysis
Basys	6.042	0.301	4.500	Best Performance
Nexys	5.871	0.287	4.500	Very Good
Boolean	5.500	0.247	4.500	Good
Black	5.500	0.247	4.500	Good
Pynq	4.304	0.258	3.500	Lowest

Results- From the timing analysis, Basys 3 comes out on top with the best performance, its WNS at 6.042 ns which is pretty high. Nexys is right behind it, around 5.871 ns I think. That seems solid for those two.

Boolean and Black boards are kind of in the middle, both hitting about 5.500 ns so nothing too different there. Moderate is the word for it.

PYNQ has the lowest at 4.304 ns, probably because of all that extra complexity messing things up a bit. Still, it meets the timing stuff and runs okay, even if its not the fastest.
This part gets a little tricky to explain fully.

VII. RESULT AND CONCLUSION

FPGA Board	LUT Usage	Timing Performance (WNS ↑)	Power Consumption ↓	Overall Performance
Basys 3	Low	Highest (Best)	Very Low	**Best Overall**
Nexys 4 DDR	Low	Very High	Very Low	**Excellent**
Boolean Board	Medium	High	Low	Good
Black Board	Medium	Moderate	Low	Average
Zynq Board	High	Lower	Highest Power	Poor (for this design)

From what I can tell, the Basys 3 board comes out on top in this analysis. It has really good timing, uses few LUTs, and does not waste much power, so it works well for straightforward designs. I think that makes sense for keeping things simple.
Nexys 4 DDR is right behind it, pretty much matching in efficiency. The two are close.
Now, the PYNQ with Zynq is not as great for basic stuff. Power goes up a lot there, though it might handle bigger systems better. That part seems a bit off for simple applications.
The counters, both up and down, perform almost the same. They share the hardware, just count different ways. So LUT usage stays low, power is minimal, timing does not change much. It is kind of expected.
FPGA performance ties back to the architecture and what resources you have. Boards like Basys 3 fit basic designs fine. Advanced ones are for when things get complicated. I am not totally sure, but that covers the main points.

VIII. REFERENCES

[1] F. Tenca and M. D. Ercegovac, “Synchronous Up/Down Binary Counter for LUT FPGAs with Counting Frequency

Independent of Counter Size," University of California, Los Angeles.

[2] D. Gupta, D. Mahajan, A B. Pandey, A. Kaur, A. Yadav, and A. Verma, "High Performance Energy Efficient Different Counters Design and Implementation on 28nm FPGA," International Journal of Control and Automation, vol. 9, no. 9, pp. 151–162, 2016.

[3] T. Kumar, B. Pandey, S. M. M. Islam, N. Singh, S. M. Alam, and T. Das, "Mapping Based Energy Efficient Counter Design on FPGA," Advanced Materials Research, vol. 984–985, pp. 1085–1088, 2014.

[4] S. M. Rana et al., "Study on the Application of FPGA Based Frequency Counter Using VHDL Program with Nexus Board," Institute of Electronics, Bangladesh Atomic Energy Commission, 2013.

TOMATO QUALITY CLASSIFICATION

Dr. Divya Mohan
Electronics and communication engineering
Ahalia School of Engineering and Technology
Palakkad, India
divya.mohan@ahalia.ac.in

Dr.V.Balamurugan
Electronics and communication engineering
Ahalia School of Engineering and Technology
Palakkad, India
hodece.aset@ahalia.ac.in

M Padmaja
Electronics and communication engineering
Ahalia School of Engineering and Technology
Palakkad, India
padmaja.m@ahalia.ac.in

Rayan Gaffor
Electronics and communication engineering
Ahalia School of Engineering and Technology
Palakkad, India
rayan.gaffor@ahalia.ac.in

S Chandru
Electronics and communication engineering
Ahalia School of Engineering and Technology
Palakkad, India
chandru.s@ahalia.ac.in

Sreelakshmi V
Electronics and communication engineering
Ahalia School of Engineering and Technology
Palakkad, India
sreelakshmi.v@ahalia.ac.in

Abstract— **Automated tomato quality categorization is crucial for enhancing post-harvest management and guaranteeing reliable delivery of fresh fruit. A deep learning-based method for classifying tomatoes into four different quality classes—Unripe, Ripe, Old, and Damaged—is presented in this work. Using a Kaggle dataset of tagged images, we Examine how well sophisticated neural architectures do this categorization problem. Although conventional convolutional neural networks (CNNs) have been extensively employed for comparable image-based evaluations, we suggest an improved approach that combines two cutting-edge models: ResNet, a hybrid model that combines a Random Forest classifier and the Swin Transformer, is renowned for its capacity to extract intricate visual information via residual learning. While the Random Forest enhances decision-making through ensemble learning, the Swin Transformer uses a hierarchical attention mechanism to capture both local and global picture characteristics. This combination strategy tackles issues including uneven lighting, variations in variety, and varying degrees of ripeness among samples. In order to provide a dependable and scalable solution for automated tomato quality assessment, with potential applications in supply chain optimization and smart farming, we measure each model's performance using accuracy, precision, recall, and F1-score.**

Keywords— CNN, Resnet, Swin Transformer, Random Forest, Deep Learning, Image Classification.

I. INTRODUCTION

In order to minimize food waste in the agricultural supply chain and guarantee that consumers obtain fresh, ripe, and safe produce, tomato quality assessment is essential. Inspectors have traditionally assessed tomato quality by hand, which is a subjective procedure prone to mistakes and irregularities. In addition to being time-consuming, the human grading procedure is ineffective, especially when handling big amounts of tomatoes, which could result in losses in profitability and quality. Automation in quality evaluation has become crucial to addressing these issues, and deep learning methods have shown great efficacy in picture categorization tasks. The goal of this research is to use cutting-edge deep learning models to create an automated tomato quality classification system. The study aims to improve tomato quality detecting accuracy and efficiency by utilizing Convolutional Neural Networks (CNN), ResNet, and a hybrid technique that combines Swin Transformer with Random Forest (RF). Tomatoes will be divided into four groups by the proposed system: Unripe, Ripe, Old, and Damaged. By guaranteeing that only tomatoes that meet the required quality requirements are given to customers, these categories will help reduce food waste and increase supply chain efficiency. Kaggle's "Tomatoes Dataset," which includes a range of tomato photos categorized by quality, is the source of the dataset used to train and test the model. These photos will be used to train the system, which will then use deep learning architectures to extract important attributes and correctly categorize each tomato. This project's ultimate objective is to increase tomato grading speed and consistency, decrease human intervention, and streamline quality control procedures in agriculture. With potential uses in the classification of other fruits and vegetables, the system's success might go beyond tomatoes and provide a scalable solution for agricultural quality management.

II. METHODOLOGY

Using cutting-edge deep learning models, this study's methodology focuses on creating an automated tomato quality classification system. To divide tomatoes into four groups—Unripe, Ripe, Old, and Damaged—three distinct methods were used: Convolutional Neural Networks (CNN), Residual Networks (ResNet), and a hybrid Swin Transformer with Random Forest (RF).

A. Dataset Collection and Preprocessing

A labelled dataset was collected from Kaggle's "Tomatoes Dataset," containing images of tomatoes in various quality states. Pre-processing steps were applied to standardize.

- Resizing: All images were resized to 224×224 pixels.
- Data Augmentation: Techniques such as rotation, flipping, zooming, and scaling were applied to increase dataset variability and reduce overfitting.
- Normalization: Pixel values were rescaled to [0,1] range.

B. Convolutional Neural Network (CNN)

For tomato classification, CNNs were used as a baseline model. Convolutional layers for feature extraction, pooling layers for dimensionality reduction, fully connected layers for classification, and a softmax output layer for final predictions were all part of the architecture. Accuracy,

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

precision, recall, and F1-score were among the assessment metrics used to train the model using categorical cross-entropy loss and the Adam optimizer.

C. Residual Network (ResNet)

ResNet was implemented to employ residual connections (skip connections) to address vanishing gradient problems in deep networks. The tomato dataset was used to refine a pre-trained ResNet model that had been built on ImageNet. Convolutional layers, residual blocks, batch normalization, and ReLU activation functions made up the architecture. For multi-class classification, the last fully linked layer with softmax was employed. The same assessment parameters as the CNN were used to evaluate performance.

D. Swin Transformer + Random Forest (Hybrid Model)

A hybrid model was used that combined Random Forest for classification with Swin Transformer for feature extraction. The Swin Transformer divided input images into windows, applied self-attention mechanisms, and generated feature embeddings. These embeddings were then fed into Random Forest classifier, which improved classification. To maximize performance, hyper parameters as the number of trees and tree depth were adjusted.

E. Model Training and Evaluation

To ensure class balance, all models were trained using stratified train-test splits. ROC-AUC curves, accuracy, precision, recall, F1-score, and confusion matrices were used in the evaluation process. CNN's accuracy was 95%, ResNet's accuracy was 98%, and the Swin Transformer + Random Forest hybrid's accuracy was 87%. It performed well in certain classes but poorly in others.

F. Deployment

Real-time tomato quality categorization was made possible by the models' deployment via a Flask-based web application following training and assessment. Images could be uploaded by users, and the algorithm offered confidence scores and anticipated quality categories.

III. WORKING PRINCIPLE

Deep learning-based image analysis is the foundation of the suggested tomato quality classification system. categorization, assisted by Swin Transformer with Random Forest, CNN, and ResNet. The procedure starts with gathering tomato photos from the Kaggle dataset, which are then pre-processed by resizing, to provide consistency and resilience against changes like lighting and orientation, normalization and augmentation are used. Once preprocess, the images are passed through the classification models:

1. CNN extracts spatial features using convolutional and pooling layers and classifies tomatoes into Unripe, Ripe, Old, and Damaged using fully connected and softmax layers.
2. ResNet improves feature extraction by using residual (skip) connections, which enable the training of deeper networks and prevent vanishing gradient problems, thereby enhancing accuracy.
3. Swin Transformer + Random Forest combines hierarchical feature extraction through shifted window self-attention with ensemble-based decision-making, improving robustness and capturing both global and local image features

The models forecast the tomato's quality category following feature extraction and classification. Users can upload tomato photos and promptly receive classification results with confidence ratings thanks to the integration of the trained models into a Flask-based web application. In order to replace manual inspection and ensure greater accuracy, speed, and consistency in tomato quality evaluation, the working principle is based on automated feature extraction, learning, and categorization utilizing sophisticated deep learning models.

IV. SYSTEM DESIGN

A. Introduction of Input Design:

In an information system, input is the raw data that is processed to produce output. During the input design, the developers must consider the input devices such as PC, MICR, OMR, etc. Therefore, the quality of system input determines the quality of system output. Well-designed input forms and screens have following properties

- It should serve specific purpose effectively such as storing, recording, and retrieving the information.
- It ensures proper completion with accuracy.
- It should be easy to fill and straightforward.
- It should focus on user’s attention, consistency, and simplicity.
- All these objectives are obtained using the knowledge of basic design principles regarding

B. Objectives for Input Design:

The objectives of input design are to design data entry and input procedures

- To reduce input volume
- To design source documents for data capture or devise other data capture methods
- To design input data records, data entry screens, user interface screens, etc.
- To use validation checks and develop effective input controls.

C. Output Design:

The design of output is the most important task of any system. During output design, developers identify the type of outputs needed, and consider the necessary output controls and prototype report layouts.

D. Objectives of Output Design:

The objectives of input design are:

- To develop output design that serves the intended purpose and eliminates the production of unwanted output.
- To develop the output design that meets the end user’s requirements.
- To deliver the appropriate quantity of output.
- To form the output in appropriate format and direct it to the right person.
- To make the output available on time for making good decisions.

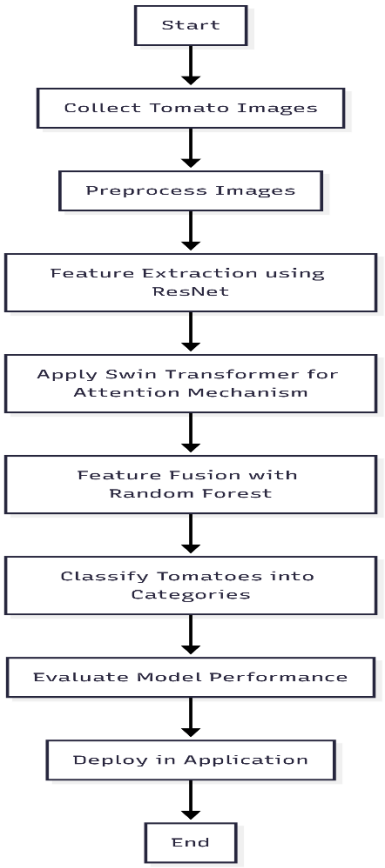


Fig. 1. Architecture

V. RESULTS AND DISCUSSION

Output Screens:

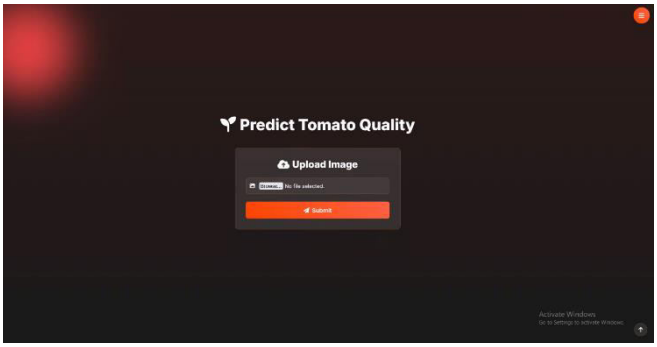


Fig. 2. image quality prediction

This is for predicting tomato quality. It allows users to upload an image of a tomato and then submit it for analysis to receive a quality prediction

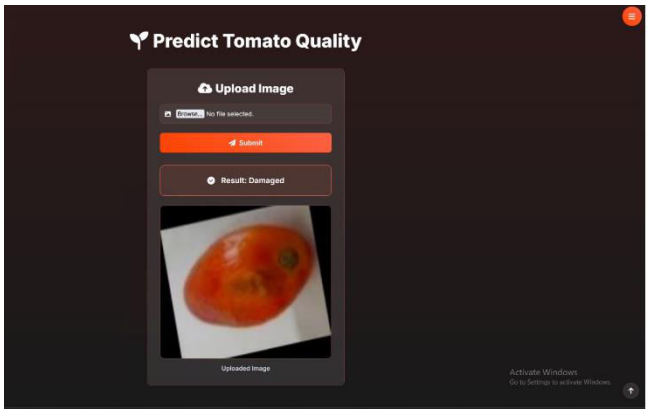


Fig. 3. Result of an image quality prediction

This shows a successful tomato quality prediction. A user has uploaded an image of a tomato, and the system has analyzed it and returned the result "Damaged," identifying a flaw in the fruit

VI. CONCLUSION

The experimental results demonstrate that all three models—CNN, ResNet, and Swin Transformer + Random Forest—are capable of classifying tomato quality into multiple categories, with varying degrees of effectiveness. The CNN model achieved a high accuracy of 95%, performing reliably across most classes, though it showed slightly lower recall for Class 2, indicating minor misclassifications. The ResNet model outperformed CNN with 98% accuracy, delivering consistently high precision, recall, and F1-scores across all classes, highlighting its robustness and suitability for tomato quality classification. In contrast, the Swin Transformer + Random Forest model reached 87% accuracy, with strong performance for some classes but poor recall for Class 2, reflecting sensitivity to class imbalance and inconsistency in predictions. Overall, ResNet emerges as the most reliable model for accurate and balanced tomato quality classification, while the other models provide valuable insights and could be further optimized to handle imbalanced classes more effectively.

REFERENCES

[1] R. K. Bania, S. Dey, and N. Buragohain, "Hybrid ResNet50-PCA based deep transfer learning approach for classification of tomato leaf diseases," Environmental Science, vol. 42, 2023. Available: www.slideshare.net.

[2] P. Ansah, S. K. Tetarave, E. Kalaimannan, and C. John, "Tomato disease fusion and classification using deep learning," International Journal of Cybernetics and Informatics, vol. 12, no. 7, pp. 31–43, 2023. Available: www.slideshare.net.

[3] J. Chen, X. Chen, and Y. Zhang, "Tomato leaf disease classification using multiple feature extraction techniques," Wireless Personal Communications, vol. 115, pp. 633–651, 2020.

[4] M. H. Yap, G. Pons, J. Marti, S. Ganau, M. Sentis, R. Zwiggelaar, A. K. Davison, and R. Marti, "Comparing vision transformers and convolutional neural networks for image classification," Applied Sciences, vol. 13, no. 9, 2023.

[5] S. Guo, Y. Zhang, and H. Liu, "Spectral-Swin Transformer with spatial attention for hyperspectral image classification," Remote Sensing, vol. 15, no. 10, 2023. Available: MDPI.

[6] M. H. Yap, G. Pons, J. Marti, S. Ganau, M. Sentis, R. Zwiggelaar, A. K. Davison, and R. Marti, "Comparing vision transformers and convolutional neural networks for image classification," Applied Sciences, vol. 13, no. 9, 2023.

TRACK-4

ELECTRICAL & ELECTRONICS ENGINEERING

SMART INFUSION PUMP USING IOT

Mohammed A
Electrical and electronics Engineering
Ahalia school of engineering and technology
Kerala.India
mohammed.a@ahalia.ac.in

Rahul R
Electrical and electronics Engineering
Ahalia school of engineering and technology
Kerala.India
rahulr666r@gmail.com

Akshay Kumar M
Electrical and electronics Engineering
Ahalia school of engineering and technology
Kerala,India
akshaykumarappu30@gmail.com

Karthika G
Electrical and electronics Engineering
Ahalia school of engineering and technology
Kerala,India
karthikagirish2022@gmail.com

Ms. Rosu Varkeyachan Padayatty
Electrical and electronics Engineering
Ahalia school of engineering and technology
Kerala.India
rosupadayatty@gmail.com

Abstract—Intravenous infusion therapy is a fundamental clinical procedure used for delivering fluids, nutrients, and medications directly into a patient's bloodstream. Although conventional infusion pumps are widely used, they still rely heavily on manual supervision, increasing the risk of dosage errors, delayed response to complications, and higher workload for healthcare professionals. This paper presents the design and implementation of a smart infusion pump that integrates embedded systems, biomedical sensors, and Internet of Things (IoT) technology to enhance patient safety and infusion accuracy. The proposed system is built around an ESP32 microcontroller and incorporates sensors for temperature, heart rate, oxygen saturation (SpO_2), pressure, and fluid weight measurement using a load cell. Wi-Fi connectivity enables remote monitoring, data logging, and centralized supervision of infusion therapy. Experimental testing demonstrates reliable operation, accurate dosage delivery, and effective safety responses. The proposed solution offers a cost-effective and scalable alternative to commercial smart infusion pumps, making it suitable for deployment in resource-constrained healthcare environments.

I. INTRODUCTION

Intravenous (IV) infusion therapy plays a crucial role in modern healthcare by enabling controlled administration of fluids, blood products, nutrients, and medications. Despite these advantages, conventional infusion pumps still require continuous monitoring by healthcare professionals to ensure patient safety and correct dosage delivery. Human errors such as incorrect parameter settings, and failure to respond promptly to abnormal patient conditions remain major concerns.

Medication errors associated with infusion therapy can lead to serious complications, including under-infusion, over-infusion, and adverse drug reactions. These risks are particularly significant in intensive care units, neonatal wards, and oncology departments, where patients are highly vulnerable and medication dosages must be extremely precise. Furthermore, the increasing patient-to-nurse ratio in many hospitals places additional pressure on healthcare staff, increasing the likelihood of oversight.

Advancements in embedded systems, biomedical instrumentation, and IoT technologies have enabled the development of intelligent medical devices capable of autonomous monitoring. Smart infusion pumps integrate patient vital monitoring with automated flow control and remote supervision, reducing reliance on manual intervention. By continuously analysing patient parameters and infusion conditions, such systems can detect abnormal situations early and initiate corrective actions.

This paper proposes a smart infusion pump system based on the ESP32 microcontroller, designed to provide accurate fluid delivery, and remote data access. The primary objective is to develop a low-cost yet reliable solution that improves patient safety, reduces healthcare provider workload, and supports modern smart healthcare infrastructure.

II. LITREATURE REVIEW

[1] (smith, J. A., Johnson, R. K. (2023))- The study explains that integrating IoT technology with medical devices enables real-time patient monitoring, remote data access, and improved treatment accuracy. It highlights that IoT-based systems can reduce human error and enhance patient safety through automated monitoring and alerts. These principles support the development of smart healthcare devices such as the IoT-enabled infusion pump.

[2] (Williams, M. T., Davis, L. P., Brown, K. S. (2024)) The study explains Journal of Biomedical Engineering and Technology reports the development of a smart infusion pump using the ESP32 microcontroller for automated and accurate fluid delivery. The findings show that integrating sensors and wireless connectivity enables real-time monitoring, improved dosage accuracy, and remote supervision, thereby enhancing patient safety and reducing manual workload in healthcare environments.

[3] (Chen, H., Zhang, Y., Liu, X. (2023).) The study reviews modern real-time patient monitoring systems used in healthcare. It finds that integrating biomedical sensors, IoT communication, and data processing technologies enables monitoring of vital signs and faster detection of critical conditions. This improves patient safety, supports remote healthcare management.

[4] (Anderson, P. R., Martinez, C. D. (2022) The study discusses advanced safety mechanisms used in modern infusion pump technology. It finds that features such as controlled drug delivery systems help prevent medication errors and equipment failures. These safety mechanisms significantly improve patient safety and reliability in infusion therapy

[5] (Patel, S. K., Kumar, V., Singh, A. (2024)) The study explains the integration of biomedical sensors for patient monitoring in healthcare systems. It finds that combining multiple sensors enables accurate measurement of vital parameters such as heart rate, oxygen saturation, and temperature in real time. This improves early detection of health issues and supports more effective patient care.

[6] (Thompson, E. L., Roberts, M. A. (2023)) The study explains the role of embedded systems in the design and development of modern medical devices. It finds that microcontrollers, sensors, and control algorithms enable accurate monitoring, automation, and reliable operation of medical equipment. These technologies improve the efficiency, safety, and performance of healthcare devices.

[7] (Lee, J. H., Kim, S. Y., Park, D. W. (2024)) The study explains the use of load cell sensors for precise measurement in medical fluid delivery systems. It finds that load cell technology enables accurate monitoring of fluid volume and dosage during infusion therapy. This improves the reliability of fluid delivery and helps prevent dosage errors in medical treatments.

[8] The report explains the global regulations and standards for the safety and performance of medical devices. It highlights the importance of quality control, risk management, and regulatory compliance to ensure safe and effective use of medical technologies. These guidelines support the development and approval of reliable healthcare devices worldwide.

[9] (Garcia, R. F., Wilson, T. M. (2023)) The study explains the role of smart infusion technology in preventing medication administration errors. It finds that automated infusion systems with monitoring and alert mechanisms help ensure accurate drug dosage and reduce human errors during treatment. This improves patient safety and reliability in medication delivery.

[10] (Miller, D. K., Jackson, P. L. (2023)) he study explains the importance of smart infusion technology in reducing medication administration errors. It shows that automated infusion systems with monitoring, alerts, and dosage control can improve accuracy in drug delivery and reduce human mistakes. This helps enhance patient safety and treatment reliability in healthcare settings.

[11] (Miller, D. K., Jackson, P. L. (2023)) The study discusses the development of cost-effective healthcare solutions using open-source hardware and software platforms. It finds that open-source technologies enable affordable medical devices and healthcare systems while maintaining good performance and flexibility. This approach helps improve access to healthcare technology, especially in resource-limited settings.

[12] (Kumar, A., Sharma, R. (2024)) The report explains the importance of cybersecurity frameworks for protecting medical IoT devices and healthcare data. It highlights the need for secure communication, data protection, and risk management strategies to prevent cyber threats. These measures ensure the safe and reliable operation of connected medical devices in healthcare systems.

III. BLOCK DIAGRAM

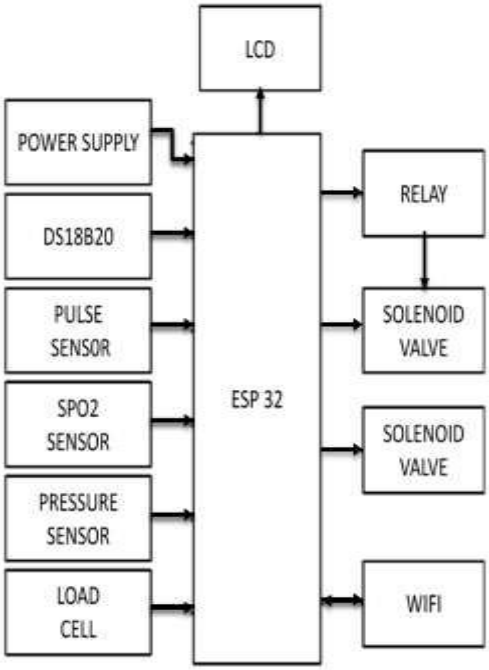


Fig-3.1

- The smart infusion pump system is designed as an integrated solution comprising sensing, control, actuation, display, and communication modules. Figure representations of the system architecture illustrate the interaction between these components. The ESP32 microcontroller serves as the core processing unit, coordinating data acquisition, control logic execution, and communication tasks.

- The sensing subsystem includes temperature, pulse, SpO_2, pressure, and load cell sensors. These sensors continuously monitor patient vitals and infusion parameters. The actuation subsystem consists of solenoid valves and relay modules that regulate fluid flow based on control decisions. A local display unit provides real-time feedback, while the communication subsystem enables remote monitoring via Wi-Fi.

IV. HARDWARE DESIGN

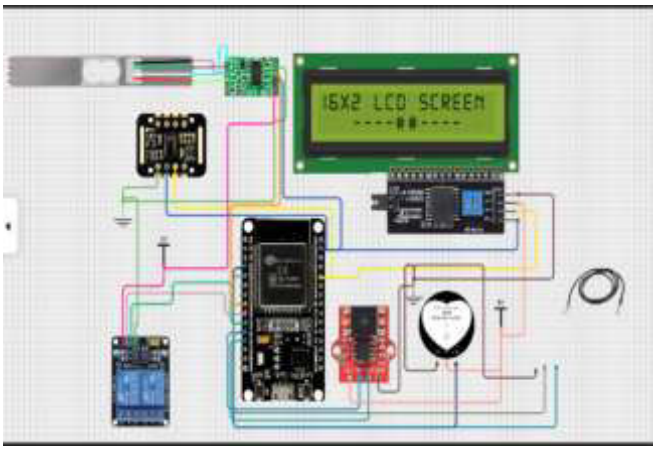


Fig-4.0.1

A. Esp 32

- The ESP32 microcontroller is selected due to its dual-core processing capability, integrated Wi-Fi and Bluetooth modules, low power consumption, and extensive peripheral support. These features make it suitable for real-time data processing and wireless communication in medical applications.
- The ESP32 continuously collects and processes data from various sensors such as temperature, pulse, SpO_2, pressure, and load cell. It analyzes these readings in real time and compares them with predefined threshold values. Based on this

processing, it determines whether the system is operating under normal or abnormal condition.

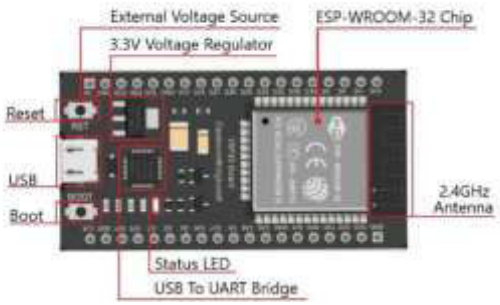


Fig-4.0.2

B. Biomedical sensor

- The system integrates multiple biomedical sensors to monitor patient condition comprehensively. A DS18B20 digital temperature sensor measures body or fluid temperature with high accuracy.

- A pulse sensor detects heart rate using photoplethysmography, while a SpO_2 sensor measures blood oxygen saturation levels. A pressure sensor installed in the infusion line detects occlusions, blockages, or abnormal pressure variations. A load cell measures the weight of the infusion fluid, enabling precise calculation of the volume delivered

- The sensors convert physical biological signals into electrical signals that the microcontroller (ESP32) can understand. For example, the pulse sensor detects heartbeats, the SpO_2 sensor measures oxygen levels in blood, and the temperature sensor monitors body temperature These signals are then sent to the ESP32 for processing.

Fig-4.0.3

1) Pulse sensor

- This circuit utilizes an infrared (IR) transmitter and receiver pair, clipped to the finger, to measure heart rate by detecting blood flow variations. The IR rays are interrupted by the blood flow, causing the receiver to generate variable pulse signals via a potential divider. These signals are passed through a coupling capacitor (C1) to block DC components and are then processed by an LM324 quad operational amplifier.
- The pulse sensor is used to measure the patient's heart rate (beats per minute) . It helps in monitoring the cardiovascular condition of the patient during the infusion, ensuring that the patient remains stable.

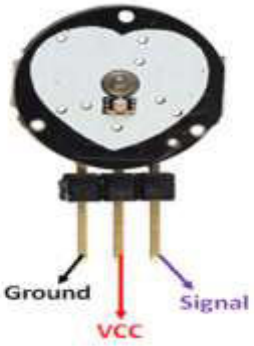


Fig-4.0.4

2) Schematic of pulse sensor

- The amplification stage, consisting of amplifiers A1 through A4, filters noise and spikes using low-pass configurations and capacitors C2, C3, and C4. A3 provides an offset voltage to A4 to refine the signal before it is sent to a switching transistor pair (BC557 and BC547). These transistors, followed by a 74C04 inverter, convert the signal into a TTL-compatible square wave. This digital output is then suitable for microcontroller interfacing to monitor heart rate, which typically ranges from 60 to 100 beats per minute in healthy adults.

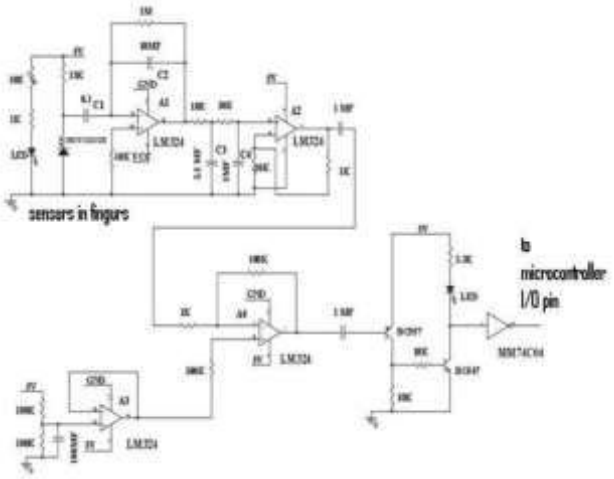


Fig-4.0.6

C. Actuators and Display

Solenoid valves controlled via relay modules regulate the flow of infusion fluid. An LCD display provides real-time visualization of sensor readings, infusion rate, volume delivered, and system status. Audible and visual alarms alert healthcare providers during abnormal conditions.

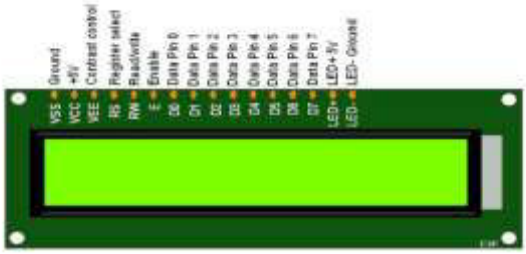


Fig-4.0.7

D. *Power supply*

- A regulated power supply ensures stable voltage levels for all components. Proper isolation and grounding techniques are applied to enhance electrical safety and system reliability.

1) Relay

- A relay is an electrically operated switch in which current flowing through a coil generates a magnetic field that actuates a mechanical lever to change the state of the switch contacts. This enables a low-power control circuit to safely switch a high-power or high-voltage circuit while maintaining complete electrical isolation between them. Since relay coils typically require higher current than most integrated circuits can supply, a transistor driver is commonly used. Relays are available in configurations such as SPDT and DPDT, with terminals designated as common (COM), normally closed (NC), and normally open (NO), allowing flexible control of connected loads.

Fig-4.0.8

a) Schematic of relay

- The proposed circuit controls an electrical load, such as a motor, using a relay driven by a pair of BC547 switching transistors. The relay coil is connected to the collector of transistor Q2, while the relay contacts include common (COM), normally closed (NC), and normally open (NO) terminals. When a high-level (5 V) input is applied to the base of transistor Q1, Q1 conducts and drives Q2 into the OFF state, thereby de-energizing the relay and keeping the load disconnected. Conversely, when a low-level input is applied to Q1, Q2 is driven into conduction, energizing the relay and connecting the COM and NO contacts, which supplies power to the load.

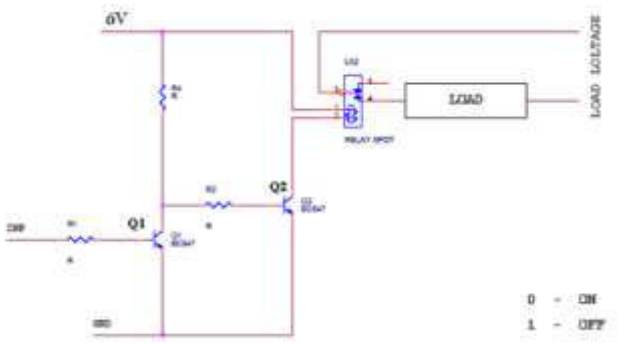


Fig-4.0.9

E. Arduino Pressure Sensor | MPS20N0040D

- The MPS20N0040D pressure sensor operates as a Wheatstone bridge in which one resistor varies with applied air or liquid pressure. When no pressure is applied, the bridge remains balanced and produces zero output voltage.

- Application of pressure causes a slight change in resistance, unbalancing the bridge and generating a small differential voltage proportional to the applied pressure. Since this voltage is extremely low and beyond the resolution of the Arduino's 10-bit ADC, an external high-resolution ADC is required. The HX710B, a 24-bit analog-to-digital converter, is used to amplify and digitize the sensor output.

- The pressure sensor will measure body pressure such as both systolic and diastolic.

Fig-4.1.1

F. PCF8574 I2C I/O Expander

- The PCF8574 is an I2C-based 8-bit general-purpose input/output (GPIO) expander designed to extend the digital I/O capability of microcontrollers using only two communication lines, namely SDA and SCL. It operates over a wide voltage range of 2.5 V to 6 V and supports standard (100 kHz) and fast (400 kHz) I2C data rates. Each of its eight I/O pins is configured as open-drain, requiring external pull-up resistors for logic-high operation.

- The device address can be configured using three address pins (A0–A2), allowing multiple PCF8574 devices to share the same I2C bus. Owing to its low power consumption and minimal wiring requirements, the PCF8574 is commonly used for interfacing LEDs, switches, relays, and other digital peripherals in GPIO-constrained embedded systems.

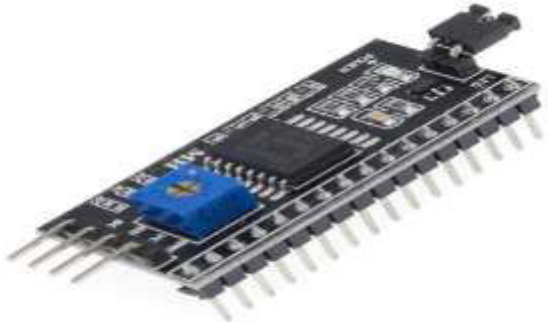


Fig-4.1.2

G. TAL220 Load Cell

- The TAL220 is a straight-bar strain gauge load cell with a rated capacity of 10 kg, commonly used for precise force and weight measurement. It is typically fabricated from aluminum or alloy steel and produces a low-level millivolt output proportional to the applied load. The sensing mechanism is based on strain gauges bonded to the load cell structure, which undergo a change in electrical resistance when subjected to deformation under load, in accordance with Hooke's law.

- These strain gauges are arranged in a Wheatstone bridge configuration, and any resistance variation unbalances the bridge, generating a measurable differential voltage. After amplification and signal conditioning, this output enables accurate measurement of both tension and compression forces. Owing to its reliability and accuracy, the TAL220 is widely employed in industrial weighing systems, material testing, process control, medical devices, and research applications.

- The load cell is used to measure the weight of the IV fluid bag continuously. This helps in tracking how much fluid is remaining . With help of the load cell the iv bag weight can be measured and find how much drug is remaining

in the bag if the drug is below 15 percentage we will get an notifications in blynk app

- There are two size of bag one is 350 gram and another is 1000 gram so the load cell will detect and pass message to the esp.the app will notify when the bag weight is less than 15 percentage.

Fig-4.1.3

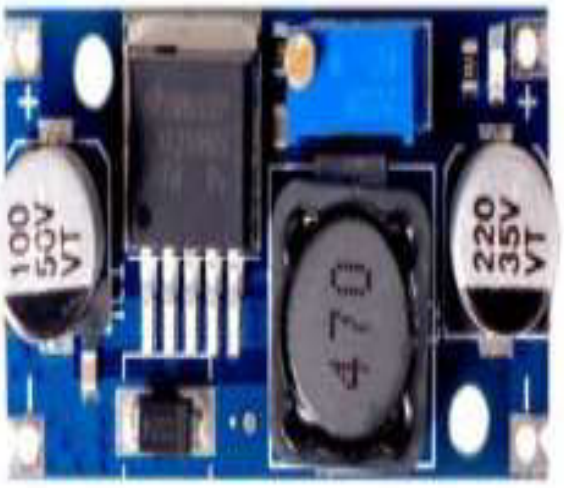


Fig-4.1.4

H. LM2596 DC–DC Buck Converter

- The LM2596 is a SIMPLE SWITCHER step-down (buck) DC–DC voltage regulator capable of delivering up to 3 A of output current with excellent line and load regulation. It supports a wide input voltage range of up to 40 V and is available in fixed output versions of 3.3 V, 5 V, and 12 V, as well as an adjustable output variant with a tolerance of ±4%.

- Operating at a fixed switching frequency of 150 kHz, the regulator enables the use of compact external filter components while requiring only a minimal number of external parts. The device integrates internal frequency compensation, a fixed-frequency oscillator, thermal shutdown, current-limit protection, and an external shutdown feature with low standby current. Owing to its high efficiency, ease of implementation, and robust protection features, the LM2596 is widely used in power supplies for appliances, industrial systems, grid infrastructure, and consumer electronics.

1) Block diagram of buck converter

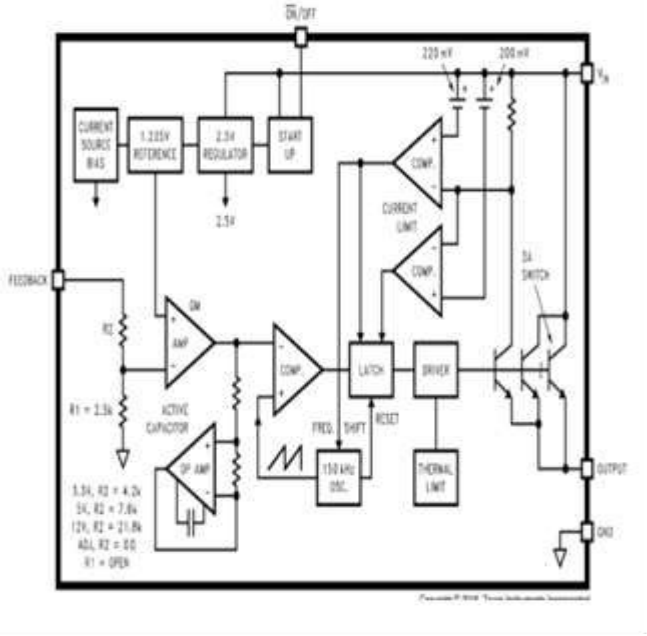


Fig-4.1.5

V. SOFTWARE ARCHITETURE

- The software architecture follows a modular design approach to enhance maintainability and scalability. The firmware running on the ESP32 is responsible for sensor data acquisition, signal conditioning, control logic execution, alarm handling, display updates, and wireless communication.

- Sensor readings are periodically sampled and filtered to reduce noise. The control algorithm evaluates sensor data against predefined thresholds to determine appropriate actions. When abnormal conditions are detected, the system triggers alarms and connectivity enables data transmission to a remote monitoring interface for centralized supervision.

I. ARDUINO IDE

1) Introduction

- Arduino IDE stands for Integrated Development Environment. It is an open-source official software provided by Arduino.cc used for writing, compiling, and uploading programs to Arduino boards. The IDE supports C and C++ programming languages and is easy to use even for beginners. It is compatible with operating systems like Windows, macOS, and Linux. The software works with various Arduino boards such as Arduino Uno, Mega, and Leonardo. It consists mainly of an editor for writing code and a compiler for compiling and uploading it to the microcontroller.

2) Connect the Arduino Board

- Plug in the Arduino using the USB cable
- Open Arduino IDE
- Go to Tools > Board and select your board model (e.g., Arduino Uno)
- Go to Tools > Port and select the correct COM port (e.g., COM3 on Windows)

3) Arduino Software (IDE)

- The Arduino IDE includes built-in support for many Arduino boards based on the AVR core, such as Arduino Uno, Nano, Mega, Leonardo, and Micro. It supports boards using microcontrollers like ATmega328P, ATmega32u4, and ATmega2560 running mainly at 16 MHZ The Boards Manager allows users to add support for newer boards based on different cores like Arduino Due, Zero, Edison, and Galileo. These boards provide various numbers of analog inputs, digital I/O pins, and PWM outputs. Features like auto-reset and built-in bootloaders make programming easier. This wide compatibility makes Arduino IDE suitable for many embedded applications.

J. BLYNK APPLICATION

1) HOW BLYNK WORK

- Blynk is an Internet of Things (IoT) platform used to control hardware remotely and monitor sensor data. It allows users to visualize, store, and manage data through a smartphone app. The platform has three main components: the Blynk App, Blynk Server, and Blynk Libraries. The Blynk App helps create user-friendly interfaces using widgets. The Blynk Server manages communication between the mobile app and hardware devices. Blynk Libraries enable smooth data exchange between the hardware and server in real time.

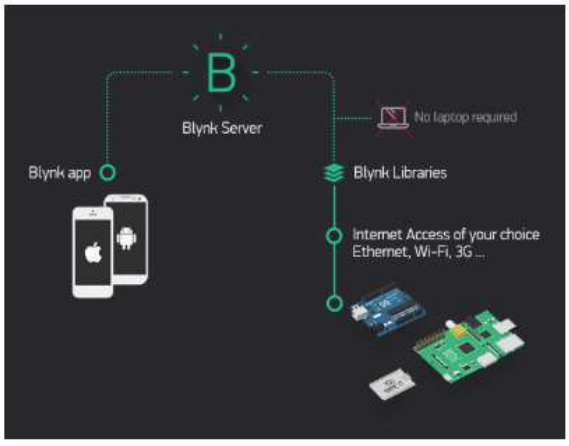


Fig-5.1

2) FEATURES OF BLYNK APPLICATION

- Provides a common API and user interface for all supported hardware

- Supports cloud connectivity via Wi-Fi, Bluetooth/BLE, Ethernet, USB (Serial), and GSM
- Offers easy-to-use widgets for creating interactive interfaces
- Allows direct pin control without writing code
- Supports virtual pins for adding advanced functionality
- Enables data monitoring, device-to-device communication, and notifications (email, push, etc.)

3) HOW TO USE BLYNK

- Auth Token Security: Do not share your Auth Token with anyone, as it provides full access to your hardware. You can send the token to your registered email using the email option or copy it directly by tapping on the Token line.

Fig-5.2

- Project Creation: After verifying the Auth Token and completing the basic setup, press the Create button to create your Blynk project.

Fig-5.3

- Adding Widgets: Once the project canvas opens and is empty, tap anywhere on the screen to open the Widget Box, which contains all available widgets.

Fig-5.4

- Button Widget Selection: Select a Button widget from the widget box to control devices such as an LED.

VI. RESULTS AND DISCUSSION

- The developed Smart Infusion Pump Monitoring System successfully collected real-time patient health parameters using the IoT platform. The system retrieves live data from the Blynk server and displays the patient information and physiological parameters.
- For the experimental test, the patient details were entered as Jancy, aged 34 years. The system monitored several parameters including heart rate (BPM), temperature factor (TF), oxygen saturation (SpO_2), load cell measurement (Grams), pressure, and systolic pressure (SYS).

- The weight of the bag used here is 350grams and the weight in which app will notify is approximately 50 grams

- In the first test case(fig-6.1), the obtained values were PULSE = 64bpm, TF = 95deg F, SpO_2 = 97%,WIEGHT OF THE BAG= 52g, SYSTOLIC PRESSURE= 110mm/Hg, and DIASTOLIC PRESSURE = 66mm/Hg. (fig-6.1).The weight of the bag is less 52 grams which is less than 50% of the total weight of the bag ie 350 grams , the app notified as shown in the fig-6.2

- In the second test case(fig6.3), the system demonstrated another scenario where PULSE = 64bpm, TF = 95f, SpO_2 = 98%, BAG WEIGHT = 250g, DIASTOLIC = 71mm/Hg, and SYSTOLIC=119mm/Hg. the system processed the available data and generated the patient report here the bag weight shown is 250 grams hence there is no notification.

- From these results it can be understood that the app will notify when the weight of the bag is less than 15 percentage of total weight

SAMPLE-1

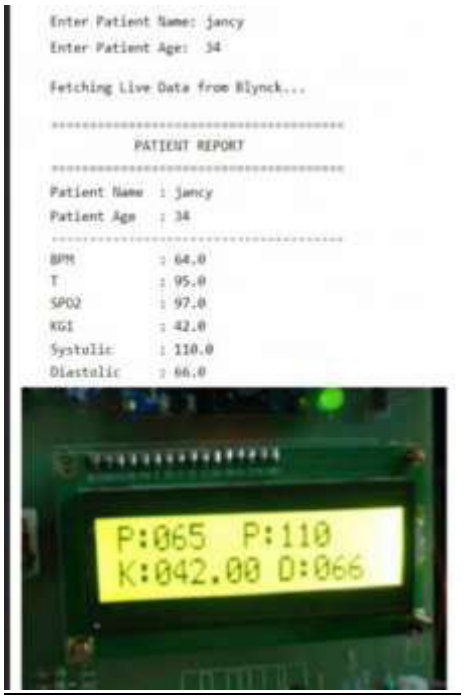


Fig-6.1

Fig-6.2

SAMPLE-2

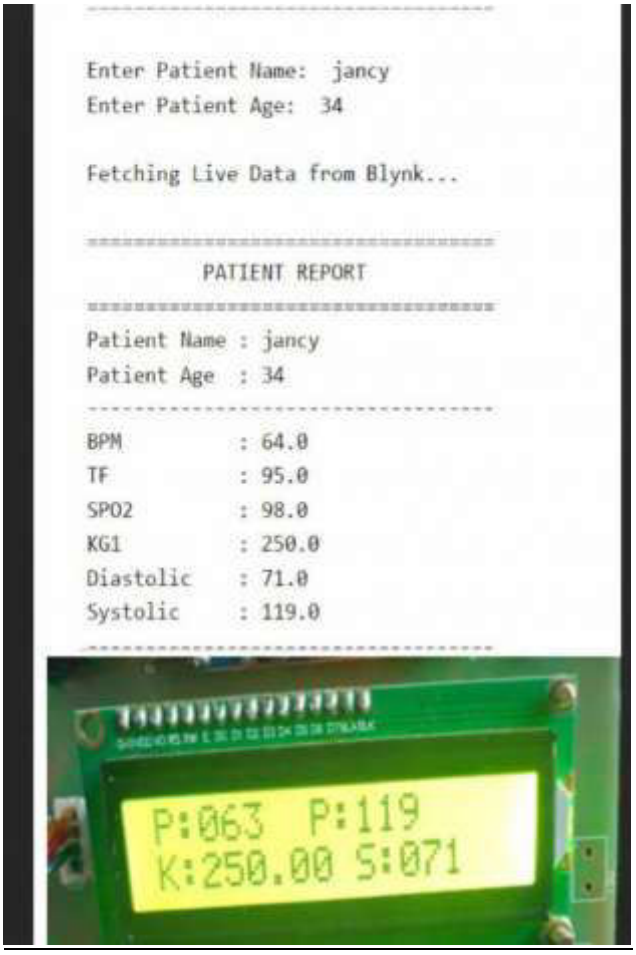


Fig-6.3

VII. ADVANTAGES

- Monitors vital parameters like temperature, pulse rate, and SpO_2
- Doctors can monitor multiple patients remotely using Wi-Fi, improving efficiency and response time.
- Controls infusion, reducing human errors in manual adjustments.
- Uses low-cost components like ESP32, making it affordable compared to commercial pumps.
- Includes automatic shut-off notification and fault detection for patient safety.
- Minimizes manual monitoring, allowing healthcare staff to focus on other tasks.

VIII. FUTURE SCOPE

- Can use ML algorithms to predict patient complications and adjust infusion automatically.
- Develop mobile apps for real-time monitoring and alerts on smartphones.
- A centralized dashboard to monitor many patients simultaneously in hospitals.
- Add sensors like ECG, glucose, and respiratory sensors for better monitoring.
- Make the device smaller and portable for use in ambulances and home care.
- Fully automatic system that adjusts drug delivery based on patient condition without manual input.

IX. CONCLUSION

- The Smart Infusion Pump project demonstrates an effective use of embedded systems and IoT technology to improve healthcare delivery. By integrating multiple biomedical sensors with the ESP32 microcontroller, the system ensures accurate monitoring of patient vital signs along with fluid delivery.
- The system significantly enhances patient safety and accuracy compared to conventional infusion pumps. Features like monitoring, load cell-based measurement, help reduce human errors.
- Another important outcome of this project is its cost-effective design. By using affordable components such as ESP32 and standard sensors, the system provides advanced functionality at a much lower cost than commercial infusion pumps, making it suitable for small hospitals and rural healthcare centers.
- Overall, the project provides a reliable, efficient, and scalable solution for modern healthcare systems. With features like remote monitoring, data logging, and automation, it contributes to improved patient care and lays a strong foundation for future smart medical device development.

X. REFERENCES

1. B. D. Shah and R. A. Shah, “Design and development of a smart infusion pump for precise drug delivery,” IEEE International Conference on Biomedical Engineering, pp. 112–116, 2018.
2. S. K. Sahoo, P. Mohanty, “Microcontroller based smart infusion pump system,” International Conference on Medical Electronics and Devices, pp. 45–49, 2019.
3. A. Patel and M. Mehta, “Automated intravenous infusion monitoring system using IoT,” IEEE International Conference on Smart Healthcare, pp. 210–214, 2020.
4. J. G. Webster, “Medical instrumentation and infusion pump technologies,” IEEE Transactions on Biomedical Engineering, vol. 65, no. 4, pp. 850–856, 2017.
5. R. S. Khandpur, “Design considerations of medical infusion devices,” Biomedical Instrumentation Conference, pp. 67–72, 2016.
6. L. M. Bui and T. Nguyen, “Development of an intelligent infusion pump with safety monitoring,” IEEE Healthcare Innovation Conference, pp. 134–138, 2019.
7. P. Sharma and N. Gupta, “IoT based infusion pump monitoring system,” International Conference on Internet of Medical Things, pp. 98–103, 2021.
8. D. Arney et al., “Challenges in developing safe and reliable infusion pumps,” IEEE Engineering in Medicine and Biology Conference, pp. 547–550, 2017.
9. M. Rahman and K. Islam, “Design of a low-cost programmable infusion pump for hospitals,” International Conference on Biomedical Systems, pp. 120–124, 2018.
10. H. Park and J. Lee, “Smart medical devices for automated drug delivery systems,” IEEE

International Conference on Healthcare Informatics, pp. 301–305, 2020.
11. Kumar, A., Sharma, R. (2024). ”Cost-Effective Healthcare Solutions Using
12. National Institute of Standards and Technology (2023) Cybersecurity Framework for Medical IoT Devices. NIST Special Publication 1800-8

FINGERPRINT ELECTRONIC VOTING MACHINE

VAISAKH RS
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala , India
rsvaisakh093@gmail.com

ADITHYAN P
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala , India
adithyanparameswaran@gmail.com

AMJITH K
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala , India
amjitho10t@gmail.com

MIDHUN J
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala , India
jmidhuns23@gmail.com

Ms.SOWMYA S
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala , India
sowmya.s@ahalia.ac.in

***Abstract*— The Fingerprint based Voting system is a voting program in which the user is identified by his finger pattern. Because each person's finger pattern is unique, the voter can be easily validated. In this study, based on a voter's fingerprint, which is an Aadhar card number in a central government database. The government collects demographic data from residents in Aadhar's central database and assigns each individual a 12- digit unique identity number. Fingerprints from the Aadhar database are employed for candidate verification and are saved on the computer on which the rest of the process is carried out. The Aadhar card has all biometric details using that if fingerprint is recognized, user details are shown and verified. This study is premeditated for the electronic voting machine using the fingerprint identification method., If it matches then the system will allow the voter to poll his vote and otherwise avert the votes from polling second time polling will not be allowed**

I. INTRODUCTION

In a democratic country like India, voting is a fundamental right and an essential process that ensures the selection of leaders by the people. However, traditional voting systems—whether manual ballot or standard electronic voting machines (EVMs)—often face challenges such as fake voting, multiple voting by a single person, and identity verification issues. To overcome these problems and enhance the transparency, security, and reliability of the election process, the concept of an. Fingerprint Electronic Voting Machine (EVM) is introduced. This project integrates biometric authentication using Aadhar data with an electronic voting system. Each voter's fingerprint is verified against the Aadhar database to confirm their identity before allowing them to cast a vote. This eliminates the possibility of duplicate or fraudulent voting and ensures that only eligible voters participate. The system is designed using components such as a microcontroller, fingerprint sensor, LCD display, and voting interface, along with secure database connectivity for Aadhar verification. The main objectives of this project are to make the voting process more secure, reduce human errors, increase efficiency, and strengthen public confidence in the electoral system. By implementing an. fingerprint electronic voting machine, the election process can become more transparent, tamper-proof, and technologically advanced, paving the way for a more robust and reliable digital.

II. LITERATURE REVIEW

1. N. Kaliammal — "Aadhar Based Fingerprint Electronic Voting Machine"

Aim: Propose a low-cost EVM that authenticates voters via Aadhar fingerprint. Method: Prototype implementation (Arduino + fingerprint module) and conceptual description. Findings: Describes voter authentication using Aadhar fingerprint templates to prevent impersonation and speed up polling. Limitations: Prototype-level, limited or no discussion of privacy/UIDAI rules or large-scale deployment issues.

2. CH. Srilatha (2024) — "Fingerprint-based biometric smart electronic voting

Aim: Address electoral fraud using fingerprint biometric EVMs. Method: System design and simulation/implementation details described for a biometric EVM. Findings: Presents improved voter authentication and lower risk of rigging. Limitations: Paper discusses potential benefits but lacks field trials and deep privacy/security analysis.

3. A. Jatain — "Design and Development of Biometric Enabled Advanced EVM" (2020)

Aim: Combine fingerprint and iris biometrics into an EVM to raise identification accuracy. Method: System design + prototype; advocates multi-modal biometrics to reduce false matches. Findings: Multi-modal biometrics can improve authentication accuracy and reduce duplicate registration risks.

4. N. Ranjan — "An Electronic Voting Machine (EVM) System Integrated with Biometrics" (2023)

Aim: Propose EVM integrating fingerprint and facial biometrics for secure voting. Method: System architecture, proposed workflow and limited experimental evaluation. Findings: Describes how biometrics can be integrated into

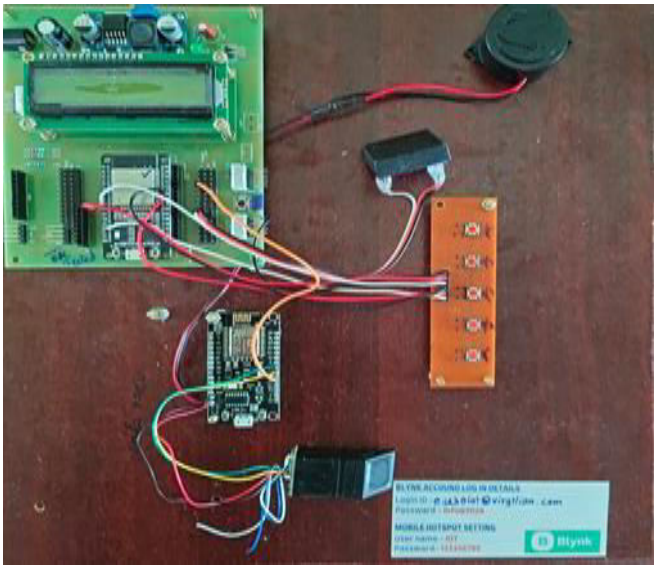
existing EVM workflows to reduce fraud. Limitations: Largely conceptual; lacks deployment case studies and legal/operational analysis

5. D. Ghosh B. Banerjee — "Fingerprint based electronic Voting machine: a review"

Aim: Survey fingerprint-based EVM approaches and hardware/software choices. Method: Review of several implementations and comparison of modules (fingerprint sensors, microcontrollers). Findings: Fingerprint verification reduces duplicate voting; many studies propose Arduino/Raspberry Pi prototypes.

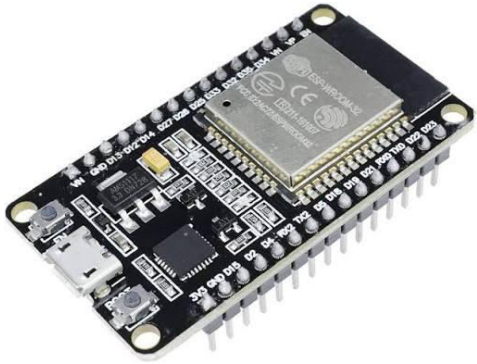
III. SYSTEM OVERVIEW

The Fingerprint Electronic Voting Machine is designed to enhance the security, accuracy, and reliability of the voting process by integrating biometric authentication with electronic voting technology. The methodology adopted for this project involves the design and implementation of a system that authenticates a voter's identity through fingerprint verification linked to their Aadhar data before allowing them to cast a vote. The system is built around a intelligent device unit that serves as the main control component, interfacing with peripheral devices such as a fingerprint sensor, LCD display, and voting database. When a voter arrives at the polling station, they are prompted to place their finger on the fingerprint sensor. The sensor captures the biometric pattern and compares it with the pre-stored fingerprint data associated with the voter's Aadhar number in the system's database. If the fingerprint matches successfully, the intelligent device enables access to the voting interface displayed on the LCD screen, where the voter can select their preferred candidate using the input keypad. Once the vote is cast, the system records it securely in the database and automatically marks the voter as having voted, preventing multiple entries by the same individual. In case of an authentication mismatch, the system denies access and displays an appropriate error message, ensuring that only valid and registered voters can proceed.



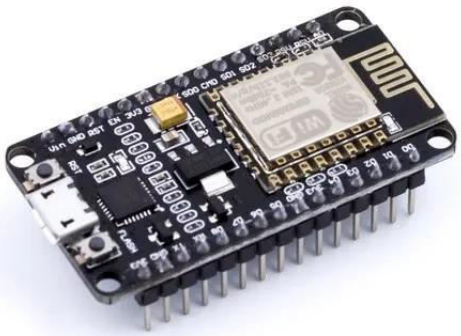
IV. HARDWARE COMPONENTS

A. *ESP32 MICROCONTROLLER*



The ESP32 microcontroller is the main control unit of the system. It processes data received from different sensors and modules and manages communication between them. The ESP32 reads GPS data, monitors temperature values, detects SOS button activation, and sends commands to the GSM module to transmit SMS alerts. Key features of ESP32 include high processing speed, multiple communication interfaces, low power consumption, and compatibility with IoT applications.

B. *ESP82 MICROCONTROLLER*

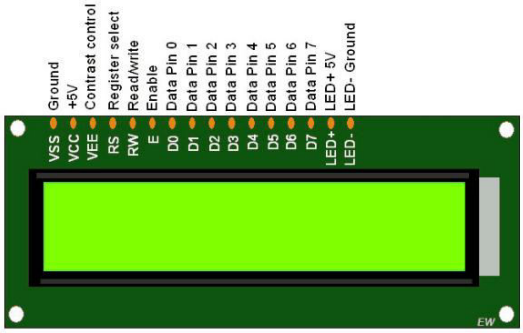


The ESP8266, developed by Espressif Systems, is a highly versatile and cost-effective system-on-chip (SoC) designed specifically for Wi-Fi-enabled embedded applications. At its core, it features a 32-bit Tensilica L106 RISC processor operating at 80 MHz (with an option to run at 160 MHz), supported by on-chip SRAM (typically around 50 KB available for user applications) and an external flash memory interface (commonly 512 KB to 4 MB or more depending on the module). The ESP8266 integrates a full IEEE 802.11 b/g/n WiFi transceiver, TCP/IP protocol stack, and networking libraries, allowing it to handle all networking functions independently without requiring a separate microcontroller.

C. *FINGERPRINT MODULE*

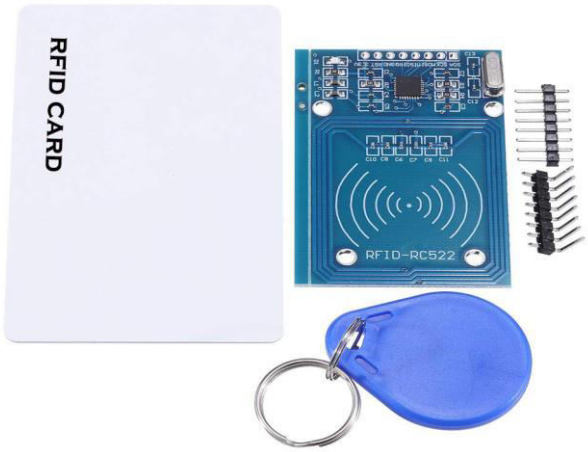
A fingerprint module is a biometric sensor device used to scan, store, and verify fingerprints for identification and security purposes. It captures the unique fingerprint pattern of a person, converts it into digital data, and stores it in memory as a template. When a user places a finger on the sensor, the module compares the scanned fingerprint with the stored data to check for a match. If the fingerprint is matched successfully, it sends the result to a controller such as ESP32 or Arduino for further action. Fingerprint modules are widely used in attendance systems, door locks, access control, and IoT-based security systems because they provide fast, accurate, and reliable authentication.

D. *LCD DISPLAY*



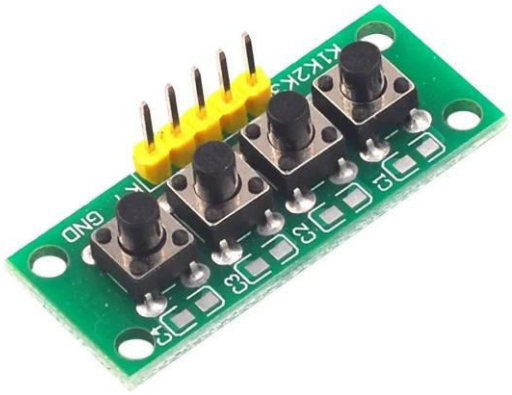
An LCD (Liquid Crystal Display) is an electronic display device used to present text, numbers, and symbols clearly to the user. It acts as an output unit in many embedded and IoT systems by displaying the status and results of the system operation. The LCD receives commands and data from a microcontroller such as ESP32, Arduino, or other controllers and then shows the required information on the screen. In a fingerprint-based attendance system, the LCD is used to display messages such as “Scan Finger”, “Fingerprint Matched”, “Attendance Recorded”, or “Access Denied”, which helps the user understand what is happening in the system. It provides a simple and effective way of communication between the machine and the user. LCD displays are widely used because they are easy to use, consume less power, provide good visibility, and improve the overall user interaction in electronic projects and automation systems.

E. *RFID*



Radio Frequency Identification (RFID) is a wireless communication technology used to automatically identify and track objects, animals, or people using radio waves. An RFID system mainly consists of three components: an RFID tag, an RFID reader, and a backend system or database. The RFID tag, which can be active (battery-powered) or passive (no internal power source), contains a microchip and antenna that store and transmit data. When the tag comes within the electromagnetic field generated by the RFID reader, it gets activated and sends its stored information, such as a unique identification number, back to the reader. The reader then converts this radio signal into digital data and forwards it to a computer system for processing and storage.

F. *BALLOUT UNIT*



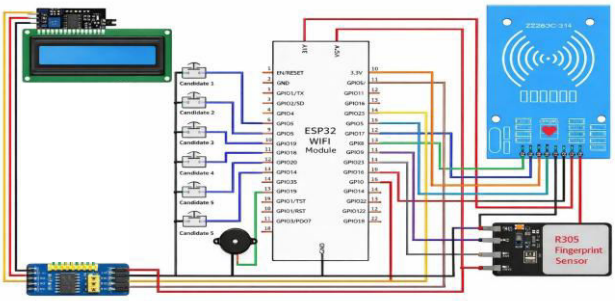
A ballot unit is the part of an Electronic Voting Machine (EVM) through which voters cast their vote during an election. It contains the names of candidates, their symbols, and a blue button beside each candidate’s name. When a voter presses the button next to their chosen candidate, the vote is recorded in the control unit. The ballot unit is placed inside the voting compartment so that the voter can vote

secretly and securely. It is connected to the control unit by a cable and works only when the presiding officer activates it. Thus, the ballot unit provides a simple, reliable, and user-friendly method for recording votes electronically.

G. BUZZER

A buzzer is an electronic sound-producing component used to provide audio indication or alert signals in an embedded system. In this biometric voting system, the buzzer is connected to the ESP32 and is used to indicate different system operations such as successful authentication, invalid voter detection, fingerprint mismatch, or vote confirmation. It produces a beep sound whenever a specific action occurs, helping the user understand the status of the system without only depending on the display. Thus, the buzzer improves user interaction, warning indication, and system feedback in the voting setup.

V. CIRCUIT DAIGRAM



The ESP32 Wi-Fi module is the main controller in this biometric voting system and is connected to all the other components in the circuit. It controls the overall working of the system by receiving inputs from the keypad, RFID reader, fingerprint sensor, and candidate buttons, and then processing them accordingly. It also sends output signals to the LCD display and buzzer to provide user instructions and system alerts.

The RFID reader and R305 fingerprint sensor are used for voter authentication and security. First, the RFID module reads the voter's card details, and then the fingerprint sensor verifies the voter's fingerprint with the stored database. This two-step verification helps ensure that only authorized and registered voters are allowed to vote, thereby improving the reliability and security of the voting system.

The LCD display, keypad, candidate buttons, and buzzer are used for user interaction and voting operation. The LCD shows messages such as voter instructions, authentication status, and vote confirmation. The keypad is mainly used for entering passwords or administrative functions, while the candidate buttons allow the voter to select and cast a vote for the desired candidate. The buzzer provides sound indications for successful actions or warning messages during the voting process.

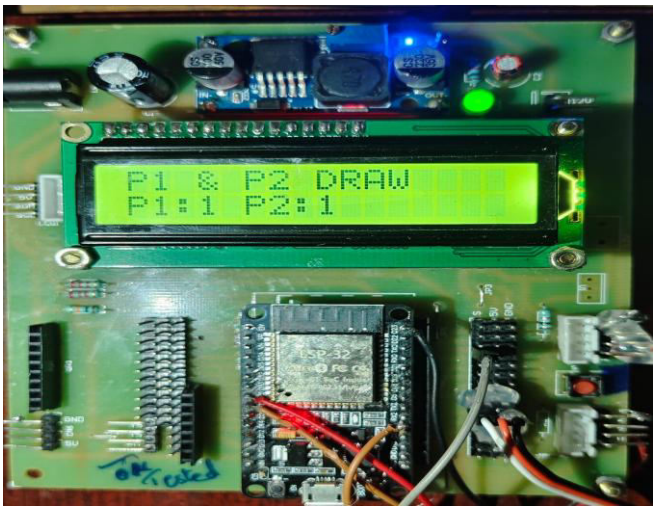
VI. RESULT AND DISCUSSION

The developed biometric electronic voting system was successfully implemented and tested using the ESP32 controller, fingerprint sensor, LCD display, candidate buttons, and buzzer. The system was able to perform the required voting operations such as voter authentication, candidate selection, and vote recording in an efficient manner. The LCD displayed the voting instructions and status messages clearly, while the buzzer provided audio indication for different operations. The overall hardware setup worked properly and demonstrated the basic functions of a secure electronic voting machine.

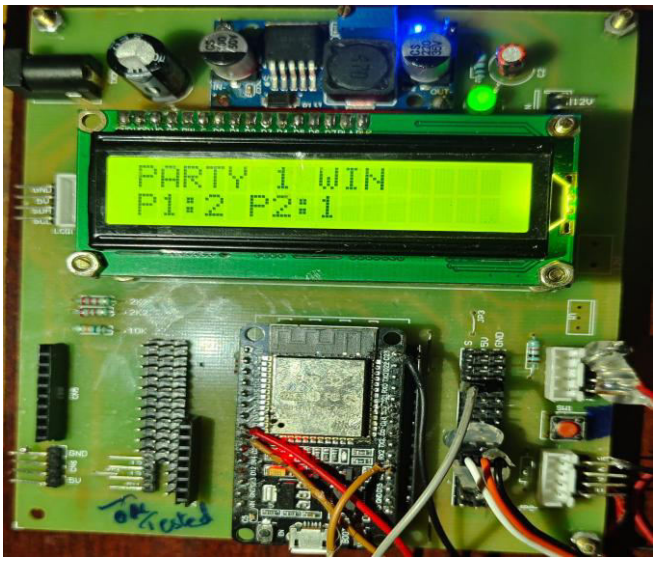
During testing, the fingerprint authentication module accurately identified registered users and restricted access to unauthorized persons. After successful verification, the voter was allowed to cast a vote by pressing one of the candidate buttons. The vote was then recorded successfully, and the system confirmed the action through the display and buzzer indication. This shows that the system is capable of reducing fake voting and ensuring that only valid users can participate in the election process.

From the discussion, it is observed that the proposed system provides a simple, low-cost, secure, and user-friendly voting solution suitable for smart election applications. The integration of biometric verification with electronic vote casting improves transparency and reliability compared to conventional voting methods. Although the prototype performs effectively, the system can be further enhanced by adding cloud data storage, voter database management, real-time result monitoring, and higher security features for practical large-scale implementation.

In addition, the system showed good coordination between both the hardware and software sections, ensuring smooth operation throughout the voting process. The response time of the fingerprint verification and button selection was satisfactory, making the system convenient for practical use. The prototype proves that the use of IoT and biometric technology in voting applications can improve efficiency, reduce manual errors, and provide a more trustworthy election process.



The result shown in the image indicates a draw condition in the electronic voting system, where both Party 1 (P1) and Party 2 (P2) have received an equal number of votes. The LCD display shows the message "P1 & P2 DRAW" along with the vote count P1:1 and P2:1, confirming that each party has secured one vote. This output proves that the ESP32-based voting system is able to accurately count votes, compare the total results, and correctly identify a tie situation. It also demonstrates that the system can display the final result clearly and automatically, making the voting process more reliable, simple, and error-free.



This image shows a compact electronic voting machine prototype built around an ESP32, which acts as the central controller to record and process votes. The system includes a regulated power supply section at the top, indicator LEDs for status, and a 16×2 LCD that displays real-time results—here indicating "PARTY 1 WIN" with vote counts P1:2 and P2:1. Push buttons or input connectors are used for casting votes, and each press updates the count stored in the microcontroller. Overall, it demonstrates a simple, efficient, and reliable digital voting system suitable for small-scale or educational applications.

REFERENCE

1. 1. S. Barra, A. Casanova, M. Fraschini, and M. Nappi, "Fusion of physiological measures for multimodal biometric systems," Multimedia Tools Appl., vol. 76, no. 4, pp. 4835–4847, Feb. 2017.

2. M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A new multi-level approach to EEG based human authentication using eye blinking," Pattern Recognit. Lett., vol. 82, pp. 216–225, Oct. 2016.

3. F. S. Liwen, X. A. Cai, and J. Ma, "A dual-biometric-modality identification system based on fingerprint and EEG," in Proc. 4th IEEE Int. Conf. Biometrics, Theory, Appl. Syst., Sep. 2010, pp. 1–6.

4. C.McCool,S.Marcel,A.Hadid,M.Pietikainen, P. Matejka, J. Cernock ¨ y, N. Poh, J. ´ Kittler, A. Larcher, C. Levy, D. Matrouf, J.-F. Bonastre, P. Tresadern, and T. ´ Cootes, "Bi-modal person recognition on a mobile phone: Using mobile phone data," in Proc. IEEE Int. Conf. Multimedia Expo. Workshops, Jul. 2012, pp. 635–640.

5. M. Wang, H. A. Abbass, and J. Hu, "Continuous authentication using EEG and face images for trusted autonomous systems," in Proc. 14th Annu. Conf. Privacy, Secur. Trust (PST), Dec. 2016, pp. 368–375

6. T. Sim, S. Zhang, R. Janakiraman, and S. Kumar, "Continuous verification using multimodal biometrics," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 4, pp. 687–700, Apr. 2007.Mohammed Abdur Rahman, R. Das, and K. Priya, "Baggage Tracking Using IoT," IEEE Access Journal, 2025.

7. S. Yazji, X. Chen, R. P. Dick, and P. Scheuermann, "Implicit user re authentication for mobile devices," in Proc. Int. Conf. Ubiquitous Intell. Comput., 2009,

8. J. J. C. Ying, Y. J. Chang, C. M. Huang, and V. S. Tseng, "Demographic prediction based on users mobile behaviors," Mobile Data Challenge, vol. 2, no.

9. pp. 1–4, 2012. 9. A. Crivellari and E. Beinat, "From motion activity to geo-embeddings: Generating and exploring vector representations of locations, traces and visitors through large-scale mobility data," ISPRS Int. J. Geo-Inf., vol. 8, no. 3, p. 134, Mar. 2019.

10. T.J.Neal,D.L.Woodard,andA.D.Striegel,"Mobilede viceapplication, Bluetooth, and Wi-Fi usage data as behavioral biometric traits," in Proc. IEEE 7th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS).

DESIGN AND IMPLEMENTATION OF SMART BAG TRACKING AND SAFETY SYSTEM FOR KIDS

RIKAS M
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
rikasriku7102@gmail.com

BABHITHA ROSH B
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
babhitha.rosh@gmail.com

GOWTHAM S
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
gowthamnest@gmail.com

HRIDYA R
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
hridyaradhakrishan@gmail.com

Mr.A.MANICKAVASAGAM
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
manick1.a@gmail.com

***Abstract*—** This project presents the design and implementation of a Smart Bag Tracking and Safety System for kids using Internet of Things (IoT) technology. The system is developed using an ESP32 microcontroller integrated with a Neo-6M GPS module, SIM800L GSM module, temperature sensor, and an SOS emergency switch. It provides real-time location tracking and sends SMS alerts with a Google Maps location link to predefined contacts during emergency situations. The device is compact, low cost, and powered by a rechargeable 3.7 V battery with a buck converter. The proposed system offers an effective solution for improving student safety and basic health monitoring.

I. INTRODUCTION

In today's rapidly evolving technological environment, ensuring the safety and security of school students has become a significant societal concern. Increasing incidents of accidents, health emergencies, missing children, and other safety threats have created an urgent need for reliable monitoring and emergency response systems. Parents and educational institutions require efficient solutions that can provide real-time information about a child's location and condition, especially during critical situations. Traditional safety measures and manual monitoring systems often lack real-time communication capabilities and accurate location tracking features. In many cases, delays in information transmission can lead to serious consequences. Therefore, integrating smart embedded technologies into daily-use objects has emerged as an innovative and practical approach to enhancing personal safety. The advancement of the Internet of Things (IoT) has enabled the development of intelligent systems capable of sensing, processing, and transmitting data instantly. IoT-based systems allow interconnected devices to communicate over wireless networks, enabling real-time monitoring and rapid alert mechanisms. By combining embedded systems, wireless communication modules, and sensors, it is possible to design compact and portable safety devices suitable for everyday applications. This project presents the design and implementation of a Smart Bag Tracking and Safety System specifically developed for school students. The system is integrated into a school bag as a compact electronic module capable of providing real-time location tracking, emergency alert communication, and basic health monitoring. The main objective of this system is to ensure immediate assistance during emergencies by automatically notifying predefined contacts such as parents or guardians. The system is built around the ESP32 microcontroller, which functions as the central control unit. A GPS module is used to continuously acquire real-time geographical coordinates, while a GSM module enables SMS-based communication. In case of an emergency, activation of the SOS switch triggers an alert message containing the student's location in the form of a Google Maps link. Additionally, a temperature sensor is incorporated to monitor body temperature and generate alerts if abnormal readings are detected. The device operates on a rechargeable battery and includes a voltage regulation mechanism to ensure stable performance. The overall design emphasizes portability, energy efficiency, cost-effectiveness, and reliability. By using GSM-based communication instead of internet-dependent applications, the system ensures functionality even in areas with limited internet connectivity. In summary, this project demonstrates the effective integration of IoT technology, embedded systems, and wireless communication to create a practical and scalable student safety solution. The proposed system addresses real world safety challenges and provides a foundation for further development in applications such as women's safety devices, elderly monitoring systems, and emergency alert systems.

II. LITERATURE REVIEW

1. S. Karthick, M. Rajesh, and V. Divya – "Smart Luggage Tracking and Alert System Using Arduino"

Karthick et al. proposed a smart luggage tracking and alert system based on the Arduino platform. The main objective of the system was to prevent luggage loss during travel by continuously monitoring the bag's location. The system used sensors and wireless communication to alert the user whenever the luggage moved beyond a predefined distance. The Arduino microcontroller was used to control and process all operations in the system. This approach demonstrated that low-cost embedded systems could effectively be used for luggage monitoring and security applications.

2. Yuvaraj S. A., Monisha P., and Kavin S. – "Smart Bag Tracking and Alert System Using RFID and GSM"

Yuvaraj et al. developed a smart bag tracking system using RFID technology for identification and GSM for communication. In their system, RFID tags were attached to the luggage, and RFID readers detected the bag's presence within a particular area. If the luggage moved away from the designated location, the system sent an alert message to the user through GSM communication. This method improved luggage security but had limitations in long-distance tracking since RFID operates only within a limited range.

3. Anuradha S. and Prakash M. – "IoT-Based Child Safety Wearable Using GPS and GSM"

Anuradha and Prakash introduced an IoT-based child safety wearable device that uses GPS and GSM modules to monitor the real-time location of children. The system continuously tracked the user's location and transmitted the information to guardians through mobile communication networks. In emergency situations, the device could send alerts along with the exact GPS coordinates. This study highlighted the importance of combining GPS and GSM technologies for reliable real-time tracking and safety monitoring.

4. Deepika R., Kavitha K., and Soundarya B. – "IoT-Based Women Safety Device Using GPS, GSM and Sensors"

Deepika et al. proposed a safety device designed for women using GPS, GSM, and different sensors. The device was capable of detecting emergency situations and immediately sending alert messages along with location details to predefined contacts. The system demonstrated how IoTbased monitoring devices can provide safety and security in critical situations. The integration of sensors and communication modules in this research provides a useful concept for designing intelligent tracking systems.

5. Meenakshi S., Haritha N., and Rohit A. – "AI-Powered Vehicle Tracking and Route Deviation Detection"

Meenakshi et al. presented an intelligent vehicle tracking system that uses artificial intelligence to monitor vehicle movement and detect route deviations. The system analyzes location data and identifies abnormal movement patterns using intelligent algorithms. The study shows how advanced technologies like AI can enhance tracking systems by improving accuracy and providing predictive monitoring capabilities.

III. SYSTEM OVERVIEW

The system overview of the proposed Smart Bag Tracking and Safety System explains the overall operation and integration of the hardware components used to ensure child safety. The system is built around an ESP32 microcontroller, which acts as the central controller responsible for processing sensor data and coordinating communication between different modules. A Neo-6M GPS module is used to obtain real-time geographical coordinates of the bag by receiving signals from GPS satellites. These coordinates are transmitted to the ESP32 through serial communication. The system also includes a SIM800L GSM module, which enables the device to send SMS alerts containing the child's location to predefined contacts such as parents or guardians.

The system continuously monitors the location of the smart bag during operation. When the SOS emergency switch is pressed, the ESP32 immediately retrieves the latest GPS coordinates and sends an alert message through the GSM module containing a Google Maps location link. This allows guardians to instantly identify the child's exact location. Additionally, a temperature sensor is included to monitor environmental conditions around the child. The entire system is powered by a rechargeable 3.7 V battery, and a buck converter regulates the voltage to ensure stable operation of all electronic components. The compact design allows the device to be easily integrated into a school bag without affecting usability.

IV. PROPOSED SYSTEM

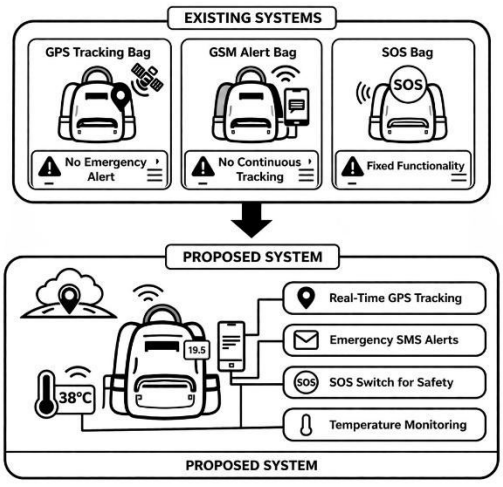


Fig 01: Proposed System

In today's rapidly advancing technological world, ensuring the safety and well-being of school students has become a major concern. Increasing cases of accidents, health emergencies, and missing children highlight the need for a reliable and intelligent monitoring system. Traditional safety methods lack real-time tracking and immediate communication, making them less effective during critical

situations. With the development of the Internet of Things (IoT), it has become possible to design smart systems that can monitor, process, and transmit data in real time. By integrating embedded systems with wireless communication technologies, compact and efficient safety solutions can be developed for everyday use.

This project presents a Smart Bag Tracking and Safety System designed specifically for school students. The system is integrated into a school bag and provides real-time location tracking, emergency alert communication, and basic health monitoring. It ensures that parents or guardians can quickly respond during emergencies by receiving instant alerts along with the student's location. The system uses an ESP32 microcontroller as the core unit, along with GPS and GSM modules for tracking and communication. Additional features such as an SOS switch and temperature sensor enhance safety by enabling manual and automatic alert generation. The overall system is designed to be portable, cost-effective, and reliable for real-world applications.

The proposed Smart Bag Tracking and Safety System is a compact IoT-based solution designed to ensure the safety of school students by integrating real-time location tracking, emergency alert communication, and basic health monitoring into a school bag. The system is built around an ESP32 microcontroller, which acts as the central control unit and coordinates the functioning of all components. A Neo-6M GPS module continuously acquires the student's geographical location, while a SIM800L GSM module enables communication by sending SMS alerts to predefined contacts such as parents or guardians. In emergency situations, pressing the SOS button triggers an immediate alert message containing a Google Maps link of the current location. Additionally, a temperature sensor monitors the student's body temperature and automatically sends alerts if abnormal readings are detected. The system operates on a rechargeable battery with proper voltage regulation and uses UART communication between modules. Overall, the proposed system provides a reliable, cost-effective, and portable safety solution that works even without internet connectivity, making it suitable for real-world applications.

Furthermore, the system is designed with low power consumption to ensure longer operational time between charges. Its compact and lightweight structure allows seamless integration into school bags without affecting usability. The use of widely available components makes the system economical and easy to maintain. It can also be enhanced in the future by integrating a mobile application for real-time tracking and notifications. Thus, the system offers a scalable and efficient solution for improving student safety and security. In addition, the system ensures quick emergency response through instant alert notifications. It is designed to function reliably even in areas with limited connectivity using GSM communication. The modular architecture allows easy upgrades and integration of additional sensors if required. It also provides continuous monitoring, which enhances overall system effectiveness.

V. BLOCK DIAGRAM

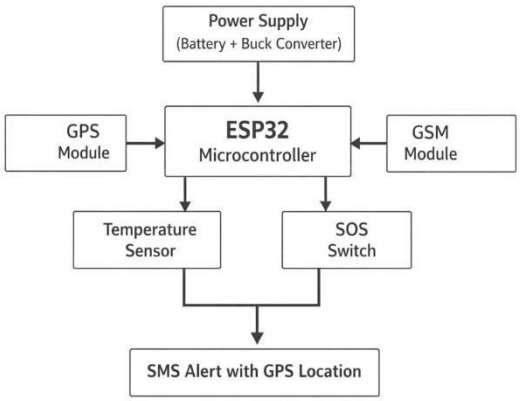


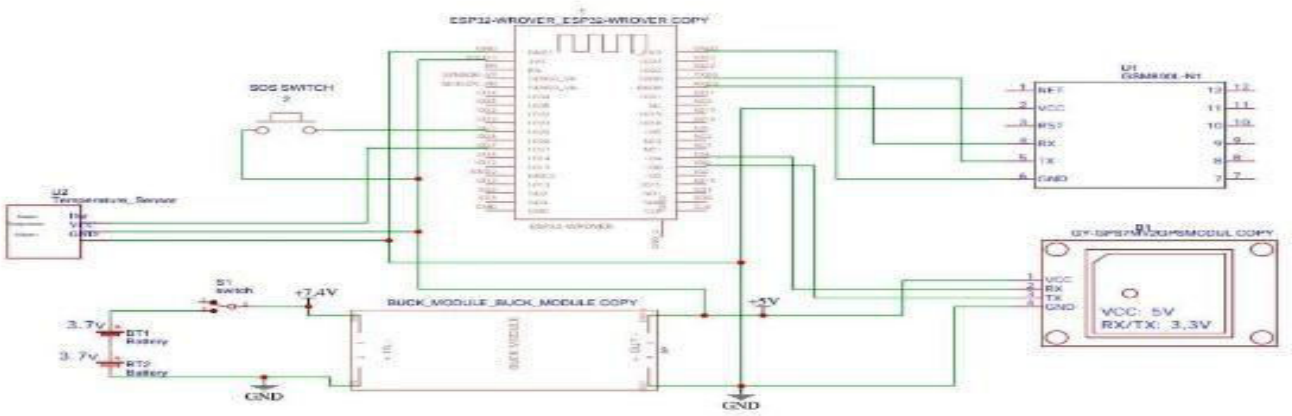
Fig 02: Block Diagram

The block diagram represents the working of the Smart Bag Tracking and Safety System, where the ESP32 microcontroller acts as the central processing unit. The system is powered by a battery combined with a buck converter to provide a stable voltage supply to all components. The GPS module is connected to the ESP32 to continuously provide real-time location data in the form of latitude and longitude coordinates. The GSM module is also interfaced with the ESP32 to enable communication by sending SMS alerts to predefined contacts. A temperature sensor is used to monitor the student's body temperature, while an SOS switch allows the user to manually trigger an emergency alert. The ESP32 continuously processes inputs from the GPS module, temperature sensor, and SOS switch, and whenever an emergency condition is detected—such as pressing the SOS button or abnormal temperature—it activates the GSM module to send an SMS alert containing the current GPS location. This integrated system ensures real-time monitoring, quick response, and enhanced safety for students.

Additionally, the system is designed to operate with low power consumption, ensuring longer battery life for continuous usage. The compact and lightweight design makes it suitable for easy integration into school bags without causing inconvenience. The use of wireless communication enhances reliability and allows remote monitoring by parents or guardians. The system can also be further upgraded by integrating a mobile application for live tracking and notifications. Overall, this smart solution provides a cost-effective, efficient, and user-friendly approach to improving student safety and security.

It also ensures quick emergency response through instant alert mechanisms, reducing potential risks. Moreover, the system's modular design allows easy customization and future expansion based on user requirements. The system is highly reliable in real-time applications due to its continuous monitoring capability. It also enhances peace of mind for parents by providing constant updates on the student's safety.

VI. CIRCUIT DAIGRAM



The overall circuit diagram represents a Smart Bag Tracking and Safety System designed to monitor location and provide emergency alerts. The system is built around the ESP32WROVER, which functions as the main controller that manages all connected components. A GY-GPS7MV2 GPS Module is connected to the ESP32 through serial communication pins to obtain real-time geographical coordinates such as latitude and longitude from satellites. For communication purposes, a SIM800L GSM Module is interfaced with the ESP32, allowing the system to send SMS alerts containing the location information to a predefined mobile number. A temperature sensor is also connected to the ESP32 to monitor the surrounding temperature inside the bag. Additionally, an SOS push button is interfaced with a digital input pin of the ESP32, enabling the user to manually trigger an emergency alert.

The power supply for the system is provided by two 3.7 V rechargeable batteries, which together produce approximately 7.4 V. This voltage is passed through a buck converter module that steps down the voltage to a regulated 5 V supply, ensuring safe and stable operation of the electronic components. The regulated power is then distributed to the ESP32, GPS module, GSM module, and the sensor circuit.

During operation, the ESP32 continuously receives location data from the GPS module and monitors the temperature sensor. If the SOS switch is pressed, the ESP32 immediately reads the latest GPS coordinates and sends them through the GSM module as an SMS alert to a registered mobile number. This allows guardians or authorities to track the location of the bag in real time during emergencies. Thus, the integration of the ESP32 controller, GPS tracking, GSM communication, and sensor monitoring creates a compact and efficient safety system for tracking and emergency notification.

The system ensures real-time response with minimal delay. Its compact design allows easy portability and also enhances safety through continuous monitoring and instant alerts.

VII. SYSTEM COMPONENTS

Fig 04: System Components

The system consists of several essential components integrated into a compact setup. A push button switch is used as an SOS trigger for emergency situations. A temperature sensor is included to continuously monitor environmental conditions. The power supply is provided by rechargeable lithium-ion batteries to ensure portability and uninterrupted operation. The core controller is the ESP32 microcontroller, which processes data and manages communication. A GPS module is used to obtain real-time location information, enabling tracking and alert functionality. All components are embedded within a backpack, with proper wiring and placement to ensure user comfort, reliability, and efficient system performance.

RESULT AND DISCUSSION

The implemented system successfully provided real-time location tracking using the Neo-6M GPS module integrated with the ESP32 microcontroller. During testing, the GPS module was able to determine the device position with an approximate accuracy of 5–10 meters under normal outdoor conditions. The obtained coordinates were correctly converted into a Google Maps link and transmitted to predefined contacts. The results indicate that the system provides reliable location monitoring suitable for student safety applications.

The SOS emergency mechanism was tested to evaluate the responsiveness of the alert system. When the SOS switch was activated, the ESP32 immediately triggered the GSM module to send an SMS alert containing the GPS location link. The message delivery time was observed to be approximately 5–8 seconds, depending on network conditions. This rapid alert mechanism ensures that guardians can quickly identify the child's location during emergency situations.

The SIM800L GSM module demonstrated stable communication performance throughout the experimental testing phase. The module consistently transmitted SMS alerts without significant transmission failures when adequate cellular signal strength was available. The use of GSM communication ensures compatibility with widely available mobile networks without requiring internet connectivity. This makes the system practical for deployment in both urban and rural environments.

The device was powered using a 3.7 V rechargeable lithium battery, regulated through a buck converter to ensure stable voltage supply. Power consumption measurements indicated that the system operates efficiently with minimal energy usage during continuous monitoring. The ESP32 microcontroller maintained stable performance while simultaneously handling GPS data acquisition and GSM communication. The results confirm that the system can operate for extended periods on battery power.

The integrated IoT-based smart bag system demonstrated reliable performance during prototype testing. All hardware components, including the ESP32, GPS module, GSM module, and sensors, operated cohesively without system failures. The compact design allows the device to be easily embedded inside a school bag without affecting portability. Overall, the system provides a cost-effective and practical solution for enhancing child safety and monitoring.

The system was able to deliver accurate real-time location data during testing. Emergency alerts were successfully transmitted without noticeable delay. The temperature sensor responded effectively to abnormal conditions. Power consumption was optimized to support longer usage with the available battery. These results indicate that the system is reliable and suitable for real-world implementation.

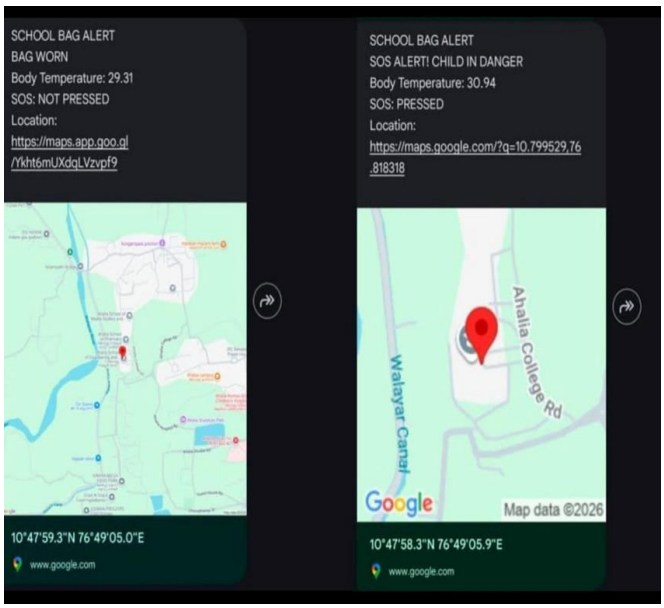


Fig 05: Location Tracking

REFERENCES

1. S. Karthick, M. Rajesh, and V. Divya, “Smart Luggage Tracking and Alert System Using Arduino”, International Journal of Engineering Research Technology (IJERT), Vol. 9, Issue 4, 2020.
2. Yuvaraj S. A., Monisha P., and Kavin S., “Smart Bag Tracking and Alert System Using RFID and GSM,” International Research Journal of Engineering and Technology (IRJET), Vol. 7, Issue 6, 2022.
3. Anuradha S. and Prakash M., “IoT-Based Child Safety Wearable Using GPS and GSM,” International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), Vol. 10, Issue 3, 2021.
4. Deepika R., Kavitha K., and Soundarya B., “IoT-Based Women Safety Device Using GPS, GSM and Sensors,” International Journal of Engineering Trends and Technology (IJETT), Vol. 71, Issue 3, 2023.
5. Meenakshi S., Haritha N., and Rohit A., “AI-Powered Vehicle Tracking and Route Deviation Detection,” in Proceedings of the International Conference on Smart Computing and Communications, 2024.
6. Shrivastava P., Ahmed S., and Kiran R., “IoT-Based Health Monitoring System for Patients,” Journal of Intelligent Systems, Vol. 8, Issue 5, pp. 1123–1132, 2020.
7. Mohammed Abdur Rahman, R. Das, and K. Priya, “Baggage Tracking Using IoT,” IEEE Access Journal, 2025.
8. Arduino Documentation and Reference, “Arduino Official Website,” [Online]. Available: https://www.arduino.cc
9. u-blox AG, “NEO-6M GPS Module Datasheet,” Switzerland, 2023.
10. SIMCom Wireless Solutions, “SIM800L GSM Module User Manual,” Shenzhen, China, 2022.

DESIGN OF SMART OXYGEN CONCENTRATOR WITH REDUCED GRID RELIANCE

AJITH M
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
ajithpkd2001@gmail.com

ANIRUDH REJI
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
anirudhreji69@gmail.com

FASIRALI
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
fasirali888@gmail.com

SREYA P
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
sreya4738@gmail.com

Ms. LAKSHMY SURESH
Asst. Professor
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala, India
lakshmy.suresh@ahalia.ac.in

***Abstract*—The increasing demand for reliable medical oxygen supply, especially during emergencies and in remote areas, highlights the need for energy-efficient and selfsustaining oxygen generation systems. This paper presents the design of a smart oxygen concentrator with reduced grid reliance, aimed at ensuring uninterrupted oxygen delivery while minimizing dependence on conventional electrical power. The proposed system integrates renewable energy sources, such as solar photovoltaic panels, with an intelligent energy management unit to power the oxygen concentrator. A microcontroller-based monitoring system continuously observes oxygen purity, flow rate, battery status, and power availability, enabling automatic switching between solar power, battery storage, and grid supply when required. The system also incorporates IoT-based monitoring for real-time data tracking and alerts to ensure safe operation and timely maintenance. By optimizing energy consumption and utilizing renewable energy resources, the proposed design improves reliability, reduces operational costs, and enhances accessibility of oxygen therapy in rural and off-grid healthcare facilities. The implementation of this smart oxygen concentrator contributes to sustainable healthcare infrastructure while ensuring continuous oxygen supply during power interruptions.**

I. INTRODUCTION

Medical oxygen plays a critical role in healthcare systems for the treatment of respiratory diseases, emergency care, and critical medical conditions. Oxygen concentrators are widely used devices that extract oxygen from ambient air and supply concentrated oxygen to patients. However, conventional oxygen concentrators depend heavily on continuous electrical power from the grid. In many rural and remote areas, frequent power outages and unstable electricity supply can interrupt the operation of these life supporting devices, posing serious risks to patient safety..The global health challenges experienced during the COVID-19 Pandemic highlighted the urgent need for reliable and continuous oxygen supply systems. Hospitals and healthcare center in many regions faced shortages of medical oxygen and reliable power sources. This situation emphasized the importance of developing oxygen generation systems that can operate with minimal dependence on conventional grid electricity.To address this challenge, the design of a smart oxygen concentrator with reduced grid reliance is proposed. The system integrates renewable energy sources such as solar photovoltaic (PV) systems with battery storage to provide an alternative and sustainable power supply. A smart control unit monitors system parameters such as oxygen purity, flow rate, battery level, and power source availability. Based on these parameters, the system can intelligently switch between solar power, battery storage, and grid supply to ensure uninterrupted operation.

In addition, modern technologies such as microcontroller based monitoring and IoT connectivity enable real-time data monitoring, remote diagnostics, and improved system reliability. By combining renewable energy utilization with intelligent power management, the proposed system aims to reduce dependency on grid electricity while ensuring continuous and efficient oxygen supply.This approach not only enhances the reliability of oxygen therapy devices but also supports sustainable healthcare infrastructure, particularly in rural healthcare centers, emergency facilities, and disaster-affected areas where stable electricity supply is limited.

II. Ease of Use

The proposed smart oxygen concentrator is designed with a strong focus on simplicity, user-friendliness, and minimal technical complexity, making it suitable for use in hospitals, clinics, and home healthcare environments. The system includes an intuitive control interface that allows healthcare providers or caregivers to easily monitor and adjust oxygen flow rate and system parameters. Clear indicators and a digital display provide real-time information about oxygen purity, battery level, power source, and system status Automatic power source switching is incorporated to improve ease of operation. The system intelligently switches between solar power, battery backup, and grid supply without requiring manual intervention. This ensures continuous oxygen delivery even during power outages, reducing the burden on medical staff and caregivers.

In addition, the concentrator is equipped with smart alerts and alarms that notify users about low oxygen purity, system faults, or low battery levels. These alerts help users take timely action and maintain safe operation. The integration of remote monitoring features also enables healthcare professionals to check system performance and receive notifications through connected devices, improving operational convenience.The design also emphasizes low maintenance requirements. Modular components and easily replaceable filters simplify routine servicing, allowing quick maintenance without specialized technical expertise. Compact design, lightweight structure, and simple installation further enhance usability, making the device suitable for both hospital settings and remote healthcare centers.Overall, the smart oxygen concentrator ensures reliable oxygen delivery with easy operation, automatic power management, and user-friendly monitoring, making it an effective solution for healthcare facilities with limited or unstable grid power supply.

III. Related Work

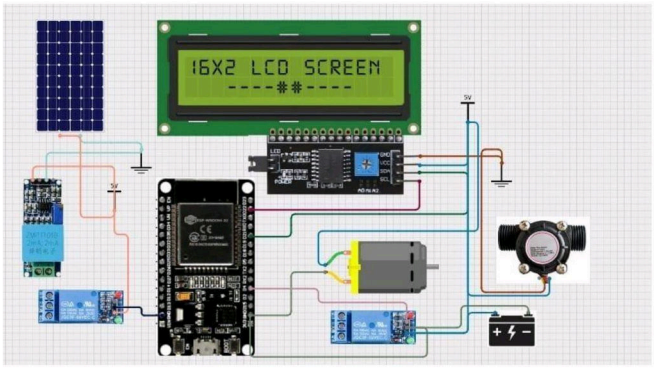
Oxygen concentrators have become essential medical devices, especially after the global impact of the COVID-19 pandemic, which highlighted the urgent need for reliable and continuous oxygen supply systems. Several research efforts have focused on improving efficiency, portability, and energy management of oxygen concentrators. Traditional oxygen concentrators primarily use Pressure Swing Adsorption (PSA) technology, which separates oxygen from atmospheric air using zeolite materials. Studies have shown that while PSA-based systems are effective, they are heavily dependent on uninterrupted grid power, limiting their usability in rural or power-deficient areas.

Recent work has explored solar-powered oxygen concentrators to address energy dependency. Researchers have proposed photovoltaic (PV) integrated systems where solar panels supply power directly to the concentrator or charge battery storage units. These systems significantly reduce reliance on grid electricity; however, challenges such as intermittent solar availability and energy storage inefficiencies remain. Advancements in battery management systems (BMS) and energy storage technologies have further improved system reliability. Lithium-ion batteries are widely adopted due to their high energy density and longer lifecycle. Some studies demonstrate hybrid energy systems combining solar, battery, and grid supply to ensure uninterrupted operation.

In addition, the integration of Internet of Things (IoT) technologies has enabled smart monitoring and control of oxygen concentrators. IoT-based systems allow real-time tracking of oxygen purity, flow rate, power consumption, and fault detection. Remote monitoring improves maintenance efficiency and ensures patient safety, particularly in home-care environments. Research has also been conducted on energy-efficient compressors and control algorithms to optimize power consumption. Variable speed compressors and intelligent control strategies help reduce energy usage while maintaining consistent oxygen output.

Despite these advancements, most existing systems lack a fully integrated solution that combines renewable energy utilization, smart monitoring, and adaptive power management in a cost-effective manner. Therefore, this work proposes a smart oxygen concentrator system designed to minimize grid reliance by incorporating renewable energy sources, efficient energy storage, and intelligent control mechanisms.

IV. Hardware Design

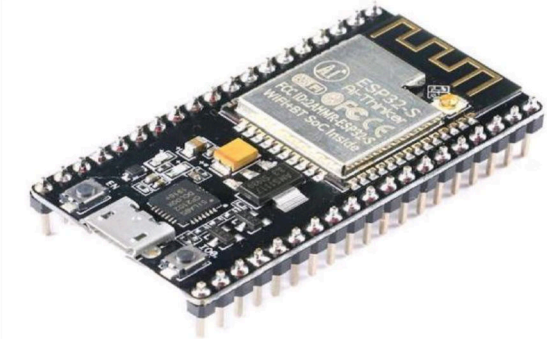


A. LCD SCREEN

LCD (Liquid Crystal Display) screen is an electronic display module and find a wide range of applications. A 16x2 LCD display is very basic module and is very commonly used in various devices and circuits. These modules are preferred over seven segments and other multi segment LEDs. The reasons being: LCDs are economical; easily programmable; have no limitation of displaying special & even custom characters (unlike in seven segments), animations and so on.

B. MICROCONTROLLER UNIT

The ESP32 is a low-cost, low-power, system-on-chip (SoC) series developed by Express if Systems, based on the Xtensa® 32-bit LX6 microprocessor. It is widely used in embedded and IoT applications due to its integrated Wi-Fi and Bluetooth capabilities, rich peripheral set, and flexible power management



C. LITHIUM ION BATTERY

The 3V 2200mAh rechargeable lithium-ion battery is designed for portable electronics, IoT devices, medical equipment, and backup power applications. It offers high energy density, long cycle life, and low self-discharge compared to traditional chemistries.

D. FLOW SENSOR

A flow sensor (or flow meter) measures the volumetric or mass flow rate of liquids or gases.
Common types include Turbine, Ultrasonic, Electromagnetic, Thermal Mass, Coriolis, Vortex, and Hall effect flow sensors.

E. RELAYS

A relay is an electrically operated switch. Current flowing through the coil of the relay creates a magnetic field which attracts a lever and changes the switch contacts. The coil current can be on or off. So relays have two switch positions and they are double throw (changeover) switches. Relays allow one circuit to switch a second circuit which can be completely separate from the first. The link is magnetic and mechanical. The coil of a relay passes a relatively large current, typically 30mA for a 12V relay, but it can be as much as 100mA for relays designed to operate from lower voltages. Most ICs (chips) cannot provide this current and a transistor is usually used to amplify the small IC current to the larger value required for the relay coil. The maximum output current for the popular 555 timer IC is 200mA so these devices can supply relay coils directly without amplification.

F. VOLTAGE SENSOR

The Arduino analog input is limited to a 5 VDC input. If you wish to measure higher voltages, you will need to resort to another means. One way is to use a voltage divider.

G. DC MOTOR

The dc generators and dc motors have the same general construction.
motor principle:
an electric motor is a machine which converts electrical energy to mechanical energy.

all dc machines have five principal components viz (i) field system (ii) armature core (iii) armature winding (iv) commutator (v) brushes

H. SOLAR PANNEL

A solar panel is a device that collects and converts solar energy into electricity or heat. It known as Photovoltaic panels, used to generate electricity directly from sunlight Solar thermal energy collection system used to generate electricity through a system of mirrors and fluid-filled tubes solar thermal collector, used to generate heat solar hot water panel used to heat water. It is energy portal.

A solar power technology uses solar cells or solar photovoltaic arrays to convert light from the sun directly into-electricity.Photovoltaics, is in which light is converted into electrical power.It is best known as a method for generating solar power by using solar cells-packaged in photovoltaic modules, often electrically connected in multiplesas solar photovoltaic arrays to convert energy from the sun into electricity.

The photovoltaic solar panel is photons from sunlight knock electrons into a higher state of energy, creating electricity.

V.SOFTWARE ARCHITETURE

The software architecture follows a modular design approach to enhance maintainability and scalability. The firmware running on the ESP32 is responsible for sensor data acquisition, signal conditioning, control logic execution, alarm handling, display updates, and wireless communication. Sensor readings are periodically sampled and filtered to reduce noise. The control algorithm evaluates sensor data against predefined thresholds to determine appropriate actions. When abnormal conditions are detected, the system triggers alarms and connectivity enables data transmission to a remote monitoring interface for centralized supervision.

A. ARDUINO IDE

1) Introduction

Arduino IDE stands for Integrated Development Environment. It is an open-source official software provided by Arduino.cc used for writing, compiling, and uploading programs to Arduino boards. The IDE supports C and C++ programming languages and is easy to use even for beginners. It is compatible with operating systems like Windows, macOS, and Linux. The software works with various Arduino boards such as Arduino Uno, Mega, and Leonardo. It consists mainly of an editor for writing code and a compiler for compiling and uploading it to the microcontroller.

2) Connect the Arduino Board

- Plug in the Arduino using the USB cable
- Open Arduino IDE
- Go to Tools > Board and select your board model (e.g., Arduino Uno)
- Go to Tools > Port and select the correct COM port (e.g., COM3 on Windows)

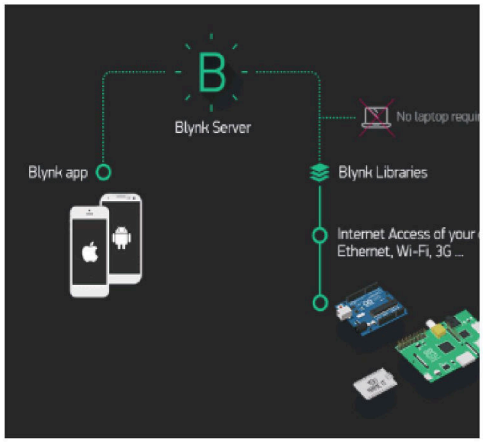
3) Arduino Software (IDE)

The Arduino IDE includes built-in support for many Arduino boards based on the AVR core, such as Arduino Uno, Nano, Mega, Leonardo, and Micro. It supports boards using microcontrollers like ATmega328P, ATmega32u4, and ATmega2560 running mainly at 16 MHZ The Boards Manager allows users to add support for newer boards based on different cores like Arduino Due, Zero, Edison, and Galileo. These boards provide various numbers of analog inputs, digital I/O pins, and PWM outputs. Features like autoreset and built-in bootloaders make programming easier. This wide compatibility makes Arduino IDE suitable for many embedded applications.

B. BLYNK APPLICATION

1) HOW BLYNK WORK

Blynk is an Internet of Things (IoT) platform used to control hardware remotely and monitor sensor data. It allows users to visualize, store, and manage data through a smartphone app. The platform has three main components: the Blynk App, Blynk Server, and Blynk Libraries. The Blynk App helps create user-friendly interfaces using widgets. The Blynk Server manages communication between the mobile app and hardware devices. Blynk Libraries enable smooth data exchange between the hardware and server in real time.

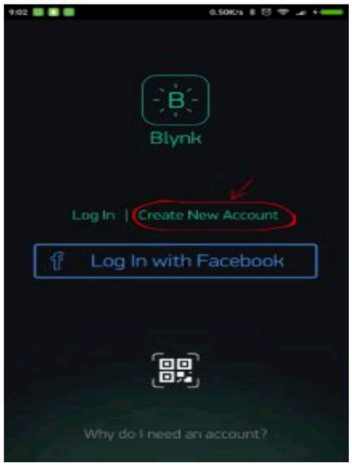


2) FEATURES OF BLYNK APPLICATION

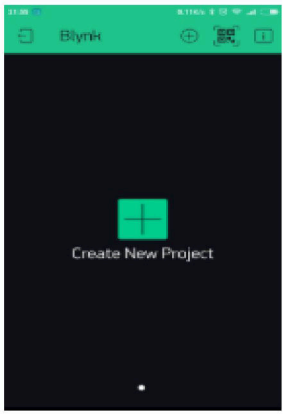
- Provides a common API and user interface for all supported hardware
- Supports cloud connectivity via Wi-Fi, Bluetooth/BLE, Ethernet, USB (Serial), and GSM
- Offers easy-to-use widgets for creating interactive interfaces
- Allows direct pin control without writing code
- Supports virtual pins for adding advanced functionality
- Enables data monitoring, device-to-device communication, and notifications (email, push, etc.)

3) HOW TO USE BLYNK

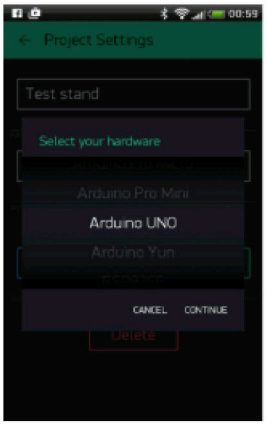
- Auth Token Security: Do not share your Auth Token with anyone, as it provides full access to your hardware. You can send the token to your registered email using the email option or copy it directly by tapping on the Token line.



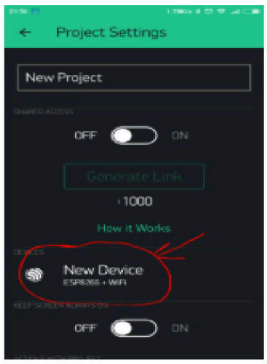
- Project Creation: After verifying the Auth Token and completing the basic setup, press the Create button to create your Blynk project.



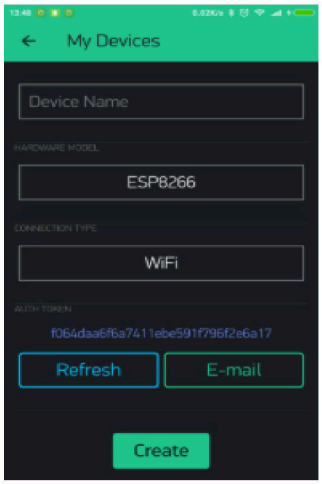
- Adding Widgets: Once the project canvas opens and is empty, tap anywhere on the screen to open the Widget Box, which contains all available widgets.



- Button Widget Selection: Select a Button widget from the widget box to control devices such as an LED.



- Widget Placement: Tap and hold the widget to drag and drop it to any desired position on the project canvas.



- Widget Settings: Each widget has its own configuration options. Tap on the widget to open its settings, where the most important parameter to set is the PIN for hardware control.

ACKNOWLEDGMENT

We express our sincere gratitude to all those who have contributed to the successful completion of this project titled **"Design of Smart Oxygen Concentrator for Reduced Grid Reliance."**

First and foremost, we would like to thank our Head of the Department, faculty members, and our project guide for their valuable guidance, continuous support, and encouragement throughout the development of this work. Their insights and suggestions have been instrumental in shaping the direction and quality of this project.

REFERENCES

1) *Implementation and 8-year follow-up of an uninterrupted oxygen supply system in a hospital in The Gambia-Published 12 March 2016*
2) *Design of Automatic Transfer Switch System Solar Power Plant – PLN-Published 02 December 2023*
3) *Grid connected solar panel with battery energy storage system-Published 01 March 2024*
4) *Energy Monitoring and Control of Automatic Transfer Switch between Grid and Solar Panel for Home System Published 2023*
5) *A Solar PV, BES, Grid and DG Set Based Hybrid Charging Station for Uninterruptible Charging at Minimized Charging Cost - Published September 2018*
6) *Design of Automatic Transfer Switch System Solar Power Plant – PLN Published December 2023*
7) *Energy Monitoring and Control of Automatic Transfer Switch between Grid and Solar Panel for Home System - Published 2023*

SMART HOME ILLUMINATION USING SMART SENSORS AND IOT

NEHA BABU
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala , India
nehababu9655@gmail.com

THEJUS DAS V
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala , India
thejus944@gmail.com

SUDARSH S
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala , India
sudarsh_s@ahalia.ac.in

ARJUN A
Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Kerala , India
arjun.a@ahalia.ac.in

Mr. SUBIN CHACKO
Asst. Professor
Electrical and Electronics Engineering
Ahalia school of Engineering and Technology
Kerala , India
subin.chacko@ahalia.ac.in

Abstract—The project involves the design and implementation of an IoT-based smart home system using the ESP32 microcontroller to optimize energy efficiency and safety. The system integrates an LDR for ambient light detection and a DHT22 for climate monitoring. Key features include automatic lamp dimming via a TRIAC circuit, relay-controlled cooling fans, and remote management through the Blynk IoT platform. For enhanced safety, the design includes a dedicated earth fault detection circuit to protect users from electrical ground leaks.

I. INTRODUCTION (*HEADING 1*)

Energy consumption in residential buildings has significantly increased due to extensive use of lighting and climate control systems. Conventional manual switching systems often result in unnecessary energy wastage when lights and fans remain ON in unoccupied spaces.

With the advancement of IoT technology and low-cost microcontrollers such as ESP32, intelligent automation systems have become practical and affordable. This project proposes an automated smart home system that dynamically adjusts lighting and ventilation based on environmental and occupancy conditions while ensuring electrical safety through fault detection mechanisms.

II. LITERATURE REVIEW

In recent years, smart home automation has become a rapidly growing area of research due to its potential to improve energy efficiency, convenience, and safety in residential and commercial environments. Several studies have investigated the use of IoT, sensor networks, and microcontrollers for intelligent control of lighting and environmental systems.

[8] proposed an IoT-based lighting control system that adjusts illumination based on ambient light intensity and remote user commands. Their approach demonstrated energy savings but lacked integration of occupancy detection or safety features. Similarly, [1] implemented an Arduino-based smart lighting architecture that employed light sensors and mobile app control, focusing primarily on user convenience without environmental monitoring for thermal conditions.

Other researchers have examined climate control using embedded platforms. [3] Developed a thermostat system using temperature sensors and cloud connectivity for real-time monitoring. Although it provided effective thermal regulation, the system operated independently of lighting and presence detection.

More comprehensive frameworks have been attempted; for example, [6] integrated motion detection, light, and temperature sensors into an IoT platform for unified control. While their system combined multiple sensing modalities, it did not include provisions for electrical safety such as earth fault detection.

Despite the advancements, existing work often addresses individual aspects—lighting control, occupancy sensing, or thermal regulation—rather than a unified energy-smart and safety-aware solution. In contrast, the present work integrates smart illumination, occupancy detection, environmental monitoring, remote control via the cloud, and electrical safety in a single modular architecture using the ESP32 microcontroller. This combined approach enables better energy optimization and user protection than previously reported systems.

III. SYSTEM OVERVIEW

The proposed system is designed to provide intelligent control of indoor lighting and climate, driven by environmental conditions and user activity. It consists of integrated hardware modules, sensor interfaces, actuation circuits, and cloud connectivity, all coordinated by the ESP32 microcontroller.

IV. HARDWARE DESIGN

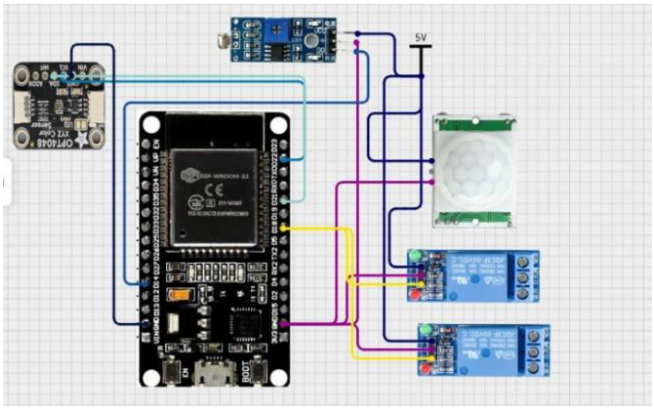


Fig 4.1 Circuit diagram

A. Microcontroller Unit

The ESP32 microcontroller is selected due to its dual-core processing capability, integrated Wi-Fi and Bluetooth modules, low power consumption, and extensive peripheral support. These features make it suitable for real-time data processing and wireless communication in medical applications.

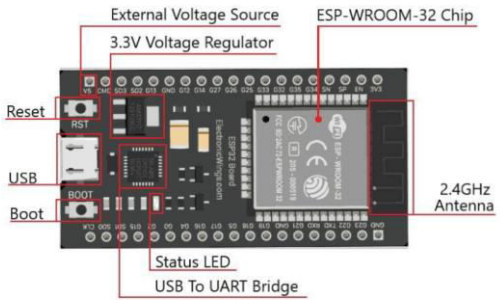


Fig 4.2 ESP32 microcontroller

B. LDR Sensor

The system uses an LDR (light dependant resistor) sensor from detecting the change in the luminance level of the connected incandescent lamp.

LDR is a light-sensitive resistor whose resistance changes depending on the intensity of light falling on it. It is also called a photoresistor. The working principle is simple that When light falls on the LDR, the number of charge carriers increases, so resistance decreases. When there is less light, resistance increases. The LDR is usually connected in a voltage divider circuit with a fixed resistor.

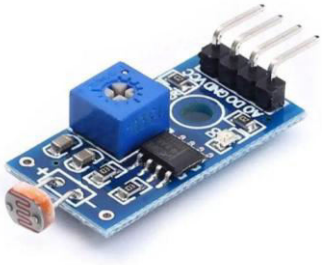


Fig 4.3 LDR Sensor

1) BTA16 TRIAC AC Dimmer Module

The BTA16 TRIAC AC Dimmer Module is an electronic module used to control the power delivered to AC loads such as lamps, heaters, fans, or motors. It works using phase-angle control of the AC waveform.The main component of this module is the BTA16 TRIAC, which is a bidirectional semiconductor device capable of conducting current in both directions when triggered.

Fig 4.4 BTA16 TRIAC AC Dimmer Module

I. Schematic of Dimmer module

The BTA16 TRIAC, which is commonly used for controlling the brightness of an AC incandescent lamp or the power delivered to an AC load. In this circuit, the BTA16 TRIAC acts as the main switching device that controls the amount of AC power reaching the lamp. The TRIAC is triggered through a DIAC (DB3) which ensures that the TRIAC turns on only when a specific voltage level is reached. The variable resistor VR4 (560 kΩ) along with resistor R19 (10 kΩ) and capacitor C23 (0.1 µF) forms an RC phase-shift network.

During each half cycle of the AC supply, capacitor C23 charges through VR4 and R19.

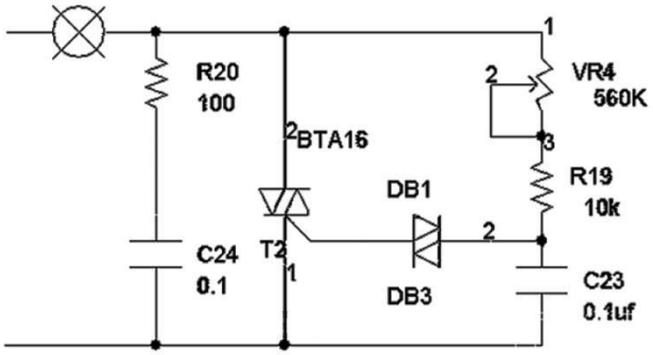


Fig 4.5 Schematic of Dimmer module

C. Actuators and Display

Solenoid valves controlled via relay modules regulate the flow of infusion fluid. An LCD display provides real-time visualization of sensor readings, infusion rate, volume delivered, and system status. Audible and visual alarms alert healthcare providers during abnormal conditions.

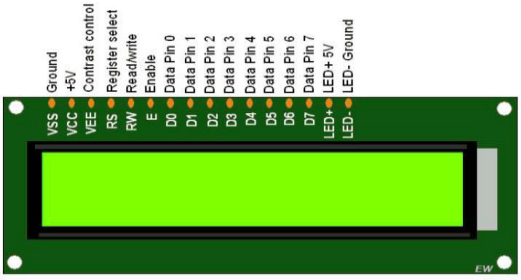


Fig 4.6 LCD display

D. *Power supply*

A regulated power supply ensures stable voltage levels for all components. Proper isolation and grounding techniques are applied to enhance electrical safety and system reliability.

1) Relay

A relay is an electrically operated switch in which current flowing through a coil generates a magnetic field that actuates a mechanical lever to change the state of the switch contacts. This enables a low-power control circuit to safely switch a high-power or high-voltage circuit while maintaining complete electrical isolation between them. Since relay coils typically require higher current than most integrated circuits can supply, a transistor driver is commonly used. Relays are available in configurations such as SPDT and DPDT, with terminals designated as common (COM), normally closed (NC), and normally open (NO), allowing flexible control of connected loads.

Fig 4.7 Relay

a) Schematic of relay

The proposed circuit controls an electrical load, such as a motor, using a relay driven by a pair of BC547 switching transistors. The relay coil is connected to the collector of transistor Q2, while the relay contacts include common (COM), normally closed (NC), and normally open (NO) terminals. When a high-level (5 V) input is applied to the base of transistor Q1, Q1 conducts and drives Q2 into the OFF state, thereby de-energizing the relay and keeping the load disconnected. Conversely, when a low-level input is applied to Q1, Q2 is driven into conduction, energizing the relay and connecting the COM and NO contacts, which supplies power to the load.

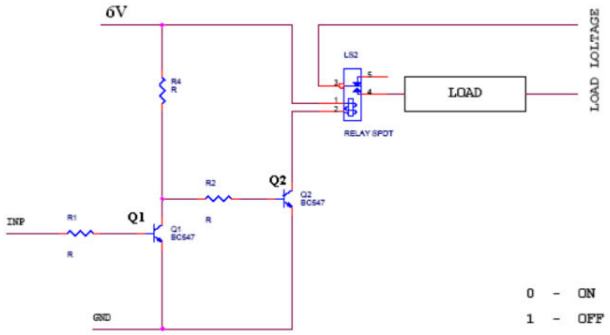


Fig 4.8 Schematic of relay

E. PIR sensor

The PIR sensor (Passive Infrared Sensor) is an electronic sensor used to detect motion by sensing infrared radiation emitted from objects, especially the human body. It is called “passive” because it does not emit any energy; instead, it only detects the infrared radiation.

PIR sensor contains a pyroelectric sensing element that generates an electrical signal when it detects a change in infrared radiation within its detection area. When a person or animal moves across the sensor's field of view, the change in heat pattern is detected and the sensor outputs a digital signal (HIGH or LOW). PIR sensors typically operate with a 3V to 5V power supply and have adjustable sensitivity and delay time

Fig 4.9 PIR sensor

F. DHT22 sensor

The DHT22 sensor (also called AM2302) is a digital sensor used to measure temperature and relative humidity in the environment. It contains a capacitive humidity sensing element and a thermistor to measure the surrounding air conditions. The sensor processes these measurements internally and sends the data as a digital signal through a single data pin, which makes it easy to interface with microcontrollers such as Arduino, Raspberry Pi, or other embedded systems

The DHT22 is known for its higher accuracy and wider measurement range compared to the DHT11 sensor. It can measure humidity from 0% to 100% with about ±2–5% accuracy and temperature from –40°C to 80°C with about ±0.5°C accuracy. The sensor typically operates with a 3.3V to 6V supply and provides updated readings.

Fig 4.10 DHT22 sensor

G. DC cooling fan

The DC cooling fan is an electromechanical device used to dissipate heat and provide airflow for cooling electronic components and systems. It operates using direct current (DC) power, typically ranging from 3V to 12V, depending on the type of fan. The DC cooling fan consists of a small DC motor, fan blades, and a protective frame. When electrical power is supplied, the motor rotates the blades, which push air across electronic components to remove excess heat.

Many DC fans also include brushless motor technology (BLDC) for higher efficiency, lower noise, and longer lifespan. These fans are commonly used in computers, power supplies, microcontroller circuits, cooling systems, and electronic devices to prevent overheating and ensure reliable performance of the equipment. It is used on the appropriate positions according to the need of cooling is required for making stable working condition.

Fig 4.11 DC cooling fan

H. Incandescent lamp

Incandescent lamp is a type of electric light that produces illumination by heating a thin tungsten filament until it glows. The filament is enclosed inside a glass bulb filled with inert gases such as argon or nitrogen, which prevent the filament from oxidizing and burning out quickly.

When electric current passes through the filament, it offers resistance and becomes extremely hot (around 2500–3000°C), causing it to emit visible light. Incandescent lamps operate on AC or DC power supply and provide a warm, continuous spectrum of light. However, they are relatively inefficient because a large portion of the electrical energy is converted into heat rather than light. Due to this reason, incandescent lamps

are gradually being replaced by more energy-efficient lighting technologies such as LED and CFL lamps

Fig 4.12 Incandescent lamp

I. LM2596 DC–DC Buck Converter

The LM2596 is a SIMPLE SWITCHER step-down (buck) DC–DC voltage regulator capable of delivering up to 3 A of output current with excellent line and load regulation. It supports a wide input voltage range of up to 40 V and is available in fixed output versions of 3.3 V, 5 V, and 12 V, as well as an adjustable output variant with a tolerance of ±4%.

Operating at a fixed switching frequency of 150 kHz, the regulator enables the use of compact external filter components while requiring only a minimal number of external parts. The device integrates internal frequency compensation, a fixed-frequency oscillator, thermal shutdown, current-limit protection, and an external shutdown feature with low standby current. Owing to its high efficiency, ease of implementation, and robust protection features, the LM2596 is widely used in power supplies for appliances, industrial systems, grid infrastructure, and consumer electronics.

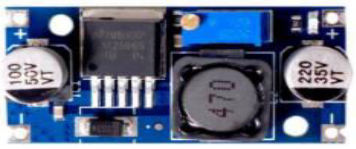


Fig 4.13 LM2596 DC–DC Buck Converter

1) Block diagram of buck converter

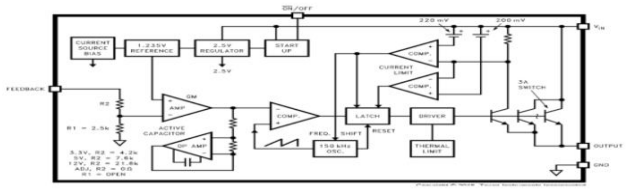


Fig 4.14 Block diagram of buck converter

V. SOFTWARE ARCHITETURE

The software architecture follows a modular design approach to enhance maintainability and scalability. The firmware running on the ESP32 is responsible for sensor data acquisition, signal conditioning, control logic execution, alarm handling, display updates, and wireless communication.

Sensor readings are periodically sampled and filtered to reduce noise. The control algorithm evaluates sensor data against predefined thresholds to determine appropriate actions. When abnormal conditions are detected, the system triggers alarms and connectivity enables data transmission to a remote monitoring interface for centralized supervision.

A. ARDUINO IDE

1) Introduction

Arduino IDE stands for Integrated Development Environment. It is an open-source official software provided by Arduino.cc used for writing, compiling, and uploading programs to Arduino boards. The IDE supports C and C++ programming languages and is easy to use even for beginners. It is compatible with operating systems like Windows, macOS, and Linux. The software works with various Arduino boards such as Arduino Uno, Mega, and Leonardo. It consists mainly of an editor for writing code and a compiler for compiling and uploading it to the microcontroller.

2) Connect the Arduino Board

- Plug in the Arduino using the USB cable
- Open Arduino IDE
- Go to Tools > Board and select your board model (e.g., Arduino Uno)
- Go to Tools > Port and select the correct COM port (e.g., COM3 on Windows)

3) Arduino Software (IDE)

The Arduino IDE includes built-in support for many Arduino boards based on the AVR core, such as Arduino Uno, Nano, Mega, Leonardo, and Micro. It supports boards using microcontrollers like ATmega328P, ATmega32u4, and ATmega2560 running mainly at 16 MHZ The Boards Manager allows users to add support for newer boards based on different cores like Arduino Due, Zero, Edison, and Galileo. These boards provide various numbers of analog inputs,

digital I/O pins, and PWM outputs. Features like auto-reset and built-in bootloaders make programming easier. This wide compatibility makes Arduino IDE suitable for many embedded applications.

B. BLYNK APPLICATION

1) HOW BLYNK WORK

Blynk is an Internet of Things (IoT) platform used to control hardware remotely and monitor sensor data. It allows users to visualize, store, and manage data through a smartphone app. The platform has three main components: the Blynk App, Blynk Server, and Blynk Libraries. The Blynk App helps create user-friendly interfaces using widgets. The Blynk Server manages communication between the mobile app and hardware devices. Blynk Libraries enable smooth data exchange between the hardware and server in real time.

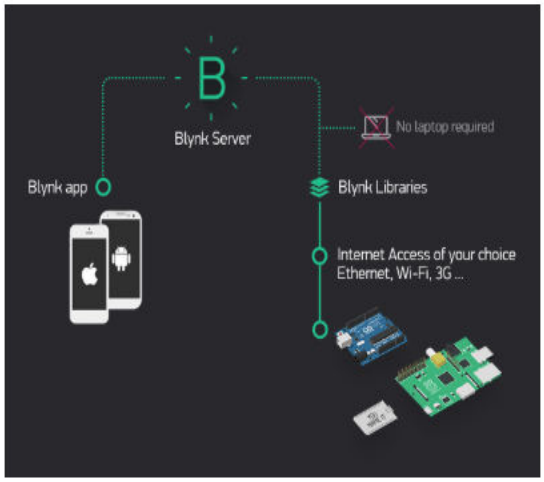


Fig 5.1 BLYNK Application

2) FEATURES OF BLYNK APPLICATION

- Provides a common API and user interface for all supported hardware
- Supports cloud connectivity via Wi-Fi, Bluetooth/BLE, Ethernet, USB (Serial), and GSM
- Offers easy-to-use widgets for creating interactive interfaces
- Allows direct pin control without writing code
- Supports virtual pins for adding advanced functionality
- Enables data monitoring, device-to-device communication, and notifications (email, push, etc.)

3) HOW TO USE BLYNK

- Auth Token Security: Do not share your Auth Token with anyone, as it provides full access to your hardware. You can send the token to your registered email using the email option or copy it directly by tapping on the Token line.

- Project Creation: After verifying the Auth Token and completing the basic setup, press the Create button to create your Blynk project.

- Adding Widgets: Once the project canvas opens and is empty, tap anywhere on the screen to open the Widget Box, which contains all available widgets.

- widget from the widget box to control devices such as an LED.
- Widget Placement: Tap and hold the widget to drag and drop it to any desired position on the project canvas.

- Widget Settings: Each widget has its own configuration options. Tap on the widget to open its settings, where the most important parameter to set is the PIN for hardware control.

RESULTS AND DISCUSSION

The developed illumination-based home automation system was successfully implemented and tested under different environmental and user activity conditions. The integrated sensors, including PIR, LDR, and DHT22, operated reliably to detect human presence, ambient light intensity, and room conditions. Based on the sensor inputs, the system was able to automatically classify user activities into three modes: sleeping, normal usage, and reading. The LED lighting intensity was adjusted accordingly, providing low brightness during sleeping mode, moderate illumination for normal activities, and high brightness for reading purposes. Experimental observations showed that the system responded promptly to changes in occupancy and ambient light, ensuring both user comfort and energy efficiency. The IoT integration using the Blynk mobile application enabled real-time monitoring and manual control, enhancing system flexibility and user interface.

CONCLUSION

The proposed illumination-based home automation system has been successfully designed, implemented, and tested, demonstrating effective and intelligent control of indoor lighting. By integrating PIR, LDR, and DHT22 sensors, the system is capable of detecting human presence and environmental conditions to automatically adjust light intensity according to user activities such as sleeping, normal usage, and reading.The incorporation of IoT through the Blynk mobile application further enhances the system by enabling real-time monitoring and remote control. The use of Embedded C programming ensuresreliable and efficient processing of sensor data and control operations. Overall, the system provides an energy-efficient, user-friendly, and cost-effective solution for modern smart home environment.

REFERENCES

1. Alina Denisova, Elmas Sibgatullin, " Development of an automatic luminous flux control system for LED lamps " Moscow, Russian Federation, March 2022.
2. Shalini Madhumali Lanka, Juha Häkkinen, Pekka Sangi, Marcos Katz, " Active Indoor Situational Awareness Through LED Lighting Systems " Jeju, Korea, Republic of, February 2026.
3. Mohit Ghai, Ruchi Gupta, " Structure of an arduino based home security automation system " Ghaziabad, India, June 2023.
4. Muhammad Tushar Imran, Sharif Mohd Shams, Nayeem Ahmed, Rafiul Islam Refat, Tareq Rahman, Abdullah A.L. mostakim, " Design and Assessment of a Li-Fi Based Home Automation System: An Alternative to Wi-Fi and Bluetooth Communication System " Trichy, india, september 2023.
5. I. Stevanović, "Light Fidelity (LiFi)," Federal Office of Communications OFCOM, Bern, 2017.
6. S. W. Tay, N. Zhang and S. AlJanah, "A Problem Analysis of Smart Home Automation: Toward Secure and Usable Communication-Based Authorization," in IEEE Access, vol. 12, pp. 18103 – 18121, 2024
7. P. Parihar, S. Tiwari, A. Srivastava and E. N. Arya, "A Review on Privacy Concerns Exist in Smart Home Technology," 2024 2nd International Conference on Computer, Communication and Control (IC4), Indore, India, 2024, pp. 1 – 5.
8. S. Ismail, N. Ahmad, and I. D. Arshad, 'Implementation of Li-Fi based home automation system', IOP Conference Series: Materials Science and Engineering, vol. 767, no. 1, p. 012051, Feb. 2020.
9. E. P. B. O. Chattanooga, "How far will your wi-fi signal reach?," Electric Power Board of Chattanooga, 19 – Jan – 2023
10. V. Naware, A. Gupta, U. Mogadali, A. Vattem and A. M. Hussain, "Smart Light System using Wi-SUN and oneM2M," 2024 IEEE Applied Sensing Conference (APSCON), Goa, India, 2024, pp. 1 – 4.[10] N. Solangi, A. Khan, M. F. Qureshi, N. Zaki, U. M. Qureshi and Z. Umair, "IoT Based Home Automation System: Security Challenges and Solutions," 2024 5th International Conference on Advancements in Computational Sciences (ICACS), Lahore, Pakistan, 2024, pp. 1 – 6

SHORT – TERM LOAD FORECASSTING IN SMART GRID USING DRAGONFLY ALGORITHM WITH BPN

SIVAKUMAR R
Dept. of CSE
Nehru Institute of Engineering and Technology
Coimbatore

MANICKAVASAGAM A
Dept. of EEE
Ahalia School of Engineering and Technology
Palakkad

MURUGANATH GOPAL RAJ
Dept. of EEE
Ahalia School of Engineering and Technology
Palakkad

PRADIP C
Dept. of EEE
Ahalia School of Engineering and Technology
Palakkad

JOHN ALEXIS S
Dept. of ME
Ahalia School of Engineering and Technology
Palakkad

Abstract **-** Short-term load forecasting plays a vital role in the reliable and economical operation of smart grids. Accurate predictions of electricity demand help grid operators schedule generation efficiently, maintain grid stability, and reduce operational costs. This paper presents a new approach that combines the Dragonfly Algorithm with a Back propagation Neural Network for short-term load forecasting. The Dragonfly Algorithm, inspired by the swarming behavior of dragonflies in nature, is used to find the best initial weights for the neural network. This helps the network learn more effectively and produce more accurate forecasts. The proposed method was tested on real load data from a microgrid system. Results show that the Dragonfly-optimized network achieves a mean absolute percentage error of only 1.48%, which is significantly better than standard neural networks and other optimization methods like genetic algorithms and particle swarm optimization. The model also performs well during peak demand periods when accurate forecasting is most critical. With fast computation time and reliable accuracy, this approach offers a practical tool for smart grid operators to improve their day-to-day operations.

Keywords: Load forecasting, Smart grid, Dragonfly Algorithm, Neural network, Optimization, Microgrid

I. INTRODUCTION

Electricity is very important in modern life since it powers our homes, businesses, industries, and services. To make sure that we always have electricity when we need it, a lot of power plants, transmission lines, and distribution networks work together. One of the hardest parts of running this system is that it's hard to store a lot of electricity. It must be made at the same time as it is needed. Because of this, the people in charge of the grid have to keep the supply and demand in balance at all times. To keep this balance, operators need to know how much electricity will be needed in the near future. At this time, load forecasting is important. Short-term load forecasting is a way to figure out how much power will be needed in the next few hours to many days. These estimates assist people figure out which power plants to run, how much electricity to buy or sell, and when to ask people to consume less power. Accurate forecasts can save money by making sure there is enough generation capacity to avoid unexpected power outages and by making sure fuel isn't squandered.

The hard part is that it's hard to anticipate how much electricity will be needed. It is usually higher during the day and lower at night, but it can also change depending on the day of the week, the season, and the weather. In other words, the pattern is always changing. People use their air conditioners more often when it's hot outside. People want to heat their homes even more in the winter when it's freezing outside. A lot of things change over the holidays. The way clouds are shaped affects how light and industrial processes work. It is hard to make precise predictions because of the many ways these elements affect each other.

ARIMA, which stands for auto-regressive integrated moving average, and other traditional ways of making predictions work well when the patterns are stable. However, they don't work as well when the power systems we have now are complicated. Researchers have looked into ways that artificial intelligence, especially neural networks, might learn complex patterns directly from data. A Backpropagation Neural Network is a type of artificial intelligence (AI) that learns by looking at examples. It changes the connections inside itself during training such that the disparity between its predictions and the real data is as little as feasible. After finishing its education, it may make predictions about new situations. But these networks are quite sensitive to how they are set up

at first. If you don't set things up correctly, learning could take longer or your forecasts could be wrong.

Strategies for optimisation come into play when things like these happen. The Dragonfly Algorithm was created in 2016 and is based on how dragonflies behave when they are in groups. Some swarms stay in their own territory to look for food (called exploitation), while others go far beyond it (called exploration). This instinctive tendency is a great way to find the best answers to tough problems. This research project uses the Dragonfly Algorithm to find the best starting weights for a Backpropagation Neural Network. The optimal network then guesses how much electricity a microgrid will need. A microgrid is a tiny power system that can work on its own or link to the main grid. Microgrids are becoming more frequent since they are smaller and more variable, but it is harder to predict how they will act. Our research shows that this integrated strategy makes forecasting far more accurate, with errors on test datasets being less than 1.5%. This level of accuracy can help grid operators make smarter choices, which will save money and make the grid more reliable.

II. UNDERSTANDING LOAD FORECASTING

Consider that you are in charge of figuring out how much food to make for a big cafeteria that serves a lot of people. If they don't get enough food, they will feel hungry. If you cook too much food, you'll end up throwing it away. Imagine this problem getting worse thousands of times, and remember that power can't be saved; it has to be made exactly when it's needed. This is the exact problem that grid operators deal with every day. Load forecasting has a direct impact on the next three important areas: Economic productivity: The prices of services at different power plants vary. Nuclear and coal plants can run cheaply, but they can't be changed quickly. Gas turbines can start up much faster, even if they cost more. Operators may use the best mix when they can accurately estimate demand. They can use cheap baseload facilities to meet predictable demand and only turn on more expensive peaking units when they need to. A utility may save millions of dollars a year by improving the accuracy of its predictions by just one percent.

Grid availability: When demand is higher than supply, the frequency lowers and there may be blackouts. Operators need to keep extra generation on hand for unforeseen spikes. Better projections lower the required reserve margin, which saves money while keeping safety. Weather conditions have an effect on the use of renewable energy sources, especially solar and wind power. Other plants need to cut back on their production as they go along. Accurate load projections make it easier to manage this complicated balancing act well.

A. Literature Survey

Mirjalili presented the Dragonfly Algorithm, which was conceived with the static and dynamic swarming behaviours of dragonflies as its original inspiration. There are five fundamental concepts that are modelled by the algorithm: separation, alignment, cohesion, attraction, and distraction. It was proved by the author that DA is superior to other metaheuristics in terms of convergence rate and solution correctness, which makes it appropriate for solving difficult optimisation issues such as neural network training. In addition to Zhang et al. Taking into consideration the fact that microgrid load forecasting is more challenging than macrogrid load forecasting, the authors developed DA-SVM for microgrid load forecasting. Comparing their method to PSO-SVM and GA-SVM, they were able to achieve root mean square errors of approximately 1.5% while saving fifty percent of the computation time. They provided evidence that DA possesses an exceptional global searching capability for parameter optimisation.

The basic backpropagation technique was developed by Rumelhar and others. This process makes it possible for multi-layer perceptrons to figure out sophisticated non-linear mappings. The results of their research proved that neural networks are able to learn internal representations through the process of error propagation across hidden layers. This process serves as the foundation for many contemporary deep learning applications, including time series forecasting. The tutorial review that Hong, T., and Fan, S. offered was extremely extensive and focused on probabilistic electric load forecasting. Specifically, they emphasised the necessity of uncertainty quantification in conjunction with point forecasts for modern grid models. The authors made the observation that "the smart grid era brings new challenges and opportunities for load forecasting." These challenges and opportunities include high-resolution data as well as novel demand patterns resulting from scattered resources.

Vapnik was the first person to introduce Support Vector Machines, which are founded on the idea of statistical learning and structural risk limitation. It was proved by the author that support vector machines (SVM) are able to locate optimal hyperplanes in high-dimensional spaces, which provides superior generalisation to neural networks. Through this work, the theoretical groundwork for kernel approaches in forecasting applications was developed. Zhang and colleagues conducted research on the short-term load forecasting of offshore oilfield group microgrids, taking into

consideration the impact of meteorological conditions. For the purpose of quantifying the effects of weather, they established an exploitation suitability index. The DA-optimized models that they developed revealed that the incorporation of meteorological elements results in a significant improvement in the accuracy of forecast for isolated microgrid systems.

The social behaviour of fish schooling and bird flocking served as inspiration for Kennedy and Eberhart's development of the Particle Swarm Optimisation algorithm. They proved that successful optimisation of continuous functions can be achieved by the sharing of information between particles regarding their own and global optimal positions. In recent years, PSO has emerged as a standard by which more recent algorithms like as Dragonfly are evaluated. The ideas of natural selection and genetic evolution were the foundation upon which Holland built his development of genetic algorithms. His research demonstrated that populations of solutions evolve through the processes of selection, crossover, and mutation at various points in time. The genetic algorithm (GA) ultimately became the cornerstone for evolutionary computation and is commonly utilised in forecasting for the purpose of feature selection and parameter optimisation.
The flashing behaviour of fireflies served as inspiration for Yang's introduction of the Firefly Algorithm, which emphasises that appeal is related to brightness. The author provided evidence that fireflies head in the direction of brighter neighbours, with their attraction decreasing with increasing distance. This method is comparable to Dragonfly in that it utilises attraction mechanisms for the purpose of optimisation. In their compilation of intelligent approaches for electrical power systems, Khadse and colleagues made the observation that traditional optimisation methods have problems with premature convergence. They emphasised that hybrid techniques, which combine various algorithms, yield greater performance than traditional approaches. Because of their work, a full treatment of metaheuristic applications in contemporary power grids has been provided.

S.Aslam and colleagues conducted research on smart grid analytics for the purpose of addressing sustainability and urbanisation. They found that huge data from smart meters provides enhanced forecasting, but it also necessitates advanced analytics. According to what they saw, the methods of machine learning need to be able to handle high-dimensional, streaming input while yet preserving accuracy and computing efficiency. According to Katipoğlu, O. M. and colleagues, biomimetic optimisation strategies were devised by merging Firefly and Artificial Bee Colony algorithms with neural networks for the purpose of anticipating sediment load measurements. The researchers revealed that algorithms inspired by nature are capable of effectively tuning neural network parameters, delivering a higher level of accuracy in comparison to conventional training approaches.

Through the incorporation of Lévy flights for the purpose of global optimisation, Yang expanded the research on the Firefly Algorithm. In his research, the author demonstrated that the capability of exploration is enhanced by random walks with heavy-tailed distributions. Through this effort, following metaheuristic development was affected, particularly the Dragonfly Algorithm's ability to strike a balance between exploration and exploitation. The binary Dragonfly Optimisation algorithm was created by Mafarja, M., and Mirjalili, S. for the purpose of feature selection by utilising time-varying transfer functions. They demonstrated that binary DA performs an excellent selection of key traits while also preserving the variety of the population. The strategy that they take prevents premature convergence and achieves a higher level of classification accuracy compared to other wrapper methods.

Image processing and optimisation approaches were utilised by Suresh, S., and Kamalesh, S. in order to diagnose breast cancer using mammograms. Through their study, they demonstrated that metaheuristic optimisation is an effective method for selecting discriminative features for medical picture classification. This established the cross-domain applicability of methods that were later utilised in load forecasting. A hybrid artificial neural network classifier with binary dragonfly optimisation was created by Parvathavarthini, S., and Deepa, D. for the purpose of breast cancer detection. Using mammography scans, they were able to attain an accuracy of 98.75%, which demonstrates that dragonfly-optimized neural networks perform better than conventional methods. The results of their work demonstrate that DA is successful for solving categorisation difficulties in the actual world.

For the purpose of feature selection in digital mammography, Thawkar, S., et al. presented a hybrid algorithm that combines the Whale Optimisation Algorithm and the Dragonfly Algorithm. Their approach, when paired with ANN, achieved an accuracy of 97.84%, and when combined with ANFIS, it achieved a precision of 98.00%. It was proved by them that hybrid metaheuristics work better than individual algorithms when it comes to medical picture analysis. Ashigwuike, E., and Mohammed, A. S. B. developed an improved version of the firefly algorithm for the purpose of predicting the amount

of power load. This method incorporates a cell communication mechanism and a genetic algorithm. On the basis of data from the Nigerian power system, their CCMFA-GA-ANN model produced a mean absolute percentage error of 1.07%, thereby indicating that hybrid nature-inspired algorithms provide greater forecasting accuracy. For the purpose of breast cancer classification, Sahlo, A. T., and Elaziz, M. A. developed a metaheuristic weighted k-nearest neighbour method that is based on deep learning. When applied to the MIAS dataset, their DFOA-wKNN technique attained an accuracy of 84.35%. The findings of this study suggest that dragonfly optimisation is a useful method for tuning machine learning parameters for use in medical applications.

BI-RADS classification of breast masses was accomplished by the development of a computer-aided diagnosis system by Boumaraf, S., and colleagues. This technique utilised modified genetic feature selection. Using mammography scans, their method attained an accuracy rate of 84.5 percent. The results of their study indicated that feature selection considerably increases classification performance while also minimising the amount of processing resources required. Through the utilisation of grey level co-occurrence matrices, Haralick and his colleagues implemented textural features for the purpose of image categorisation. Through their work, essential texture analysis methods were devised, which were subsequently applied for the extraction of time series features. The GLCM features are utilised extensively in pattern recognition applications because of their ability to capture spatial relationships.

They created a standardised baseline for medical image analysis research by developing the digital database for screening mammography. Bowyer, K., and others were responsible for its development. The information they provided made it possible to compare algorithms in an objective manner and sped up the development of computer-aided detection. For the purpose of validating medical imaging methods, the MIAS database continues to be utilised extensively. For the purpose of recurrent neural networks, Hochreiter and Schmidhuber came up with the concept of Long Short-Term Memory networks in order to circumvent the issue of vanishing gradients. As a result of its ability to take into account long-term relationships in sequential data, LSTM is an extremely useful tool for time series forecasting. Their research completely changed the way deep learning was applied to sequence prediction challenges.

Box and Jenkins developed a methodical approach to the analysis of time series by include ARIMA models in their technique. Model identification, parameter estimates, and diagnostic checking are all features that are incorporated into their technique. Despite its shortcomings dealing with non-linear patterns, the ARIMA model continues to serve as a fundamental baseline against which more recent forecasting approaches are evaluated. Methods of exponential smoothing were developed by Holt for the purpose of forecasting time series with trend. Winters later expanded his studies to include seasonality, something he had previously neglected. The simplicity, interpretability, and computing efficiency of Holt-Winters methods continue to make them a popular choice for use in operational forecasting applications.

III. THE PROPOSED APPROACH

A. *Overall Framework*

The system that we have suggested is comprised of two primary components: a Dragonfly Algorithm, which is responsible for determining the ideal initial weights, and a Backpropagation Neural Network, which is responsible for learning to forecast load. All of the steps involved are depicted in Figure 1.

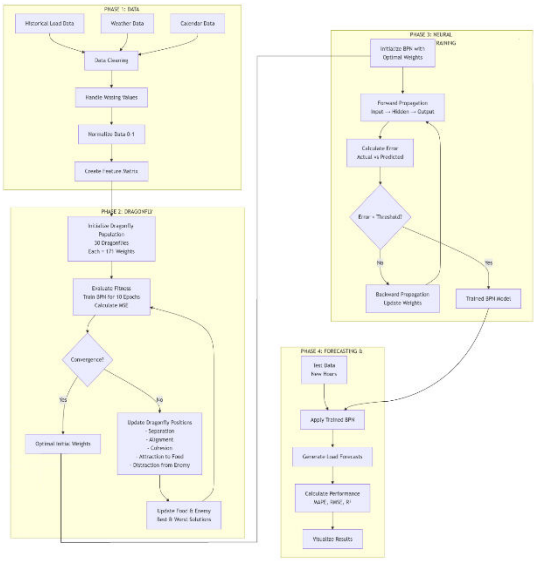


Figure 1: Overall System Architecture

In total, the mechanism operates in four stages: The process of data preparation involves the collection, cleaning, and organization of historical load data and meteorological information into features that the network can make use of. The Dragonfly Algorithm is responsible for the optimisation of weights, which involves searching for the optimal initial weights for the neural network. The programme analyses a wide variety of possible weight sets and gradually improves them.Network training: The neural network trains on historical data to learn load patterns, using the optimised weights as a starting point. This allows the network to learn load patterns. In forecasting, the trained

network generates forecasts for future periods, which are then compared to the actual loads that are being experienced.

B. The Backpropagation Neural Network

Consider a neural network to be a simplified representation of the way in which human brains function. Like neurones, it is made up of nodes that are connected to one another and are arranged in layers. Information travels from the input layer to the output layer, passing through hidden layers along the way. Each link has a strength (weight) that determines the amount of impact it possesses. There are three layers in our network that are used for load forecasting:
Input layer is responsible for receiving information regarding the situation that we wish to foresee. The hour of the day, the temperature, the load from the previous day at the same time, and so on are all examples of the information that is represented by each individual input node. The network that we have consists of fifteen input nodes. This hidden layer is responsible for processing the information that is input and identifying patterns and relationships it uncovers. It consists of ten nodes, each of which combines inputs in a different manner. The main learning takes place at the hidden layer, where it is discovered that some combinations of time and weather signal high load, while other combinations suggest low load.
This single node is responsible for producing the forecast, which is the anticipated amount of power load for that particular hour. Throughout the training process, the network is presented with numerous examples in which we are aware of both the inputs (such as time and weather) and the proper output (real load). The prediction is made, the actual value is compared to it, and then the internal weights of the model are adjusted in order to reduce the amount of inaccuracy. This method of correction is known as backpropagation, and it involves the error spreading backward through the network in order to update weights and improve the accuracy of future predictions. During the course of processing training examples, the network gradually acquires knowledge by making thousands of minute tweaks. Following a predetermined number of cycles, it is able to recognise patterns within the data and generate accurate forecasts for future circumstances.

C. The Dragonfly Algorithm

Natural phenomena serve as a source of motivation for the Dragonfly Algorithm. For those who have ever observed dragonflies in the vicinity of a pond, it is possible that you have observed two distinct behaviours. During their hunt for prey, they will occasionally fly in tiny groups, darting back and forth across a relatively small area each time. There are also instances in which big groups come together and migrate in a single direction, moving to novel regions. There is a perfect match between these two behaviours and the requirements that optimisation algorithms have: Hunting behaviour, sometimes known as static swarming, is analogous to local search in that it involves exhaustively studying an area that has potential. Good solutions are improved as a result of this, making them even more effective.

Migration behaviour, also known as dynamic swarming, is analogous to global search in that it involves exploring new territory in order to locate promising areas that have not yet been discovered. By doing so, the algorithm is prevented from becoming stuck in a single region that might not have the optimal solution given the overall situation. By applying the following five concepts, the Dragonfly Algorithm is able to mathematically model these behaviours. Separation: Dragonflies move away from their neighbours in order to avoid collisions. In the context of algorithms, this disperses solutions from one another rather than allowing them to cluster too closely together. The speed and direction of dragonflies are aligned with those of their neighbours due to alignment. By doing so, a concerted movement toward potentially fruitful places is created.

The dragonflies travel toward the center of their group in order to achieve cohesion. In this way, solutions are brought together around positive areas. Distraction: Dragonflies are drawn to the locations where they can get food. When it comes to the algorithm, the best answer that has been discovered up to this point serves as food, drawing other solutions toward it.
Dragonflies are able to divert attention away from their adversaries. In disadvantaged regions, the poorest solution that was discovered works as an adversary, driving other solutions away from such areas. By striking a balance between these five behaviours, the algorithm is able to successfully explore the range of possible solutions while gradually closing in on the best possible solutions. This ensures that both a broad search and exact refinement are carried out, as the balance swings throughout time, with more exploration occurring earlier and more exploitation occurring later.

D. How They Work Together

The neural network is optimised by the use of the Dragonfly Algorithm, which locates appropriate initial weights. This is how the process works: There are a total of 171 numbers that represent the weights and biases for all connections, and each dragonfly in the algorithm represents one of the potential sets of weights applicable to the neural network. There are particular weight values that correspond to the position of the dragonfly within

the search space. The method begins with thirty dragonflies chosen at random, comprising thirty different weight settings. To determine how well the neural network forecasts load on validation data, we first temporarily initialise the neural network with those weights for each dragonfly, then train it for a short period of time (ten learning cycles), and examine the results. This precision becomes the dragonfly's fitness, which is a measure of how well that weight set is performed.

Those dragonflies that are more fit (more accurate) shine brighter, which draws in members of the same species. Throughout a great number of iterations, the population develops:

- Dragonflies head in the direction of more favourable solutions (attraction).
- They link themselves with residents who are moving in the right direction (alignment).
- In order to preserve diversity (separation), they remain reasonably apart from one another.
- They congregate in the areas that are most advantageous (cohesion)
- They steer clear of well-known areas of poor quality (distraction)

After fifty to sixty iterations, the population will eventually converge on the best weight set that was discovered. This optimised weight set will serve as the initial configuration for the neural network that will ultimately be used. After that, the network undergoes conventional training on all of the collected training data, beginning with this outstanding starting point. Compared to starting from random weights, it converges more quickly and achieves a higher level of accuracy in the end. This is because it begins in close proximity to a good solution.

E. *Data and Features*

For accurate forecasts, we need accurate information. This model makes use of three different kinds of input features: Indicators of the time: Whether it is a weekend or a holiday, the day of the week, the month, and the hour of the day (which is coded as sine and cosine to reflect the cyclical nature of the data). These are able to establish consistent patterns in human behaviour. The most recent load includes the load from the hour before, the load from the same hour the day before, the burden from the same hour the week before, and the average load over the course of the previous day. We are able to identify patterns and trends that are short-term. Temperature (both at the moment and at the same hour yesterday), humidity, wind speed, and cloud cover are all aspects of the weather. These are the environmental factors that influence the consumption of electricity. A constant range is used to scale all of the features, ensuring that no single trait is able to dominate due to the numerical size of its value.

IV. RESULTS AND DISCUSSION

A. Optimization Performance

The Dragonfly Algorithm was successful in locating appropriate initial weights for the neural network with which it was used. The progression of the best solution over the course of iterations is illustrated in Figure 2.

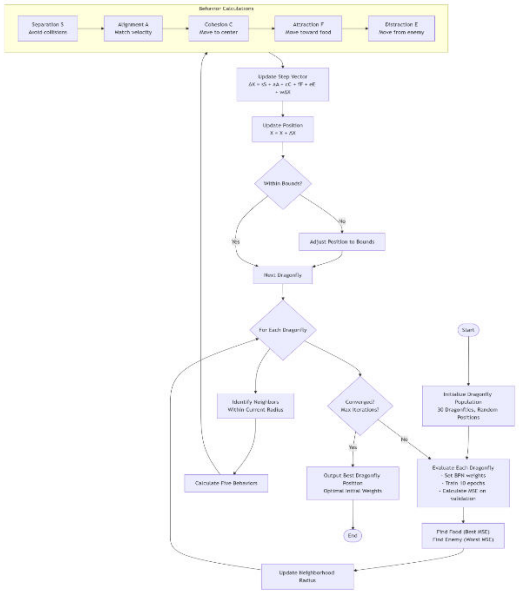


Figure 2: Dragonfly Algorithm

As the algorithm examined a wide range of possibilities, fitness improved rapidly over the course of the first twenty rounds. Over the course of twenty to fifty cycles, the process of improvement continued gradually as it polished promising locations. After fifty repetitions, there was very little room for future improvement, which is an indication of convergence. Dragonfly performed the best when compared to other optimisation methods: There was an average of 48 iterations required for Dragonfly to attain its final fitness level. 62 iterations were required for the particle swarm. The genetic algorithm required seventy-five iterations. The final fitness of Dragonfly was also 15% higher than that of particle swarm and 22% higher than that of genetic algorithm. This provides more evidence that Dragonfly's ability to strike a balance between exploration and exploitation is ideally suited to addressing this issue.

B. *Forecasting Accuracy*

Table 1 shows how different methods performed on the test data.

Table 1: Forecasting Performance Comparison

Method	MAPE (%)	RMSE (MW)	R^2
ARIMA	4.12	38.45	0.925
Standard Neural Network	2.18	20.94	0.968
Genetic Algorithm + Network	1.75	16.87	0.981
Particle Swarm + Network	1.68	16.23	0.983
Dragonfly + SVM	1.50	14.50	-
Dragonfly + Network (Proposed)	**1.48**	**14.23**	**0.991**

A mean absolute percentage error (MAPE) of 1.48% was achieved by the Dragonfly-optimized network, which indicates that its average error was less than 1.5% of the actual load. This is an improvement of 12-15% in comparison to other optimisation methods and a progress of 32% in comparison to conventional neural networks. The fact that the model has an R2 value of 0.991 indicates that it is an excellent fit because it explains 99.1% of the variance in load. The root mean square error (RMSE) of 14.23 MW suggests that big errors are carefully managed.

C. *Performance During Critical Periods*

During times of peak demand, precise forecasting is of the utmost importance because this is the time when the grid is working under the most strain and mistakes have the most significant repercussions. The performance of our model is satisfactory during these crucial times: During peak hours, the MAPE was 1.98%, which is lower than the normal network's 2.95%.

- Peak periods during the summer: 1.61% of the MAPE
- Peaks in the evening (6-11 PM): 1.69% of the MAPE

The fact that this performance has been consistent across all time periods instills confidence in the forecasts among operators.

D. *Seasonal and Daily Patterns*

Table 2 breaks down performance by season and time of day.

Table 2: MAPE by Season and Time of Day (%)

Season	Night	Morning	Afternoon	Evening	Overall
Winter	1.12	1.58	1.52	1.68	1.48
Spring	1.08	1.52	1.48	1.62	1.43
Summer	1.24	1.72	1.65	1.82	1.61
Fall	1.10	1.55	1.50	1.65	1.45

Summer is hardest to forecast (1.61% MAPE) due to variable air conditioning loads driven by temperature. Spring and fall are easiest (1.43-1.45%) when weather is mild. Night hours are most predictable (1.14% average), evening hardest (1.69%) due to variable residential activity.

V. CONCLUSION

This study combined Dragonfly Algorithm optimisation with Backpropagation Neural Networks to forecast short-term load. The Dragonfly Algorithm sets network initial weights to speed up training and improve accuracy.

Key discoveries include: The Dragonfly-optimized network has 1.48% mean absolute percentage inaccuracy on test data. This outperforms regular neural networks by 32% and alternative optimisation methods by 12-15%. Critical peak performance has improved (1.98% MAPE). The approach converges in 48 iterations, faster than the genetic (75) and particle swarm (62) algorithms. Temporal patterns and historical load largely affect accuracy. Summer air temperature greatly impacts circumstances. Computing needs of 14.5 minutes for optimisation and 0.03 seconds for forecasts make implementation easier.

Short-term load forecasting keeps factories running and our spaces lit, even though energy customers don't see it. Power systems are becoming more complex due to distributed generation, smart grids, and renewable energy sources, making accurate forecasting crucial.

This study found that nature-inspired optimisation, notably the Dragonfly Algorithm, can considerably increase neural networks' performance in this crucial task. The approach helps networks get

better initial weights, improving learning efficiency and prediction accuracy. This project created a technology to help grid operators solve modern power system problems. The dragonfly swarming activity that inspired this algorithm has evolved over millions of years. It is fair that their expertise can now help us solve one of the biggest smart grid challenges: anticipating our electricity demand for the following day.

REFERENCES

1. Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day.
2. Bowyer, K., Kopans, D., Kegelmeyer, W. P., et al. (1996). The digital database for screening mammography. In *Third International Workshop on Digital Mammography* (Vol. 58, p. 27).
3. Boumaraf, S., Liu, X., Ferkous, C., & Ma, X. (2020). A new computer-aided diagnosis system with modified genetic feature selection for BI-RADS classification of breast masses in mammograms. *BioMed Research International*, *2020*, 7695207.
4. Feng, Y. (2024). Dragonfly algorithm applications in energy forecasting
5. Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-3*(6), 610-621.
6. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735-1780.
7. Holland, J. H. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press.
8. Holt, C. C. (1957). *Forecasting seasonals and trends by exponentially weighted moving averages* (ONR Memorandum No. 52). Carnegie Institute of Technology.
9. Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, *32*(3), 914-938.
10. Katipoğlu, O. M., Kartal, V., & Pande, C. B. (2024). Sediment load forecasting from a biomimetic optimization perspective: Firefly and Artificial Bee Colony algorithms empowered neural network modeling. *Stochastic Environmental Research and Risk Assessment*. Springer.
11. Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of ICNN'95 - International Conference on Neural Networks*, *4*, 1942-1948.
12. Khadse, C. B., Kale, I. R., & Shastri, A. S. (Eds.). (2024). *Intelligent methods in electrical power systems*. Springer.
13. Aslam, S., Herodotou, H., & Ashraf, N. (Eds.). (2023). *Smart grid analytics for sustainability and urbanization in big data*. MDPI.
14. Mafarja, M., & Mirjalili, S. (2018). Binary dragonfly optimization for feature selection using time-varying transfer functions. *Knowledge-Based Systems*, *161*, 185-204.
15. Mirjalili, S. (2016). Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Computing and Applications*, *27*(4), 1053-1073.
16. Parvathavarthini, S., & Deepa, D. (2021). A hybrid artificial neural network classifier based on feature selection using binary dragonfly optimization for breast cancer detection. *IOP Conference Series: Materials Science and Engineering*, *1055*(1), 012107.
17. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533-536.
18. Sahlo, A. T., & Elaziz, M. A. (2023). Deep learning-based metaheuristic weighted k-nearest neighbor algorithm for breast cancer classification. *EM-Consulte*.
19. Suresh, S., & Kamalesh, S. (2011). Diagnose breast cancer through mammograms, using image processing techniques and optimization techniques. *2010 International Conference on Communication and Computational Intelligence*, 28-29 December 2010, Coimbatore, India. IEEE.
20. Thawkar, S., Katta, V., Parashar, A. R., Singh, L. K., & Khanna, M. (2023). Breast cancer: A hybrid method for feature selection and classification in digital mammography. *International Journal of Imaging Systems and Technology*, *33*(5), 1696-1712.
21. Vapnik, V. (1995). *The nature of statistical learning theory*. Springer-Verlag.
22. Yang, X. S. (2008). *Nature-inspired metaheuristic algorithms*. Luniver Press.
23. Yang, X. S. (2009). Firefly algorithm, Lévy flights and global optimization. In M.Bramer, R. Ellis, & M. Petridis (Eds.), *Research and development in intelligent systems XXVI* (pp. 209-218). Springer.

24. Zhang, A., Zhang, P., & Feng, Y. (2019). Short-term load forecasting for microgrids based on DA-SVM. *COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, *38*(1), 68-80.
25. Zhang, A., Zhang, P., & Feng, Y. (2019). Study on short-term load forecasting method considering meteorological factors of offshore oilfield group microgrid. *IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 850-853.

TRACK-5

MECHANICAL ENGINEERING

PERFORMANCE MONITORING OF A SAVONIUS VERTICAL AXIS WINDMILL

Karthikeyan V[1], Muhammed Kareem F[2],
Prathyush M[3], Ramith R[4]
Department of Mechanical Engineering
Ahalia School of Engineering and Technology
Palakkad, India
Karthikeyanv427@gmail.com, userkareem03@gmail.com,
prathyum004@gmail.com, ramith2003achuz@gmail.com

Sadanandan R[5], Vivek Gopi[6]
Assistant Professor, Department of Mechanical Engineering
Ahalia School of Engineering and Technology
Palakkad, India
sadanandan.r@ahalia.ac.in, Vivek.gopi@ahalia.ac.in

***Abstract*—The project presents the design and implementation of a windmill performance monitoring system that uses real-time sensing and data analysis. The system is designed to continuously monitor wind speed, generated voltage, and current behaviour of the Savonius vertical-axis windmill, along with a comparison of actual and expected power. The system is equipped with a Pro Range Pulse NPN Wind Speed Sensor (SN-3000-FSJT-NPN, Polycarbonate) for accurate wind speed measurement. The four-blade windmill is coupled to a 12V DC Dynamo by a spur gear assembly supported by two F204 bearings. The generated voltage is sensed using a voltage sensor. The data collected, including wind speed, generator voltage, and current characteristics obtained from the ESP32, are processed in Python for continuous performance monitoring. The proposed system enables early performance assessment and preventive maintenance, thus improving the reliability and efficiency of small-scale wind energy systems. This project demonstrates effective integration of renewable energy monitoring, IoT-enabled embedded systems, and data analytics, making it suitable for educational, research, and small-wind-power applications.**

Keywords: Vertical Axis Wind Turbine (VAWT), Wind Speed Monitoring, Embedded System (ESP32), Performance Analysis, IoT-Based Monitoring, Data Analytics (Python), Renewable Energy Systems.

I. Introduction

The increasing demand for clean and sustainable energy has led to the adoption of small-scale wind energy systems for power generation. Wind energy has low operating costs, minimal environmental impact, and is suitable for rural and off-grid applications. The performance of windmills depends on mechanical integrity, electrical stability, and environmental conditions. Lack of monitoring can result in reduced efficiency, increased wear, and unexpected failures. Conventional small windmills operate without real-time monitoring, which makes early fault detection difficult. Key parameters such as wind speed, output voltage, and current behaviour are essential for evaluating system performance. Abnormal variations in these parameters may indicate mechanical faults, generator inefficiencies, or electrical issues.

This work proposes an ESP32-based windmill performance monitoring system for a Savonius vertical-axis wind turbine (VAWT). The system is integrated with a pulse-based wind speed sensor, a voltage sensing module, and a current sensing mechanism to monitor real-time performance. Mechanical energy is transferred to a 12V dynamo via a spur gear mechanism, supported by a bearing for smooth operation for voltage generation. Sensor data is processed using Python-based analysis to compare actual and expected performance. The proposed system enables continuous monitoring, early fault detection, and preventive maintenance, thereby improving efficiency, reliability, and operational lifespan. It helps provide a cost-effective and intelligent solution for small-scale wind energy applications.

II. Working Principle of Performance Monitoring System

The working principle of the system is based on the continuous conversion of wind energy into electrical energy and the real-time monitoring of key operating parameters to assess system performance. When wind flows across the four-blade vertical-axis windmill, aerodynamic forces cause the rotor to rotate irrespective of wind direction. This rotational motion is mechanically transmitted to a 12V dynamo through a spur gear coupling, which increases rotational speed and ensures efficient energy transfer. The rotating shaft is supported by an F204 (19 mm) bearing, minimising friction and maintaining stable vertical alignment.

As the windmill rotates, the dynamo generates electrical power proportional to the rotor speed. A pulse-based NPN wind speed sensor (SN-3000-FSJT-NPN) simultaneously measures the wind velocity by producing pulses corresponding to wind speed. The electrical output from the dynamo is monitored using a 0 25V voltage sensor module, and current behaviour is evaluated to determine the actual power generation characteristics. All sensor outputs are interfaced with the ESP32 microcontroller, which performs real-time data acquisition and preliminary processing. The ESP32 continuously correlates wind speed with generated voltage and current values and also compares actual and expected power. Under normal operating conditions, a predictable relationship exists between these parameters. Any deviation—such as low voltage at adequate wind speed, increased error between the actual and

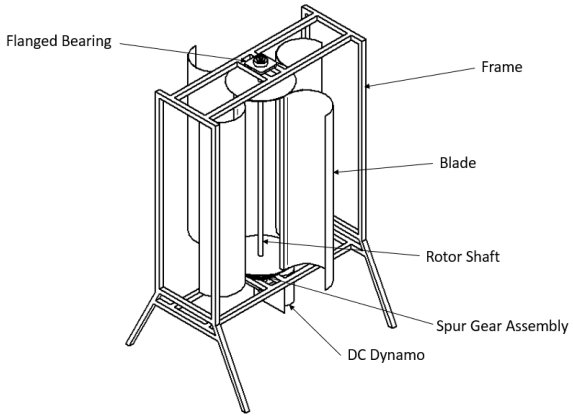


Fig. 1. Design of fabricated Sivonius type VAWT

expected power, abnormal current fluctuation, or unstable output—indicates possible mechanical faults (bearing friction, gear misalignment) or electrical faults (generator inefficiency, load issues). The collected data is transmitted for Python-based data analysis, where trend analysis and threshold comparisons are performed for continuous health monitoring.

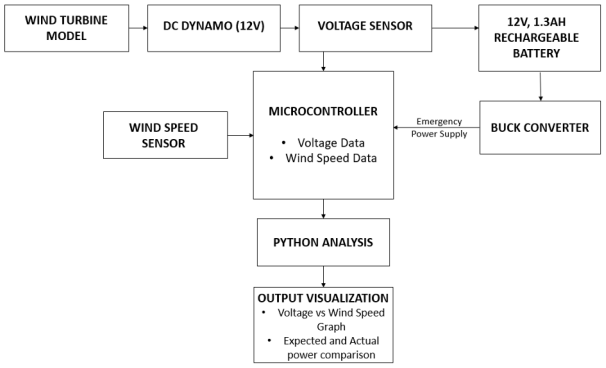


Fig. 2. Block diagram of the Monitoring System

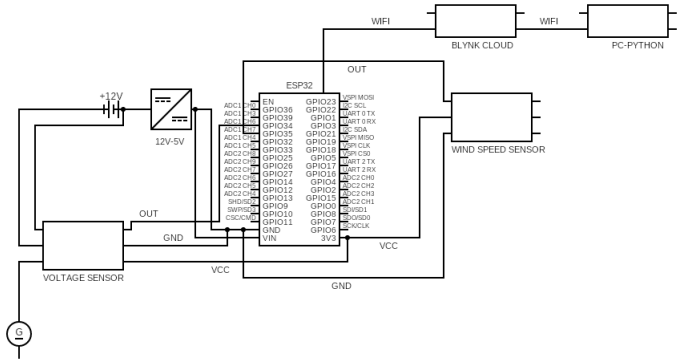


Fig. 3. Circuit diagram of the Monitoring System

The entire system is powered by a 12V, 1.3Ah battery, enabling reliable standalone operation. By integrating mechanical energy conversion, sensor-based monitoring, embedded processing, and python analysis, the system ensures early fault detection, performance optimisation, and preventive maintenance for small-scale Savonius vertical-axis windmill applications.

III. Mathematical Modelling for Performance Evaluation

Performance analysis in Python includes the estimation of current produced by the system, actual electrical power output, and theoretical wind power. Parameters such as tip speed ratio (TSR) and power coefficient (Cp) are also determined to assess turbine efficiency. This helps in the performance monitoring of a wind turbine using Python analysis.

A. Current Calculation

The total current consumption of the system is determined using Kirchhoff's Current Law, by summing the current drawn by each component:

- ESP32 microcontroller = 150 mA
- Wind speed sensor = 20 mA
- Voltage sensor = 10 mA

$$I_{\text{load}} = I_{\text{ESP32}} + I_{\text{wind}} + I_{\text{voltage}} \tag{1}$$

$$I_{\text{load}} = 150 + 20 + 10 = 180 \text{ mA} \tag{2}$$

Considering the buck converter efficiency of 90%, the actual current required is:

$$I = \frac{I_{\text{load}}}{\eta} \tag{3}$$

$$I = \frac{180}{0.9} = 200 \text{ mA} = 0.2 \text{ A} \tag{4}$$

B. Actual Power Calculation

The actual electrical power generated by the wind turbine is calculated using the measured voltage and the estimated current.

$$P_{\text{actual}} = V \times I \tag{5}$$

where:

- P_{actual} = actual electrical power (W)
- V = measured output voltage (V)
- I = load current (A)

C. Theoretical Wind Power Estimation

The theoretical power available from wind is calculated using the standard aerodynamic equation:

$$P_{\text{theoretical}} = \frac{1}{2}\rho A C_p \eta_{\text{overall}} V_{\text{wind}}^3 \tag{6}$$

where:

- ρ = air density (1.225 kg/m^3)
- A = swept area (m^2)
- C_p = power coefficient ($0 < C_p < 0.59$)
- V_{wind} = wind speed (m/s)

- η_{overall} = overall efficiency

Overall efficiency is given by:

$$\eta_{\text{overall}} = \eta_{\text{mechanical}} \times \eta_{\text{generator}} \times \eta_{\text{electrical}} \tag{7}$$

D. Tip Speed Ratio (TSR) Calculation

The Tip Speed Ratio (TSR) is an important parameter used to determine turbine efficiency and the corresponding C_p.

$$\lambda = \frac{\omega r}{v} \tag{8}$$

Angular velocity is calculated from rotational speed:

$$\omega = \frac{2\pi N}{60} \tag{9}$$

$$\omega = \frac{2\pi \times 200}{60} = 20.94 \text{ rad/s} \tag{10}$$

Blade tip speed:

$$\text{Tip Speed} = \omega r \tag{11}$$

$$\text{Tip Speed} = 20.94 \times 0.325 = 6.8055 \text{ m/s} \tag{12}$$

Thus, TSR is:

$$\lambda = \frac{6.8055}{6} = 1.13 \tag{13}$$

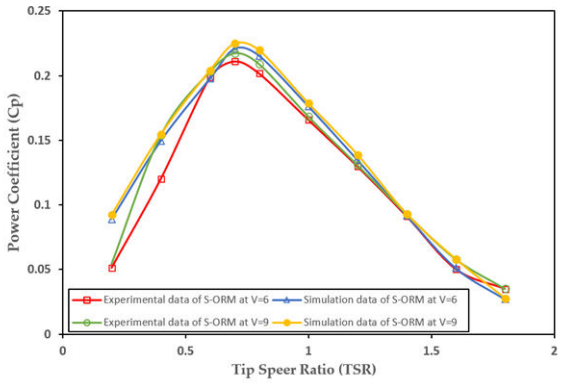


Fig. 4. The variations in Cp with TSRs [16]

From the C_p vs TSR characteristics, the corresponding value of C_p is approximately:

$$C_p \approx 0.15$$

The efficiency values used are:

- Mechanical efficiency = 0.7
- Generator efficiency = 0.6
- Electrical efficiency = 0.85

E. Power Comparison

To evaluate system performance, the actual power is continuously compared with the theoretical power using a Python-based digital twin framework.

The percentage deviation is given by:

$$\text{Deviation } (\%) = \frac{P_{\text{theoretical}} - P_{\text{actual}}}{P_{\text{theoretical}}} \times 100 \tag{14}$$

where:

- P_{actual} = measured power output (W)
- $P_{\text{theoretical}}$ = predicted power (W)

This continuous monitoring and comparison form the core of the performance monitoring framework, enabling trend visualisation and supporting predictive maintenance for both prototype and scalable wind turbine applications.

IV. Results and Discussions

The fabricated Savonius Vertical Axis Wind Turbine (VAWT) Prototype demonstrated successful electrical power generation, with the output voltage increasing proportionally with wind speed. This confirms the effective conversion of wind energy into electrical energy by the developed turbine system.

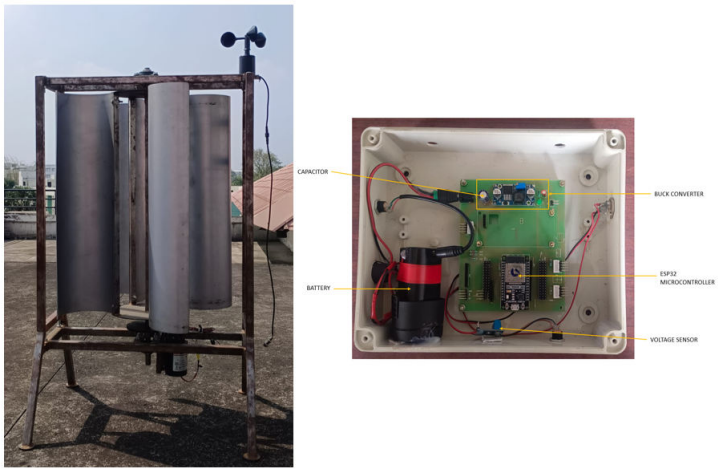


Fig. 5. Fabricated Savonius VAWT and DATA Acquisition system

The ESP32 microcontroller acquired real-time wind speed, voltage, and current data from the connected sensors, enabling continuous monitoring of turbine performance. The collected data was transmitted without interruption to the analysis platform for further processing.

The figures below [Fig. 6 & Fig. 7] shows the real-time variation of wind speed and the corresponding generated voltage of the wind turbine, illustrating the relationship between wind conditions and electrical output during system monitoring.

The Python-based data analysis platform accurately represented the expected operational behaviour of the wind turbine by estimating theoretical power output from measured wind

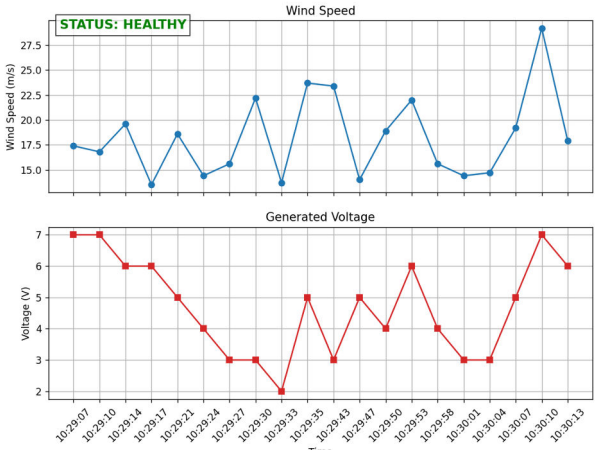


Fig. 6. Wind Speed and Voltage Correlation Graph-1

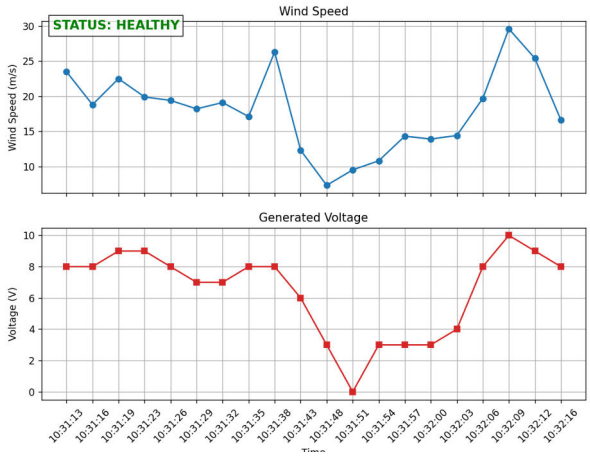


Fig. 7. Wind Speed and Voltage Correlation Graph-2

speed. Continuous comparison between the expected and actual outputs enabled effective performance evaluation of the turbine.

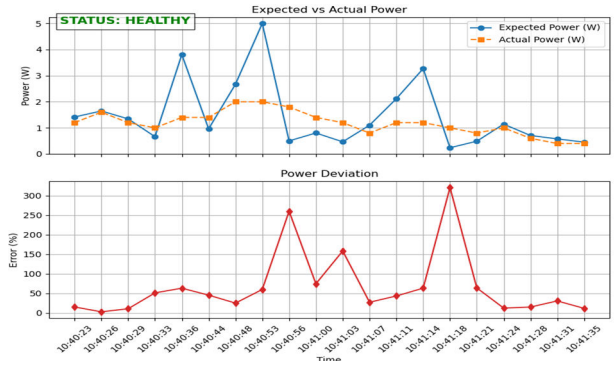


Fig. 8. Act vs Exp Power Graph without cut-in and cut-off wind specification

The figure [Fig. 8- Fig 11] shows the comparison between expected and actual power output of the wind turbine along

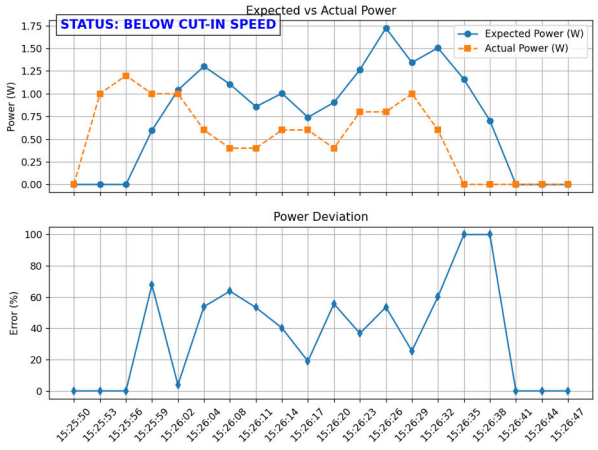


Fig. 9. Act vs Exp Power Graph with cut-in and cut-off wind specification

with the corresponding power deviation. The upper plot shows the variation of expected and measured power output, while the lower plot presents the percentage deviation between them. This analysis helps evaluate turbine performance and identify variations in power generation under changing wind conditions.

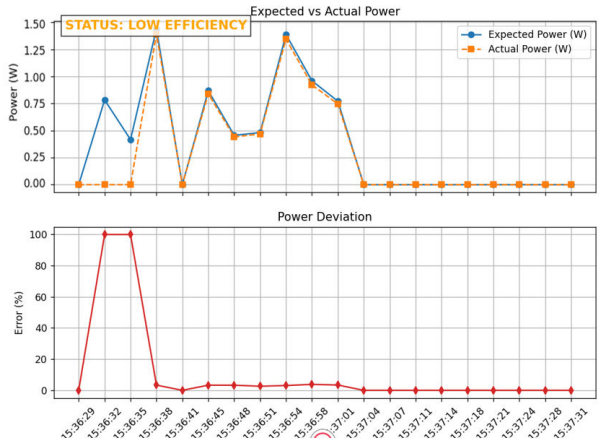


Fig. 10. Actual vs Expected Power Graph-1

For graphs (Fig. 10 & Fig. 11), the expected power was estimated by calculating the power coefficient (Cp) at selected wind speeds using the actual power data and the correlation between wind speed and generated voltage. These correlations and Cp values were implemented in a Python-based analysis code to compute the expected power accurately. The resulting data was then used to plot the expected versus actual power graph for performance evaluation.

Noticeable deviations between measured and expected values were successfully identified, indicating the presence of mechanical or electrical performance variations. Trend-based analysis and graphical visualisation clearly illustrated turbine performance under varying wind conditions. These results validate the effectiveness of the proposed continuous monitoring

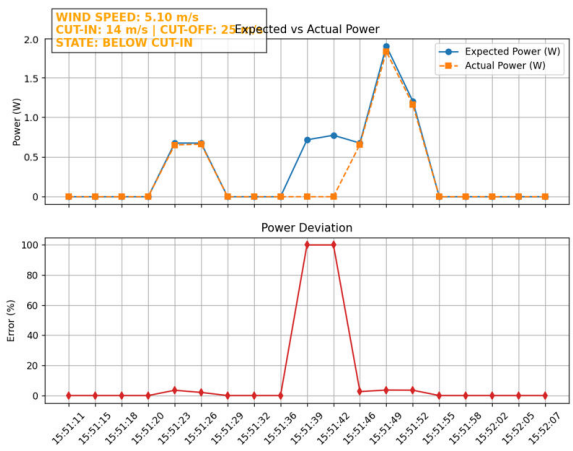


Fig. 11. Actual vs Expected Power Graph-2

framework for real-time performance assessment and early fault detection in wind energy systems.

V. Conclusion and Future Scopes

This paper presented the performance evaluation of a vertical-axis wind turbine. The system estimated the current consumption, actual electrical power output, and theoretical wind power. Parameters such as tip speed ratio (TSR) and power coefficient (Cp) were evaluated to analyse turbine performance under varying wind conditions. The continuous comparison between theoretical and actual power provided useful insights into system efficiency and highlighted deviations caused by mechanical, electrical, and environmental losses. Future work can focus on enhancing the system through the integration of advanced data analytics and machine learning techniques for predictive maintenance and fault detection. The monitoring framework can be extended with IoT-based cloud connectivity for remote access and real-time visualisation. Further optimisation of turbine design parameters and scaling of the model for larger or hybrid renewable energy systems can improve efficiency and expand its applicability in sustainable energy solutions.

References

[1] Fahim. M, Sharma. V, Cao. T. V, Canberk. B, and Duong. T. Q (2022). Machine Learning Based Digital Twin for Predictive Modeling in Wind Turbines. IEEE Access, 10, 14184-14192. https://doi.org/10.1109/ACCESS.2022.3147602.

[2] Mahmoud. M, Semeraro. C, Abdelkareem. M. A, and Olabi. A. G (2024). Designing and Prototyping the Architecture of a Digital Twin for a Wind Turbine. International Journal of Thermofluids, 22, 100622. https://doi.org/10.1016/j.ijft.2024.100622.

[3] Stadtmann. F, Rasheed. A, Kvamsdal. T, Johannessen. K. A, San. O, Kölle. K, Tande. J. O, Barstad. I, Benhamou. A, Brathaug. T, Christiansen. T, Firle. A. L, Fjeldly. A, Frøyd. L, Gleim. A, Høiberget. A, Meissner. C, Nygård. G, Olsen. J, Paulshus. H, Rasmussen. T. E, Rishoff. E, Scibilia. F, and Skogås. J. O (2023). Digital twins in wind energy: Emerging technologies and industry-informed future directions. IEEE Access, 11, 110762–110787. https://doi.org/10.1109/ACCESS.2023.3321320.

[4] Chang. C. C. W, Liu. Y. H, and Lin. M. H (2023). Recent advancements in condition monitoring systems for wind energy conversion systems: A review. Renewable and Sustainable Energy Reviews, vol. 175, pp. 1–20. https://doi.org/10.1016/j.rser.2023.113194.

[5] Chang. C. C. W, Liu. Y. H, and Lin. M. H (2023). Recent advancements in condition monitoring systems for wind energy conversion systems: A review. Renewable and Sustainable Energy Reviews, vol. 175, pp. 1–20. https://doi.org/10.1016/j.rser.2023.113194.

[6] Bhayo. M. A, Yatim. A. H. M, Khokhar. S, Aziz. M. J. A, and Idris. N. R. N (2015). Modeling of Wind Turbine Simulator for Analysis of the Wind Energy Conversion System Using MATLAB/Simulink. Conference Paper. https://doi.org/10.1109/ENCON.2015.7443315.

[7] Liu. Y, Zhang. J. M, Min. Y. T, Yu. Y, Lin. C, and Hu. Z. Z (2023). A digital twin-based framework for simulation and monitoring analysis of floating wind turbine structures. Ocean Engineering, 283, 115009. https://doi.org/10.1016/j.oceaneng.2023.115009.

[8] Ansuaya. P and Kumsuwan. Y (2013). Design of MATLAB/Simulink Modeling of Fixed-Pitch Angle Wind Turbine Simulator. Energy Procedia, 34, 362–370. https://doi.org/10.1016/j.egypro.2013.06.764.

[9] Ahmad. R. T and Abdul-Hussain. M. A (2017). Modeling and Simulation of Wind Turbine Generator Using Matlab-Simulink. Journal of Al Rafidain University College, (40), 282–284. https://iasj.rdd.edu.iq/journals/uploads/2024/12/.../3afbb2647cfedead1a06156097b56608.pdf.

[10] Hwang. S and Yoo. C, "Health monitoring and diagnosis system for a small H-type Darrieus vertical-axis wind turbine," Energies, vol. 14, no. 21, p. 7246, 2021. doi:10.3390/en14217246. https://www.mdpi.com/1996-1073/14/21/7246.

[11] Mohamed. H. E. A, Ahmed. A. A, Salama. A, and Ailelmagssabi. S. Y (2025). Wind Energy: an overview of the physical basics and the current types of technology. Eurasian Journal of Theoretical and Applied Sciences (EJTAS), 1(3), 18–28. https://eurasian-journals.com/index.php/etjas/index.

[12] Abdelsattar. M, Ismeil. M. A, Menoufi. K, Abdel Moety. A, and Emad-Eldeen. A (2025). Evaluating Machine Learning and Deep Learning models for predicting Wind Turbine power output from environmental factors. PLOS ONE, 20(1), e0317619. https://doi.org/10.1371/journal.pone.0317619.

[13] Bilendo. F, Thirion. J. M, and Guillaume. P (2022). Power curve-based fault detection method for wind turbines using statistical learning. Energy Reports, vol. 8, pp. 456–466. https://doi.org/10.1016/j.egyr.2021.12.202.

[14] Chen. Y and Lin. T (2013). Development of a small-scale vertical-axis wind turbine monitoring system using wireless sensor networks. International Journal of Distributed Sensor Networks, vol. 9, no. 4, pp. 1–10. https://journals.sagepub.com/doi/10.1155/2013/947986.

[15] El-Khozondar. H. J, El-Khozondar. R. J, and Khatib. M. J (2024). A smart energy monitoring system using ESP32 for renewable applications. Sustainable Energy Technologies and Assessments, vol. 62, pp. 1–9. https://doi.org/10.1016/j.prime.2024.100666.

[16] Al-Gburi. K. A. H, Al-Quraishi. B. A. A, Alnaimi. F. B. I, Tan. E. S, and Al-Safi. A. H. S (2022). Experimental and simulation investigation of performance of scaled model for a rotor of a Savonius wind turbine. Energies, vol.15, no.23, pp. 8808. https://www.mdpi.com/1996-1073/15/23/8808.

[17] Kurniati. S, Syam. S, and Nursalim (2025), "A novel elliptical Savonius turbine integrated with gearless axial generator," Journal Européen des Systèmes Automatisés, vol. 58, no. 2, pp. 271–282, 2025. [Online]. Available: https://doi.org/10.18280/jesa.580208.

STUDIES ON THE PRODUCTION OF COMPRESSED AIR USING WAVE ENERGY BASED AIR COMPRESSOR SYSTEM

Mohammed Noufal N[1], Akhilesh P[2],
Rebin Husain M[3], Naveen K[4]
Department of Mechanical Engineering
Ahalia School of Engineering and Technology
Palakkad, India
noufaln6235@gmsil.com, akhiprakash2004@gmail.com,
rebinfebin5@gmail.com, naveennavi906130@gmail.com

Dr. John Alexis S[5]
Professor, Department of Mechanical Engineering
Ahalia School of Engineering and Technology
Palakkad, India
John.Alexis@ahalia.ac.in,

***Abstract*—Ocean waves represent a significant source of renewable energy that can be harnessed for sustainable power generation. This work focuses on the design and fabrication of a wave-driven air compression system that converts the vertical motion of sea waves into useful compressed air energy. The system utilizes a floating buoy to capture wave motion, which is then transmitted through a mechanical power transmission arrangement including a chain and sprocket drive, gear mechanism, and a reciprocating compressor adapted from a discarded internal combustion engine. The air compressed by the system is collected and stored in a pressure tank for later use in pneumatic applications such as spray painting, cleaning operations, and small-scale industrial activities. Experimental testing shows that wave motion with an average height of approximately 0.7 m is sufficient to operate the mechanism and produce compressed air within the storage tank. The developed system demonstrates a cost-effective, environmentally friendly, and renewable approach for utilizing wave energy in coastal regions.**

Keywords: Wave energy, Renewable energy, Air compressor system, Buoy mechanism, Mechanical power transmission, Pneumatic energy storage, Ocean wave energy conversion.

I. Introduction

The rapid growth in global energy demand and the environmental impacts associated with conventional fossil fuels have increased the need for alternative and sustainable energy sources. Renewable energy technologies have become an important area of research because they offer environmentally friendly and long-term solutions for energy generation. Among the different renewable resources available, ocean wave energy is considered a promising source due to its high energy density, continuous availability, and large potential along coastal regions. Ocean waves are produced by wind transferring energy to the surface of the sea, resulting in oscillatory motion that contains both kinetic and potential energy. This energy can be converted into useful mechanical or electrical power using suitable energy conversion systems. Compared with other renewable resources such as solar and wind, wave energy has the advantage of being more predictable and consistent in many coastal areas. The purpose of this project is to develop a wave energy-based air compressor system that converts the vertical motion of ocean waves into compressed air energy. Instead of generating electricity, the proposed system directly utilizes mechanical motion to operate a reciprocating air compressor. This approach reduces system complexity and allows the energy to be stored in the form of compressed air, which can later be used for various pneumatic applications.

II. Experimental Setup

The experimental setup of the wave energy based air compressor system was developed to evaluate the ability of the prototype to convert wave motion into compressed air. The setup consists of a buoy mechanism, motion transmission system, compressor unit, air storage tank, and measuring instruments mounted on a rigid supporting frame. A floating buoy made of PVC material was placed in water to capture the vertical motion generated by waves. The buoy moves up and down according to the rise and fall of the water surface. This vertical movement acts as the input energy for the system. The buoy is connected to a chain which transfers the motion to a sprocket mounted on a rotating shaft. The chain and sprocket arrangement converts the vertical oscillating motion of the buoy into rotational motion. This rotation is transmitted through a gear mechanism consisting of spur gears and pinion gears. The gear assembly ensures smooth transmission of motion and provides the required torque to drive the compressor unit. The compressor used in the setup is a modified piston-cylinder mechanism obtained from a discarded internal combustion engine. The rotational motion from the transmission shaft drives the crank mechanism of the compressor, converting rotary motion into reciprocating motion of the piston. During the suction stroke, atmospheric air enters the cylinder through the inlet valve. During the compression stroke, the piston compresses the air inside the cylinder. The compressed air produced by the compressor is delivered through a non-return valve into an air storage tank. The non-return valve prevents backflow of compressed air

from the tank to the compressor. A pressure gauge is installed on the storage tank to measure the pressure of the compressed air generated during operation. During the experimental testing, the system was operated under wave conditions with an average wave height of approximately 0.7 m. The movement of the buoy produced continuous mechanical motion that drove the compressor mechanism. The pressure developed in the air storage tank was observed and recorded over time to evaluate the performance of the system.

III. Design and Fabrication

The design and fabrication of the wave energy based air compressor system were carried out to effectively convert the vertical motion of ocean waves into compressed air energy. The overall design focuses on creating a simple, durable, and cost-effective mechanical system capable of operating under fluctuating wave conditions. The structure was designed to ensure proper alignment of components, smooth transmission of motion, and reliable compression of air.

The system consists of several important mechanical components including a floating buoy, chain and sprocket mechanism, gear transmission, compressor unit, air storage tank, and supporting frame structure. Each component was selected based on its strength, availability, and suitability for operation in a marine environment. A floating buoy was used as the primary energy capturing element of the system. The buoy is designed to float on the water surface and move vertically with the motion of the waves. Lightweight and corrosion-resistant materials such as PVC are preferred for the buoy to ensure good buoyancy and durability. The vertical movement of the buoy serves as the input motion for the mechanical transmission system.

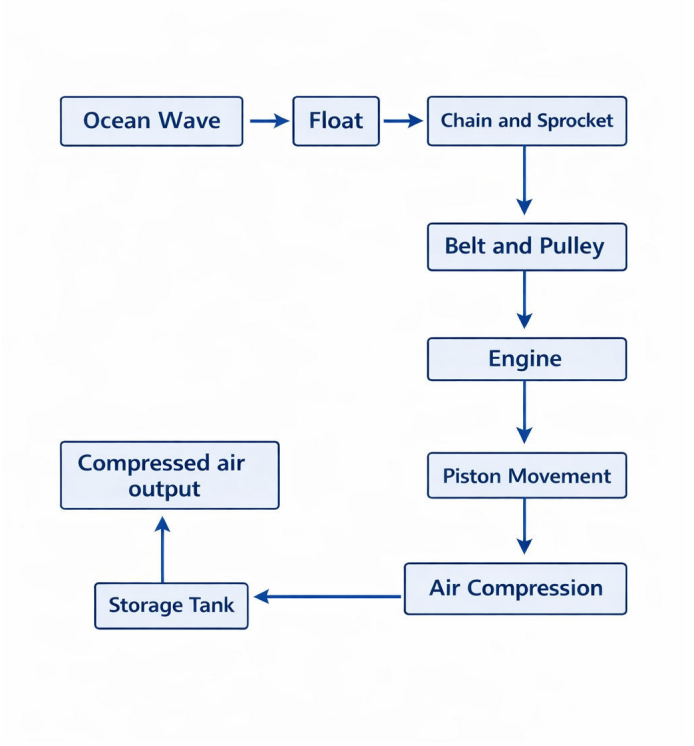
The motion captured by the buoy is transmitted through a chain and sprocket arrangement. This mechanism converts the vertical oscillating motion of the buoy into rotational motion. Chain and sprocket drives are commonly used in mechanical power transmission due to their ability to transmit motion without slipping and their capability to handle variable loads. To further transmit the motion to the compressor, spur gears and pinion gears are used. These gears help transfer rotational motion between shafts while maintaining proper speed and torque. The gear system ensures smooth mechanical operation and improves the efficiency of motion transfer from the buoy mechanism to the compressor unit.

The compression unit of the system is fabricated using a repurposed internal combustion engine cylinder and piston assembly. Instead of using the engine for fuel combustion, the piston-cylinder mechanism is utilized as a reciprocating air compressor. The crank mechanism connected to the piston converts the rotational motion from the transmission system into reciprocating motion required for air compression.

The exhaust nozzle is attached to the outlet of the combustion chamber. The high-pressure and high-temperature gases produced during detonation combustion expand through the nozzle and are accelerated to very high velocities. This conversion of thermal energy into kinetic energy produces thrust according to Newton's third law of motion.

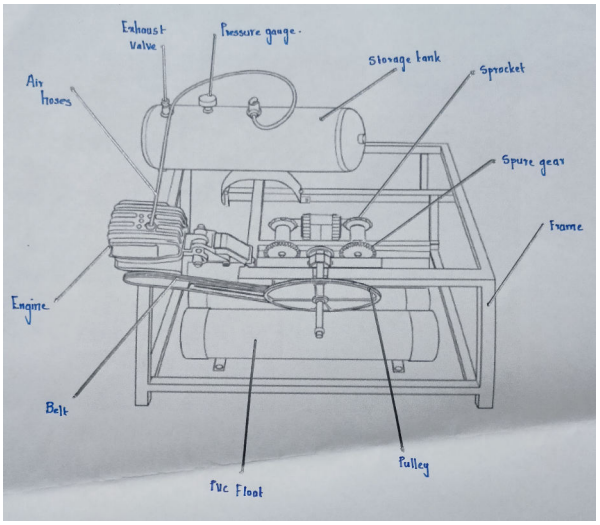
IV. Working Principle

The working principle of the wave energy based air compressor system is based on converting the vertical motion of ocean waves into mechanical energy and then transforming that mechanical energy into compressed air. Ocean waves produce continuous up and down movement on the water surface, and this motion is utilized as the primary source of energy for the system. In this system, a floating buoy is placed on the surface of the water. When waves pass through the buoy, it moves vertically according to the rise and fall of the waves. This vertical oscillating motion of the buoy acts as the input motion for the energy conversion mechanism. The buoy is connected to a mechanical linkage that transfers this motion to a chain and sprocket arrangement.



The chain and sprocket mechanism converts the vertical movement into rotational motion. As the buoy moves upward due to wave action, the chain pulls the sprocket, causing it to rotate. When the buoy moves downward, the mechanism continues the rotational movement through the mechanical transmission system. This rotation is then transmitted through gears and shafts to the compressor unit. The compressor used in the system is a modified piston-cylinder assembly obtained from a repurposed internal combustion engine. The rotary motion from the transmission system drives the crank mechanism of the compressor, which converts the rotary motion into reciprocating motion of the piston. During the suction stroke, the piston moves downward, allowing atmospheric air to enter

the cylinder through the inlet valve. During the compression stroke, the piston moves upward and compresses the air inside the cylinder.



The compressed air is forced through a non-return valve into an air storage tank. The non-return valve ensures that the compressed air flows only in one direction and prevents backflow into the compressor. As the wave motion continues, the compressor repeatedly compresses air and gradually increases the pressure inside the storage tank. A pressure gauge is attached to the tank to monitor the pressure level of the stored air. The stored compressed air can then be used for various pneumatic applications such as cleaning, spray painting, and operating pneumatic tools. Thus, the system effectively converts wave energy → mechanical motion → compressed air energy, providing a renewable and environmentally friendly method of utilizing ocean wave energy.

V. Advantages and Challenges

The wave energy based air compressor system offers several advantages as a renewable energy technology. One of the major benefits is that it utilizes ocean wave energy, which is a naturally available and sustainable energy source. Unlike fossil fuels, wave energy does not produce harmful emissions, making the system environmentally friendly. Another important advantage of the system is its simple mechanical design. The system mainly consists of mechanical components such as a buoy, chain and sprocket mechanism, gears, and a piston-cylinder compressor. Because of this simple construction, the system can be fabricated using commonly available materials and standard mechanical components.

Despite its advantages, the system also faces several challenges. One of the main limitations is the dependence on wave conditions. The performance of the system is directly influenced by wave height, wave frequency, and ocean conditions. When wave activity is low, the amount of compressed air produced by the system may also decrease. Another challenge is related to mechanical losses. Friction in the transmission components such as chains, gears, and bearings can reduce the overall efficiency of the system. Proper lubrication and maintenance are required to minimize these losses.

VI. Result and Discussion

A. Air Volume per Stroke

Piston diameter D = 50 mm = 0.05 m
Stroke length L = 48.8 mm = 0.0488 m

$$Formula = \frac{\pi}{2} D^2 L \tag{1}$$

$$\frac{3.14}{4}(0.05)^2(0.0488) = 0.0000958m^2 \tag{2}$$

B. Strokes per Minute

3 waves per minute
35 strokes per wave

$$= 3 \times 35 = 105 strokes/min \tag{3}$$

C. Air Delivery

$$Formula : Q = V_s \times N \tag{4}$$

Where.
V_s = volume per stroke
N = strokes per minute

$$Q = 0.096 \times 105 \tag{5}$$

$$FinalAirDelivery(Q) = 0.01008m^3 \tag{6}$$

D. Total Air Produced in 40 Minutes

= 0.01008 × 40 = 0.403 m^3

$$Efficiency = \frac{aircompressionwork}{Waveinputwork} \tag{7}$$

E. Input Energy from Waves

Wave force on float, F = 95 N
Wave height (float movement), h = 0.7 m
Work per wave
W = F × h = 95 × 0.7 = 66.5 J
3 wave/minu
Total waves in 40 minutes = 3 × 40 = 120
Total input energy, E_{in} = 66.5 × 120 = 7980 J

F. Useful Energy (Compressed Air)

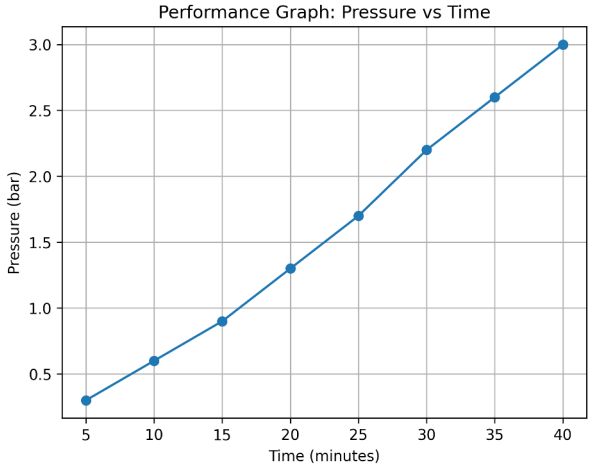
Final pressure
3 bar = 300000 Pa
Total air volume produced
0.01008 × 40 = 0.403 m^3
Convert to cubic meter
403L = 0.403 m^3
Energy stored in compressed air
E = P × V
= 30000 × 0.403 = 120900 J

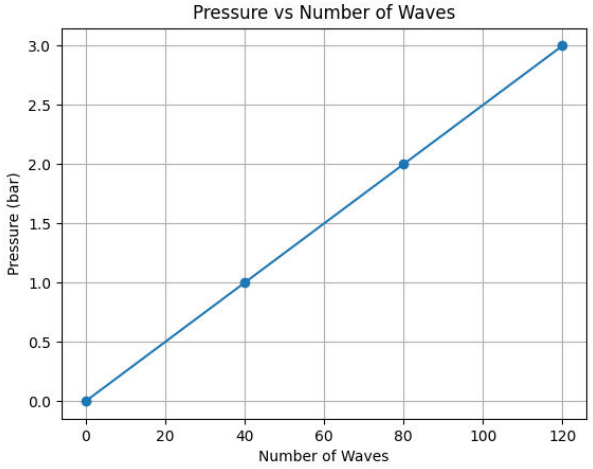
G. Efficiency

$$Efficiency = \frac{Eout}{Ein}100 \quad (8)$$

= (120900/7980) × 100 = 28%

The system was tested with an average wave height of 0.7 m, producing approximately 3 waves per minute. Each wave generated about 35 strokes, resulting in a total of 105 compressor strokes per minute. The compressor had a piston diameter of 50 mm and a stroke length of 48.8 mm, enabling effective air compression.





The pressure development in the storage tank was monitored during operation. The system produced 1 bar pressure in 13.3 minutes, 2 bar in 26.6 minutes, and reached a final pressure of 3 bar in 40 minutes. The air delivery rate of the compressor was approximately 10.08 L/min, and the total compressed air produced during the test was about 403 liters. Based on the experimental air output and estimated wave energy input, the overall practical efficiency of the prototype system was approximately 28 Percentage. The results confirm that the developed prototype successfully converts ocean wave motion into compressed air energy and demonstrates the feasibility of using wave energy for small-scale pneumatic applications in coastal regions.

VII. Conclusion

This project successfully demonstrated the design and development of a wave-powered air compressor system that converts ocean wave motion into useful compressed air energy.

The floating buoy mechanism effectively transformed the vertical movement of waves into reciprocating mechanical motion, which operated the compressor through a mechanical transmission system.

Experimental testing showed that with an average wave height of 0.7 m and 3 waves per minute, the system produced a total of 105 compressor strokes per minute. The compressor, with a 50 mm piston diameter and 48.8 mm stroke length, was capable of generating compressed air that gradually increased the tank pressure to 3 bar within 40 minutes.

References

[1] A. F. de O. Falcão, "Wave energy utilization: A review of the technologies," Renewable and Sustainable Energy Reviews, vol. 14, no. 3, pp. 899–918, 2010.

[2] J. Falnes, Ocean Waves and Oscillating Systems: Linear Interactions Including Wave-Energy Extraction. Cambridge, U.K.: Cambridge University Press, 2002.

[3] M. E. McCormick, Ocean Wave Energy Conversion. New York, USA: John Wiley Sons, 1981.

[4] M. E. McCormick, Ocean Wave Energy Conversion. New York, USA: John Wiley Sons, 1981.

[5] T. J. Hammons, "A review of tidal and wave energy," International Journal of Electrical Power Energy Systems, vol. 26, no. 2, pp. 81–92, 2004.

[6] P. A. Lynn, Electricity from Wave and Tide: An Introduction to Marine Energy. London, U.K.: Wiley, 2013.

[7] C. Boylestad and L. Nashelsky, Mechanical Engineering Design and Applications. New Delhi, India: Pearson Education, 2015.

[8] R. K. Rajput, A Textbook of Fluid Mechanics and Hydraulic Machines. New Delhi, India: S. Chand Publishing, 2016.

[9] B. S. Nagendra and K. K. Patil, "Design and development of wave powered air compression system," International Journal of Mechanical Engineering Research, vol. 9, no. 4, pp. 145–150, 2019.

Parametric Study on the Vibration Characteristics of Neoprene Rubber Isolators

Anna Inasu Chungath
Department of Mechanical Engineering
Christ College of Engineering
Thrissur, India
annainasu@gmail.com

Charles Michael Augustine
Department of Mechanical Engineering
Christ College of Engineering
Thrissur, India
charlesmichaelaugustine@gmail.com

Dawn Cletus Francis
Department of Mechanical Engineering
Christ College of Engineering
Thrissur, India
pekkaclan007@gmail.com

Srutheesh T P
Department of Mechanical Engineering
Christ College of Engineering
Thrissur, India
tpsrutheesh5@gmail.com

Rojin Mathews
Department of Mechanical Engineering
Christ College of Engineering
Thrissur, India

***Abstract*— This paper studies the influence of geometric configuration on the vibration isolation performance of neoprene-based elastomers. Conventional isolator designs primarily focus on material properties, while the effect of geometry on dynamic behaviour is often overlooked. In this work, three different geometries such as square blocks, ring, and strips were designed with a constant material volume of 72,000 mm³ and uniform thickness to ensure a fair comparison.

A numerical modal analysis was performed using ANSYS to determine the natural frequencies under a constant payload of 11 kg. The results showed a significant shift in natural frequency from 16.44 Hz for the square configuration to 21.68 Hz for the strip configuration. To understand this variation of natural frequency, the concept of shape factor was employed, which had a direct relationship between geometry, stiffness, and natural frequency.

Experimental validation was carried out on the square configuration using a vibration shaker setup. The measured resonance frequency of approximately 16.0 Hz closely matched the numerical result, confirming the accuracy of the simulation model. The study demonstrates that geometric modification alone can effectively tune the vibration response without altering material properties or mass.

These findings highlight the importance of geometry-driven design in elastomeric isolators and provide a practical approach for optimizing vibration isolation performance in engineering applications.

Keywords— Vibration isolation, Neoprene Rubber, Shape Factor, Natural Frequency, Transmissibility.

I. INTRODUCTION

Vibration is a critical issue in mechanical and structural systems, leading to noise, reduced performance, and potential structural failure. To mitigate these effects, vibration isolators are widely used to reduce the transmission of dynamic forces from the source to the supporting structure. Among various isolation materials, elastomers such as neoprene rubber are extensively used due to their high damping capacity, flexibility, durability, and resistance to environmental factors such as temperature, oil, and oxidation [1].

The performance of a vibration isolator is primarily influenced by parameters such as stiffness, natural frequency, and transmissibility. Stiffness determines the resistance of the isolator to deformation under load, while the natural frequency defines the frequency at which the system tends to vibrate when disturbed. Transmissibility represents the ratio of output vibration to input vibration and is a critical measure of isolation effectiveness. These parameters are inherently interrelated; an increase in stiffness leads to an increase in natural frequency, which directly influences the transmissibility characteristics of the system [10]. Traditional design approaches focus mainly on material selection and loading conditions, while the influence of geometric configuration is often not thoroughly explored. However, in elastomeric systems, geometry plays a crucial role in controlling deformation behavior, especially due to the nearly incompressible nature of rubber materials, where lateral bulging significantly affects stiffness [3][4].

One of the key parameters used to characterize this behavior is the shape factor, defined as the ratio of loaded area to the area available for lateral bulging. An increase in shape factor restricts bulging, resulting in higher compression modulus and stiffness, which in turn affects the natural frequency of the system [2][5].

In the present study, the effect of geometric configuration on vibration isolation performance is systematically investigated using three different neoprene geometries that are square blocks, ring, and strip configurations while maintaining constant material volume and thickness. Numerical modal analysis is carried out using ANSYS Workbench to determine the natural frequencies of the system under a constant load condition. Subsequently, experimental validation is performed using a vibration shaker setup, where the natural frequency obtained from the transmissibility curve of the square configuration is compared with the numerical results obtained from ANSYS Workbench to ensure model accuracy. The objective of this work is to demonstrate that geometric modification alone can be effectively used to tune the dynamic response of elastomeric isolators without altering material properties or mass. The findings of this study have practical significance in the design of vibration isolation systems used in applications such as machinery mounting, automotive components, and structural vibration

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

control, where performance optimisation and cost efficiency are critical.

II. METHODOLOGY

The methodology adopted in this study follows a systematic approach to evaluate the influence of geometric configuration on the vibration isolation performance of neoprene-based elastomeric systems. The study consists of geometric modelling, numerical analysis, analytical evaluation, and experimental validation to ensure both theoretical and practical accuracy

A. Geometry Design

Three different geometric configurations that are square blocks, strips, and ring were designed using neoprene rubber. All configurations were developed with a constant material volume of 72,000 mm³ and a uniform thickness of 12 mm to ensure a fair comparison.

The square configuration consists of four identical blocks, each having dimensions of 30 mm × 30 mm × 12 mm. The strip configuration consists of two rectangular strips of dimensions 25 mm × 150 mm × 12 mm.

For the ring geometry, an inner diameter of 50 mm was selected. To maintain constant volume, the outer diameter was calculated based on the required cross-sectional area, resulting in an outer diameter of approximately 100.7 mm.

These geometries were specifically chosen to produce distinct deformation characteristics under compressive loading, primarily due to differences in their shape factor and lateral bulging behaviour.

In the square configuration, each block has a relatively small loaded area and a large free surface area around its sides. Due to the nearly incompressible nature of neoprene, when compressive load is applied, the material tends to expand laterally. Since the square blocks have minimal geometric constraints along their edges, lateral bulging occurs freely, resulting in lower resistance to deformation and hence lower stiffness.

In contrast, the strip configuration has a high aspect ratio, with a significantly larger length compared to its width. This geometry increases the loaded area while reducing the proportion of free surfaces available for lateral expansion along the longer direction. As a result, lateral bulging is partially restricted, leading to higher internal resistance to deformation. This restriction increases the effective compression modulus and consequently the stiffness of the system.

The ring configuration provides a continuous circular geometry, where the material is distributed uniformly around the loading axis. Unlike discrete blocks, the ring does not have free edges in the circumferential direction, which alters the bulging pattern. The deformation is distributed more uniformly along the circumference, resulting in a moderate level of bulging restriction. This leads to stiffness values that lie between the square and strip configurations.

Thus, by maintaining constant volume and varying only the geometric layout, the study isolates the influence of geometry on deformation behaviour and stiffness. This enables a clear understanding of how geometric configuration alone can be used to control the dynamic response of elastomeric vibration isolators.

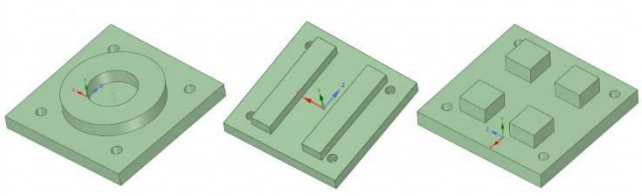


Fig. 1. Geometry models of square, ring, and strip configurations.

B. Numerical Modelling

The numerical analysis was carried out using ANSYS Workbench to evaluate the dynamic behaviour of the neoprene-based vibration isolators. A finite element modelling approach was adopted to simulate the deformation and vibration characteristics of different geometric configurations under identical loading conditions.

Neoprene was modelled as a nearly incompressible elastomeric material, with a Poisson's ratio of 0.49 to capture its resistance to volumetric change during compression. The material density was taken as 1230 kg/m³, and an equivalent Young's modulus of 1 MPa was used to represent its elastic response under small strain conditions. These properties were assigned to all neoprene components in the model. The numerical model consisted of a top and bottom plate assembly with neoprene elements placed in between. Both the top and bottom plates were modelled as structural steel to replicate the actual experimental setup. The interaction between the neoprene and steel plates was defined using a bonded contact condition, ensuring that no relative sliding or separation occurs at the interface.

A remote displacement boundary condition was applied to the bottom support face. In this condition, all translational and rotational degrees of freedom were constrained to zero except for the vertical (Y-axis) displacement, which was kept free. This setup simulates the excitation behaviour of a vibration shaker, where motion is permitted only in the vertical direction while restricting all other movements. The use of a rigid remote condition ensures uniform transmission of displacement across the selected surface.

The mass of the system was introduced through the top plate, which was modelled as a structural steel component with dimensions of 150 mm × 150 mm × 32 mm, resulting in an approximate mass of 6 kg. This approach ensures realistic representation of inertial effects, as the mass is physically distributed rather than applied as a lumped parameter. Since the natural frequency of the system depends on the mass–stiffness relationship, accurate modelling of mass is essential for reliable results. A modal analysis was performed to determine the natural frequencies and mode shapes of the system for each geometric configuration. The first mode corresponded to a rigid body motion with a near-zero frequency, arising due to the absence of restoring stiffness in that mode. Therefore, it does not represent the actual vibration behaviour of the system. The second mode was identified as the first significant deformation mode,

characterised by vertical motion of the top plate, and was considered as the fundamental mode for evaluating the dynamic response of the system. A refined finite element mesh was generated for all configurations to accurately capture the deformation behaviour of the elastomer. Three-dimensional solid elements were used to model the neoprene geometry, along with contact elements at the interfaces. The mesh statistics for each configuration are summarised in Table I below:

TABLE I. MESH STATISTICS FOR DIFFERENT CONFIGURATIONS

Configuration	Nodes	Solid Elements	Contact Elements	Total Elements
Square	7840	1374	651	2026
Strip	8352	1472	886	2359
Ring	8382	1410	1027	2438

The mesh density was selected to ensure accurate stress and deformation prediction while maintaining computational efficiency. A consistent meshing approach was followed across all configurations to ensure comparability of results.

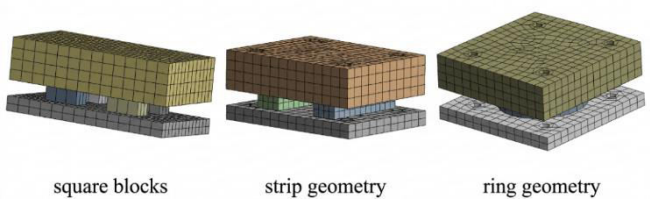


Fig. 2. Finite element model and mesh configuration.

Overall, the numerical modelling framework was developed to replicate the physical system as closely as possible, incorporating realistic material behaviour, boundary conditions, loading conditions, and contact interactions. This ensured that the obtained natural frequencies accurately represent the dynamic behaviour of the neoprene isolator system.

C. Shape Factor and Stiffness Evaluation

The variation in stiffness due to geometry was analysed using the concept of shape factor. The shape factor (S) is defined as the ratio of loaded area to the area available for lateral bulging, as shown in (1),

$$S = \text{Loaded Area} / \text{Bulge Area} \quad (1)$$

The compression modulus of the elastomer is related to the shape factor as shown in (2):

$$E_c = E_0 (1 + 2S^2) \quad (2)$$

The stiffness of the isolator is given by (3),

$$k = (E_c \times A) / t \quad (3)$$

The natural frequency of the system is calculated using (4) ,

$$f_n = (1 / 2\pi) \sqrt{(k / m)} \quad (4)$$

These relationships establish the connection between geometry, stiffness, and vibration behaviour.

The calculated geometric parameters for each configuration are summarised in Table II.

TABLE II. SHAPE FACTOR AND GEOMETRIC PARAMETERS

Configuration	Loaded Area (mm^2)	Bulge Area (mm^2)	Shape Factor (S)
Square	900	1440	0.625
Ring	6000	5346	1.05
Strip	3000	3480	0.862

The values presented in Table 2 are used to evaluate stiffness variation across different geometries

D. Experimental Setup

Experimental validation was performed using a vibration shaker system to verify the numerical results obtained from modal analysis. The setup consisted of a top and bottom structural steel plate assembly with neoprene specimens placed between them. The top plate had dimensions of 150 mm × 150 mm with a thickness of 32 mm, while the bottom plate had a thickness of 12 mm. The neoprene isolators were maintained at a uniform thickness of 12 mm to ensure consistency with the numerical model.

The experiment was conducted using a Data Physics GW V20 vibration shaker. The excitation signal was generated using a function generator, where the desired frequency input was defined. This signal was passed through a power amplifier to increase its strength before being supplied to the shaker. The shaker converted the electrical signal into mechanical vibrations, producing controlled harmonic excitation of the system.

The frequency range for testing was selected from 10 Hz to 35 Hz, based on the expected natural frequency obtained from numerical analysis (around 16 Hz). This range was chosen to ensure that the resonance region was fully captured, along with sufficient data points before and after resonance for accurate transmissibility analysis. The excitation was applied in discrete frequency steps, and corresponding response values were recorded at each step.

Non-contact laser displacement sensors were used to measure the system response. One sensor was directed at the bottom plate to capture the input displacement, while another sensor was aimed at the top plate to measure the output displacement. The signals obtained from the sensors were transmitted to a data acquisition unit (C-box), and subsequently processed using SensorTool software for analysis.

This experimental arrangement ensures accurate measurement of input and output vibrations without physical contact, thereby eliminating measurement interference. The setup closely replicates the boundary conditions used in the numerical model, enabling reliable validation of the simulated results.

Fig. 3. Experimental setup showing shaker, plate assembly, and sensor arrangement.

The transmissibility (T) of the system was calculated as the ratio of output displacement to input displacement:

$$T = X_{OUT} / X_{IN} \quad (5)$$

The experimental data obtained from the vibration shaker was recorded in time-domain format. At each excitation frequency, the peak-to-peak displacement values of the input (bottom plate) and output (top plate) were extracted.

The processed data was then imported into MATLAB for further analysis. MATLAB was used to compute the transmissibility values at each frequency and to generate a continuous transmissibility curve through interpolation of discrete data points.

This curve was used to identify the resonance frequency and to compare the experimental results with numerical predictions for validation.

III. RESULTS AND DISCUSSION

The results obtained from both numerical analysis and experimental testing are presented in this section. The primary focus is on the variation in natural frequency due to changes in geometric configuration and the validation of numerical results through experimental observations.

A. Modal Analysis Results

The natural frequencies of the three geometrical configurations were obtained through modal analysis using ANSYS Workbench. A total of the first six modes were extracted for each configuration to understand the dynamic behaviour of the system. Among these, the first mode corresponded to a rigid body motion with near-zero frequency and was therefore not considered for analysis. The second mode represented the first significant deformation mode, characterised by vertical motion of the top plate, and was identified as the fundamental mode governing the vibration response of the system.

TABLE 3: NATURAL FREQUENCIES OF DIFFERENT GEOMETRIES

Geometry	Natural Frequency (Hz)
Square Blocks	16.44
Ring Geometry	18.06
Strip Geometry	21.68

The results presented in Table 3 indicate a clear variation in natural frequency with a change in geometry, despite maintaining constant material volume and loading conditions. Since the mass of the system remains constant, this variation directly reflects the change in effective stiffness of each configuration.

The square configuration exhibited the lowest natural frequency of 16.44 Hz, indicating the least stiffness among the three geometries. This behaviour can be explained using the concept of shape factor. The square blocks have a relatively lower shape factor due to a higher proportion of free surface area available for lateral bulging. Since neoprene is nearly incompressible, this increased bulging reduces resistance to deformation, resulting in lower effective stiffness and hence lower natural frequency.

The strip configuration showed the highest natural frequency of 21.68 Hz, indicating increased stiffness. This is attributed to its higher shape factor, where the available area for lateral bulging is relatively reduced compared to the loaded area. The restriction of lateral deformation increases the compression modulus, thereby increasing stiffness and natural frequency.

The ring geometry produced an intermediate natural frequency of 18.06 Hz. Although it exhibits a higher shape factor than the square configuration, the deformation is distributed uniformly across the circular geometry. This results in moderate bulging restriction and stiffness, placing its natural frequency between the square and strip configurations.

Overall, the modal analysis results clearly demonstrate that geometric configuration directly influences on system stiffness and natural frequency. This establishes geometry as a critical design parameter in tuning the dynamic performance of elastomeric vibration isolators.

B. Experimental Results

Experimental testing was conducted on the square configuration using a vibration shaker setup to validate the numerical results. The response of the system was measured over a frequency range of 10 Hz to 35 Hz using non-contact laser displacement sensors.

The data acquisition was carried out using SensorTool software, which recorded the displacement response of both the bottom plate (input) and top plate (output) in the form of time-domain signals. The recorded data was exported as Excel files containing time-stamped displacement values for both input and output signals at each excitation frequency.

The collected data was further processed using MATLAB. For each frequency step, the peak displacement values of the input and output signals were extracted. The transmissibility (T) was calculated as the ratio of output displacement to input displacement.

To generalize the results, the excitation frequency was normalized using the frequency ratio ($r = f / f_n$), where f_n is the natural frequency obtained from numerical analysis. The transmissibility values were then plotted against the frequency ratio using Microsoft Excel to obtain the experimental transmissibility curve.

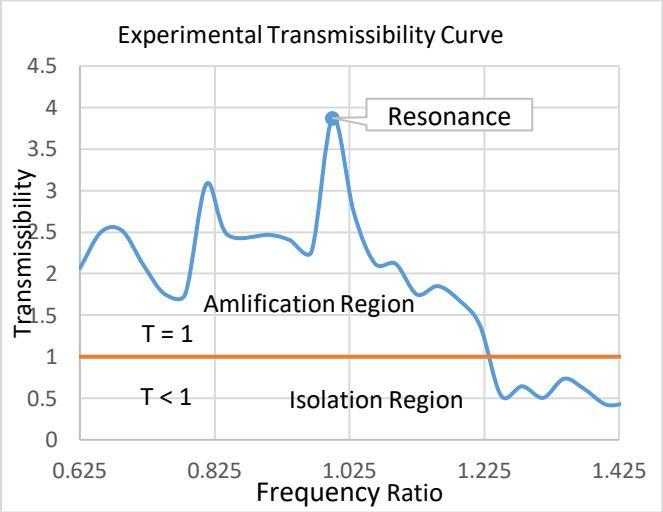


Fig. 4. Experimental transmissibility curve showing resonance behaviour.

The transmissibility curve exhibits a clear resonance peak near a frequency ratio of $r \approx 1$, corresponding to the natural frequency of approximately 16.0 Hz. Beyond this point, the transmissibility decreases and falls below unity ($T < 1$), indicating the onset of the vibration isolation region. The experimental results show close agreement with the numerical prediction of 16.44 Hz, confirming the validity of the developed model. This agreement indicates that the selected material properties, boundary conditions, and loading assumptions accurately represent the physical behaviour of the system. Minor deviations can be attributed to practical factors such as material inconsistencies, bonding imperfections, and measurement limitations.

The transmissibility curve further demonstrates that effective vibration isolation occurs when the operating frequency exceeds the natural frequency, where the transmissibility drops below unity ($T < 1$). This confirms that the isolator successfully reduces the transmission of vibrations beyond the resonance region.

Overall, the results establish that geometric configuration plays a decisive role in controlling both stiffness and isolation performance. By modifying the shape of the neoprene isolator, the natural frequency can be shifted, thereby enabling the system to operate within the desired isolation region without altering material properties. This highlights geometry as a key design parameter for optimising vibration isolation in engineering applications.

IV. CONCLUSION

This study investigated the influence of geometric configuration on the vibration isolation performance of neoprene-based elastomeric systems under constant material volume and loading conditions. Three different geometries were square, ring, and strip which were analysed to isolate the effect of shape on dynamic behaviour.

The results clearly demonstrate that geometric configuration significantly affects system stiffness and natural frequency. The square configuration exhibited the lowest natural frequency (16.44 Hz), indicating lower stiffness and making it suitable for low-frequency isolation applications. In contrast, the strip configuration showed the highest natural frequency (21.68 Hz), resulting from greater restriction to lateral deformation and increased stiffness. The ring geometry produced an intermediate response (18.06 Hz), indicating that, in addition to shape factor, the overall deformation pattern also influences system behaviour.

A comparison between numerical and experimental results for the square configuration showed a natural frequency of 16.44 Hz from simulation and approximately 16.0 Hz from experimental testing. This close agreement confirms the accuracy of the modelling approach and validates the assumptions related to material properties, boundary conditions, and loading conditions.

The transmissibility analysis further revealed that effective vibration isolation is achieved when the excitation frequency exceeds the natural frequency of the system, where transmissibility falls below unity ($T < 1$). This confirms that shifting the natural frequency through geometric modification directly influences the isolation performance.

Overall, the study establishes that geometry is a critical design parameter for tuning both stiffness and vibration isolation characteristics of elastomeric systems. The findings demonstrate that desired isolation performance can be achieved through geometric optimization alone, without altering material properties or increasing system mass. This provides a practical and efficient approach for the design of vibration isolation systems in engineering applications.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Mr. Rojin Mathews for his guidance and support throughout this work. The authors also acknowledge Christ College of Engineering for institutional support and the Vibration Testing Laboratory at the National Institute of Technology Calicut for facilitating the experimental studies.

REFERENCES

[1] Gent AN. *Engineering with Rubber: How to Design Rubber Components*. Hanser Publishers. 2012.

[2] Gent AN, Lindley PB. The compression of bonded rubber blocks. *Proceedings of the Institution of Mechanical Engineers*. 1959; 173: 111–122.

[3] Kumar P, Singh R. Dynamic characteristics of elastomeric vibration isolators. *Journal of Vibration and Control*. 2017.
[4] Zhang Y, et al. Analysis of rubber isolators under dynamic loading. *Mechanical Systems and Signal Processing*. 2018.
[5] Chen W, et al. Effect of geometry on stiffness of elastomeric bearings. *Engineering Structures*. 2019.
[6] Miehe C, et al. Computational modelling of rubber-like materials. *Computer Methods in Applied Mechanics and Engineering*. 2016.
[7] Patel H, Shah D. Finite element analysis of elastomeric pads for vibration isolation. *Materials Today: Proceedings*. 2020.
[8] Li X, et al. Experimental study on vibration isolation performance of rubber systems. *Journal of Sound and Vibration*. 2021.
[9] Wang L, et al. Recent advances in elastomer-based vibration isolation systems. *Composite Structures*. 2022.
[10] Rao SS. *Mechanical Vibrations*. 6th ed. Pearson Education. 2017.

Experimental Investigation of an Innovative System for Solar Desalination

Amrithesh Krishna R
dept. Mechanical Engineering
Ahalia School of Engineering and Technology
Palakkad, India
amritheshkrishna@ahalia.ac.in

Arjun S
dept. Mechanical Engineering
line 3: *Ahalia School of Engineering and Technology*
Palakkad, India
arjun.s@ahalia.ac.in

Shabith A
dept. Mechanical Engineering
line 3: *Ahalia School of Engineering and Technology*
Palakkad, India
shabith_a@ahalia.ac.in

Prabijith R
dept. Mechanical Engineering
line 3: *Ahalia School of Engineering and Technology*
Palakkad, India
prabijith .r@ahalia.ac.in

Dr. Bhavin K Bharath
dept. Mechanical Engineering
line 3: *Ahalia School of Engineering and Technology*
Palakkad, India
bhavin.bharath@ahalia.ac.in

Dr.P R Suresh
dept. Mechanical Engineering
line 3: *Ahalia School of Engineering and Technology*
Palakkad, India
principal.aset@ahalia.ac.in

***Abstract*—Water scarcity is becoming a major problem around the world due to rapid population growth, climate change, and the limited availability of fresh water. One promising way to address this issue is desalination, which converts salty water into drinking water, but common methods like reverse osmosis and multi-stage flash distillation consume a lot of energy. To overcome this, solar still desalination offers a simpler and more eco-friendly solution by using sunlight to purify water. In this study, a basic single-basin solar still was first tested to understand how it performs under normal conditions. Then, improvements were made to increase its efficiency, such as adding paraffin wax as a phase change material (PCM) to store heat and using fins inside the basin to enhance heat transfer. The PCM helps by storing excess heat during the day and releasing it later, allowing water evaporation to continue even when sunlight reduces. Throughout the experiment, important factors like water temperature, glass temperature, surrounding temperature, water output, and efficiency were measured and analyzed. The results clearly showed that the modified solar still produced more fresh water than the conventional one, making it a practical, low-cost, and sustainable solution for areas facing water shortages.**

Keywords— Solar still desalination, phase change material (PCM), finned absorber, renewable energy desalination, solar thermal distillation, water scarcity solution

I. INTRODUCTION (*HEADING 1*)

The shortage of freshwater is becoming a serious global concern due to factors such as rapid population growth, climate change, urban expansion, and increasing industrial demand. Many regions around the world are already facing limited access to clean and safe drinking water. India, for example, is home to nearly 18% of the world's population but has access to only about 4% of global freshwater resources. As a result, around 600 million people in the country experience high to severe levels of water stress. This situation highlights the urgent need to explore alternative ways of producing freshwater. Desalination, which involves removing salts and impurities from seawater or saline water, has gained attention as a viable solution. However, commonly used methods like reverse osmosis and multi-stage flash distillation require large amounts of energy and involve high operational costs These systems often rely on fossil fuels, contributing to environmental pollution and greenhouse gas emissions. In comparison, solar desalination provides a cleaner and more sustainable approach by using renewable solar energy. Solar stills work by imitating the natural water cycle, where water evaporates and then condenses to form purified water. They are simple in construction, cost-effective, and require very little maintenance, making them suitable for use in remote and water-scarce areas. This study focuses on the experimental investigation of an innovative solar still system incorporating thermal enhancement techniques to improve freshwater yield and system efficiency.

II. LITERATRE REVIEW

Many researchers have explored different ways to improve the performance and water output of solar stills. Common methods include adding fins, using wick materials, incorporating phase change materials (PCM), and improving absorber surfaces. For example, Samuel et al. (2022) studied solar still designs for remote island use and highlighted that better evaporation and condensation processes are key to increasing productivity. Mayilsamy et al. (2024) carried out experiments on finned solar stills and found that increasing the heat transfer surface area leads to higher evaporation rates and better water yield. Pacheco et al. (2025) focused on optimizing solar still designs and showed that factors like basin size, glass cover angle, and proper insulation greatly affect system efficiency. Kumar et al. (2021) worked on tubular solar stills with multiple wicks and reported improved performance due to better water distribution and a larger evaporation surface. Another effective method is the use of phase change materials for storing thermal energy. Kabeel et al. (2017) demonstrated that materials like paraffin wax can absorb heat during the day and release it later, allowing water production to continue even after sunset. Similarly, El-Said et al. (2023) studied the use of wire mesh absorbers and found that they enhance evaporation by increasing surface area and heat absorption. Overall, these studies show that combining different improvement techniques, such as PCM and surface modifications, can significantly enhance the efficiency and productivity of solar desalination systems.

III. SYSTEM DESIGN AND METHODOLOGY

A. Solar Still Design

The basin is coated with a material that enhances its ability to absorb solar radiation effectively. Saline water is then poured into the basin, where it is heated by the

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

absorbed solar energy, leading to evaporation. The water vapor formed rises and comes into contact with the inner surface of the glass cover, where it cools and condenses. As condensation takes place, the water droplets slide down along the inclined glass surface due to gravity and are collected in a trough in the form of distilled water. The basin dimensions used in the system are:

Length=0.60m

Width=0.45m

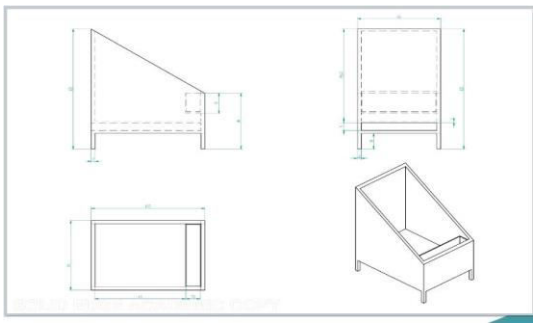


Fig.1. 2D Design of solar still

$$\text{Basin } area = 0.6 \times 0.45 = 0.27m \quad (1)$$

A shallow water depth of about 3 cm is maintained to reduce the thermal mass, allowing the water to heat up more quickly and thereby improving the rate of evaporation.

B.Glass cover inclination

- The glass cover is inclined to facilitate efficient condensation and collection of distilled water. The calculated inclination angle is approximately 28.4°, which lies within the recommended range for tropical regions (25°–30°). This angle ensures effective solar radiation absorption and smooth runoff of condensed water droplets.

Fig.2. Inclined Glass

B. Phase Change Material Integration

Paraffin wax is used as the phase change material in the system. PCM stores thermal energy during periods of high solar radiation and releases it during evening or cloudy conditions. This stored energy maintains higher water temperatures for longer durations, thereby increasing evaporation and freshwater production. The PCM tray volume used in the system is approximately 2.4 litres.

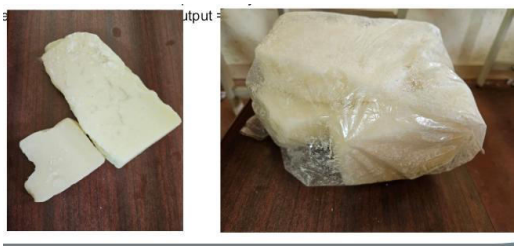


Fig.3. Paraffin Wax

D.FINS

In a solar still desalination system, fins are thin extended surfaces attached to the basin absorber plate to improve the heat transfer area. The main role of incorporating fins in the system is to enhance heat absorption and improve the transfer of heat from the absorber surface to the saline water. When sunlight passes through the glass cover, the absorber plate heats up. The fins attached to this plate capture additional heat and help transfer it more effectively to the water, causing the water temperature to rise at a faster rate. As the temperature of the water increases, the rate of evaporation also improves, resulting in a greater yield of distilled water. In addition, the fins help spread the heat more evenly throughout the basin, reducing the chances of uneven or localized heating and thereby improving the overall thermal efficiency of the solar still. Fins are generally made from materials that conduct heat well, such as aluminum or copper, although mild steel can also be used in low-cost designs. Adding fins to a solar still has been found to improve freshwater production by enhancing the evaporation process. Because of this, integrating fins into the design is seen as a practical and cost-effective way to boost the overall performance of solar desalination systems

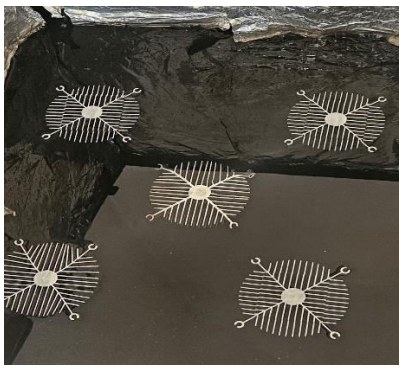


Fig.4. Fins

E. Expected Water Output

The expected freshwater production is estimated using solar radiation data. Average solar radiation in Kerala:

G = 5 kWh/m²/day

Solar energy received by the basin: Q = G × A Q = 5 ×0.27 = 1.35 kWh/day

Converting to Joules: Q = 4.86 × 10 J/day

Using the latent heat of vaporization of water: L = 2.26 × 10 J/kg

The theoretical distilled water output: m = Q / L 2.1 kg/day Considering a practical system efficiency of around 45%, the expected actual output is approximate 1.1 liters per day.

F.Heat Transfer Mechanism

The performance of the solar still depends on the heat transfer processes occurring within the system. The primary heat transfer mechanisms include:

- Solar radiation absorption by the basin liner
- Convective heat transfer between water and glass cover
- Evaporative heat transfer from water surface
- Radiative heat transfer between water surface and glass cover
- Heat loss through conduction and convection

The evaporative heat transfer plays a major role in freshwater production. Wind speed also influences heat loss from the glass surface and can be estimated using the convective heat transfer coefficient equation: $H_{(ca)} = 2.8 + 3.8V$ where V represents wind speed in m/s

IV. MODIFIED SOLAR STILL WOKING

Fig.5.Modified work

Fig.6.Modified Frame

Working of Solar Desalination Using Fins and PCM A solar desalination system that uses fins along with Phase Change Material (PCM) operates by harnessing solar energy to convert saline water into fresh water through evaporation and condensation. The main components of this system include a basin, an absorber plate fitted with fins, a layer of PCM, a transparent glass cover, and a collection channel. During the daytime, sunlight passes through the glass cover and reaches the absorber plate inside the basin. The absorber plate converts this solar energy into heat, which is then transferred to the saline water. The fins attached to the absorber plate help increase the surface area for heat transfer, allowing heat to spread more efficiently throughout the water. This leads to a quicker rise in water temperature and, in turn, enhances the evaporation process. At the same time, the PCM—commonly paraffin wax—absorbs excess heat during periods of intense sunlight by melting and storing thermal energy. Later, when the solar intensity drops in the evening or during cloudy conditions, the PCM slowly releases this stored heat as it solidifies. This helps maintain the water temperature for a longer duration, allowing evaporation to continue even after peak sunlight hours. As the water heats up, it begins evaporate, and the vapor rises toward the inner surface of the glass cover, which remains relatively cooler due to the surrounding air. The vapor then condenses into tiny droplets of freshwater on the glass surface. These droplets flow downward along the inclined glass cover and are collected in a trough. By combining the effects of fins and PCM, the system achieves better heat transfer, higher evaporation rates, and extended operating time. As a result, the modified solar still produces more freshwater compared to a conventional design.

V. RESULT AND DISCUSSION

.

A. TEMPERATURE OF SOLAR STILL IN FOUR CASES

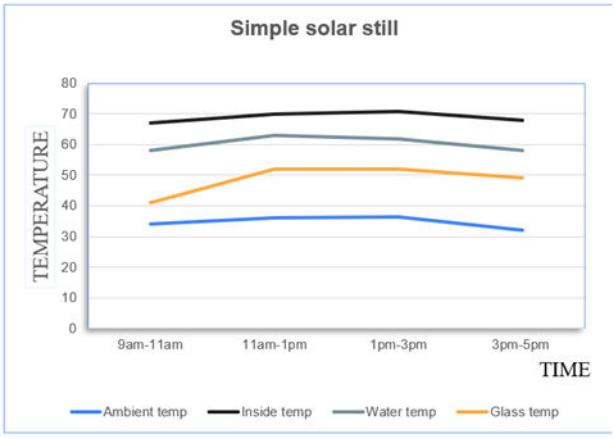


Fig. 6. Simple still temperature of solar still with 4 parameters

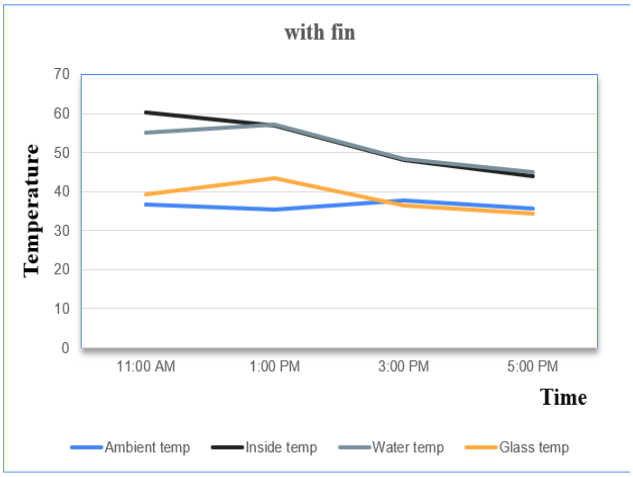


Fig.7 . with fin temperature of solar still with 4 parameters

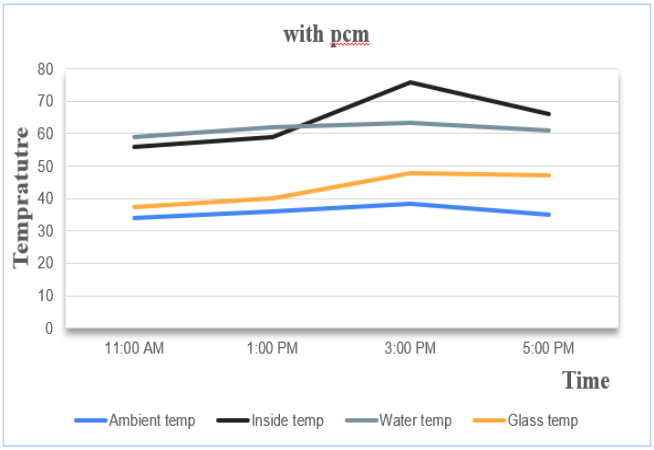


Fig.8 with pcm temperature of solar still with 4 parameters

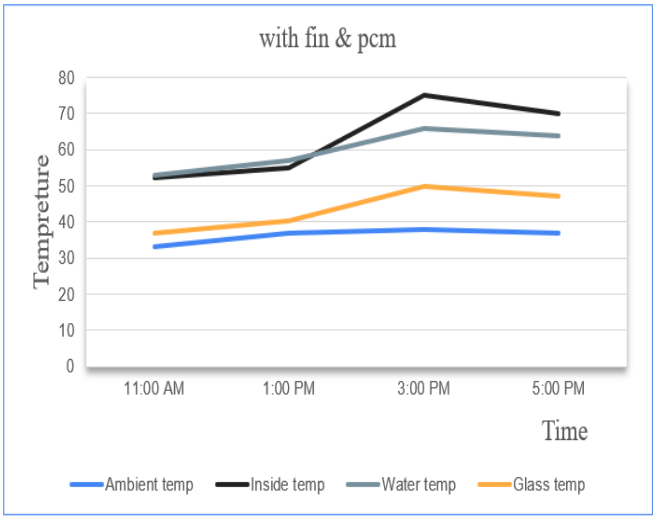


Fig .9. combination of still having fin & PCM

The Solar Still with Fin and PCM is clearly the most effective design, outperforming all other configurations in both total yield and efficiency. The synergistic combination of fins (which increase evaporative surface area) and PCM (which stores latent heat for extended operation) produces a significantly superior result. For practical solar distillation applications, integrating both fins and PCM is strongly recommended over using either enhancement individually or neither

B. YIELD OF WATER COLLECTED & CUMMULATIVE YIELD

Daily yield of water for the 4 cases as follows

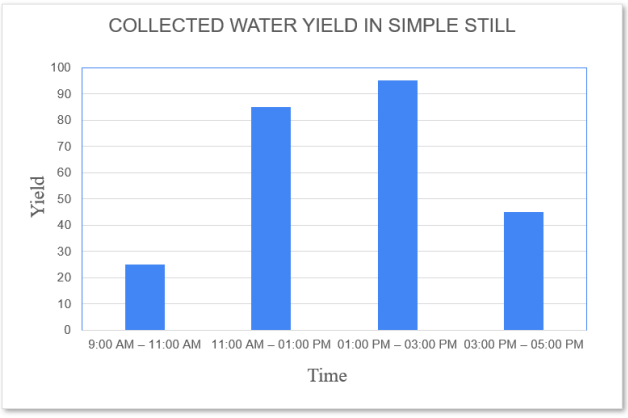
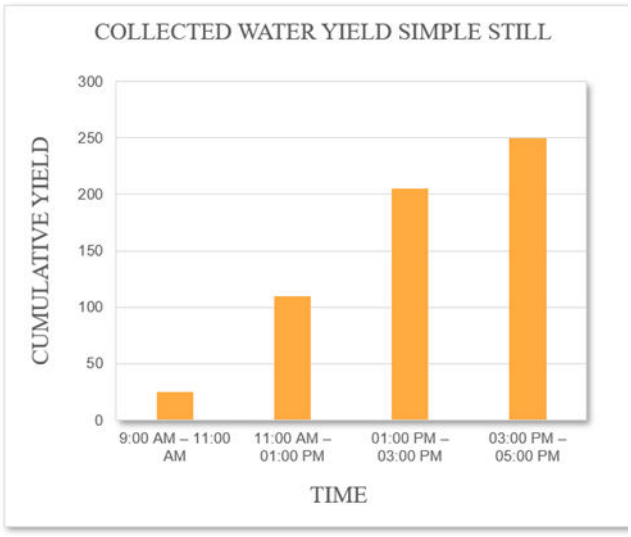


Fig.10. a) yield of water with conventional still



.Fig.10. b) Cumulative yield of water with simple still

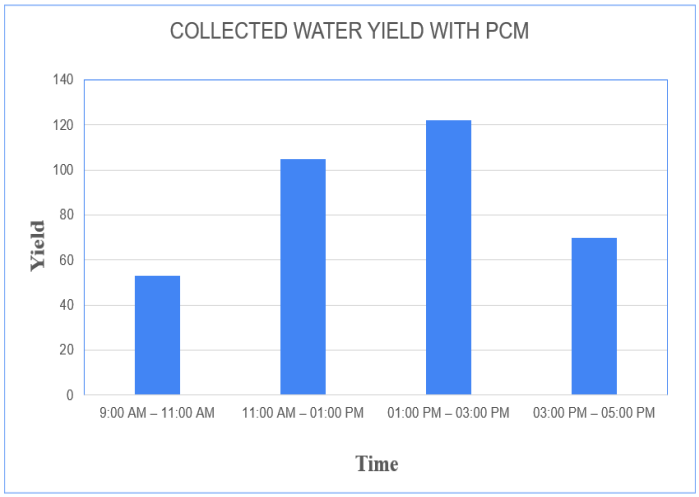


Fig.13.a) yield of water with FIN & PCM

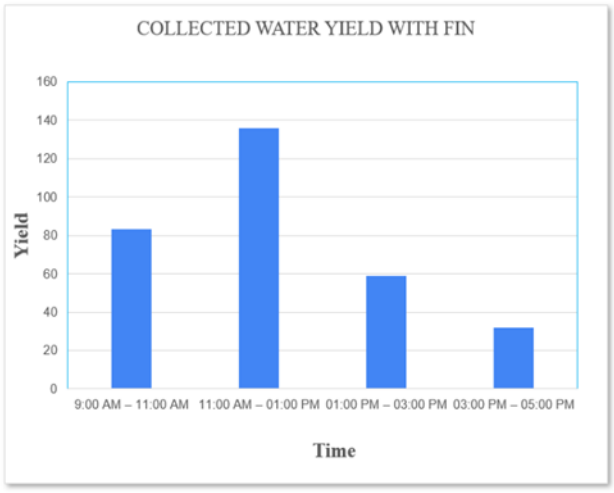


Fig.11.a) yield of water with fin

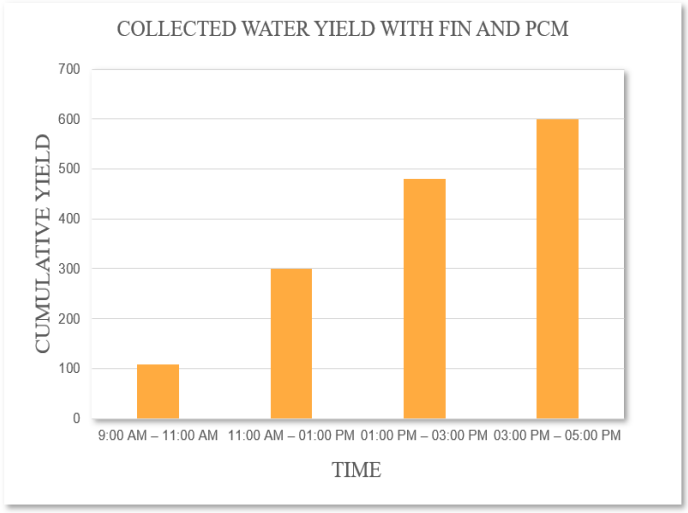


Fig.13.b) cumulative yield of water with FIN & PCM

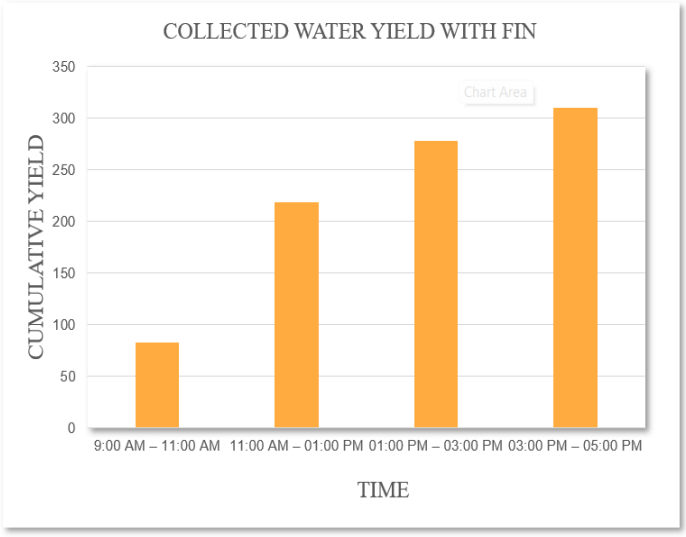


Fig.11.b) cumulative yield of water with fin

.

C. Energy Efficiency

a) Simple solar still :

.

TIME	Collected Water Yield	Efficiency
9 am - 11 am	25 ml	2.91
11 am - 1pm	85 ml	9.89
1 pm - 3 pm	95 ml	11.06
3 pm - 4 pm	45 ml.	5.24
Total 8 hr	250 ml	Overall efficiency 7.27 %

b) solar still with fins :

TIME	Collected Water Yield	Efficiency
9 am -11 am	83 ml	9.66
11 am - 1 pm	136 ml	15.84
1 pm - 3 pm	59 ml	6.87
3 pm - 5 pm	32 ml	3.73
Total 8 hr	310 ml	Overall efficiency9.01 %

b) solar still with PCM :

Time	Efficiency	Efficiency
9 am – 11am	6.12	6.12
11 am - 1 pm	12.23	12.23
1 pm - 3 pm	14.20	14.20
3 pm - 5 pm	8.15	8.15
Total 8 hr	Overall efficiency 10.18 %	Overall efficiency 10.18 %

d) solar still with fins and PCM :

Time	Collected Water Yield	Efficiency
9 am - 11 am	108 ml	12.57
11 am - 1 pm	192 ml	22.34
1 pm - 3 pm	180 ml	20.95
3 pm - 5 pm	120 ml	13.97
Total 8 hr	600 ml	Overall efficiency 17.46%

VI. Application

The system can be used for freshwater generation in coastal regions, rural communities, desert areas, and emergency water purification systems. Due to its low cost and minimal maintenance requirements, it is suitable for decentralized water supply systems.

VII . Conclusion

The present study successfully investigated the performance of both conventional and modified solar still desalination systems for sustainable freshwater production. The experimental study of the conventional solar still helped establish a baseline for its performance under local climatic conditions. However, its relatively low efficiency made it

clear that improvements were necessary. The modified solar still, which incorporates phase change material (paraffin wax) along with fins, showed a clear improvement in both thermal performance and freshwater output. The PCM played an important role by storing excess heat during periods of strong sunlight and releasing it later when solar intensity decreased, thereby allowing the evaporation process to continue for a longer time. In addition, the use of fins increased the effective surface area for heat transfer, which helped in heating the water more efficiently and boosting the evaporation rate. As a result, the modified system achieved higher productivity and better overall efficiency compared to the conventional solar still, while still remaining simple in design and cost-effective. These advantages make the system particularly suitable for remote and coastal areas, as well as regions facing water scarcity, where access to electricity and advanced desalination technologies is limited. The study ultimately shows that combining thermal energy storage with improved heat transfer techniques can significantly enhance the performance of solar stills, making them a practical and sustainable option for decentralized water purification

VIII. REFERENCES

[1] A. Samuel, J. Brizuela, K. C. Chang, and C. T. Lin, "Design and investigation of an effective solar still applicable to remote islands

[2] Y. Mayilsamy, S. Natarajan, and D. Balakrishnan, "Experimental investigation on the single slope desalination with finned pond.

[3] A. Pacheco, C. Sequeira, M. Moreira da Silva, and V. Serrao Sousa, "Enhancing efficiency in solar non-intrusive desalination," Desalination, vol. 598, 2025.

[4] A. E. Kabeel, M. A. Teamah, M. Abdelgaied, and G. B. Abdel Aziz,"Modified pyramid solar still with V-corrugated absorber plate and PCM as a thermal storage medium," Journal of Cleaner Production, vol. 161,pp. 881–887, 2017

[5] P. S. Kumar, V. Sathishkumar, and S. Manikandan, "Experimental investigation on tubular solar still with multiple wicks," Journal of Physics:Conference Series, vol. 2071, no. 1, 012007, 2021

Solar-Based Continuous Hot Water Supply System Enhanced with Phase Change Materials*

1st Abhijith M M
Department of Mechanical Engineering
Vidya Academy of Science and Technology
Thrissur,Kerala,India
abhijithmm2004@gmail.com

2ndHrithikkrishna V D
Department of Mechanical Engineering
Vidya Academy of Science and Technology)
Thrissur,Kerala,India
hrithikkrishnavd@gmail.comD

3rdFevin Francis
Department of Mechanical Engineering
Vidya Academy of Science and Technology
Thrissur,Kerala,India
francisfevin6@gmail.com

4th Justin Thomas
Department of Mechanical Engineering
Vidya Academy of Science and Technology)
Thrissur,Kerala,India
justinthomas0003@gmail.com

5th Sarath Babu Ramachandran
Department of Mechanical Engineering
Vidya Academy of Science and Technology
Thrissur,Kerala,India
email address or ORCID

6th Dr.Renjish Vijay
Department of Mechanical Engineering
Vidya Academy of Science and Technology
Thrissur,Kerala,India
email address or ORCID

***Abstract*—This study investigates the design and performance of a solar water heating system integrated with Phase Change Material (PCM) for efficient thermal energy storage. The system incorporates evacuated tube collectors to heat water, which subsequently transfers heat to the PCM via an immersion coil heat exchanger. During periods of solar availability, the PCM stores thermal energy by undergoing a phase change, while during non-sunshine periods, it releases the stored heat to sustain water temperature. Experimental analysis reveals that 11 kg of PCM melts in approximately 28 minutes at an inlet water temperature of 75 °C and a flow rate of 1 L/min. Furthermore, during the absence of solar radiation, the system increases water temperature from 32 °C to 50 °C, demonstrating effective thermal energy recovery. The results confirm that PCM integration significantly improves system efficiency and ensures a reliable hot water supply under intermittent solar conditions.**

***Index Terms*——Phase change material , Solar water heater, Paraffin wax, stearic acid, Evacuated tube**

I. Introduction

Increasing energy demands and decreasing conventional energy resources have led to the increasing importance of renewable energy systems. Solar energy is one of the most abundant and environmentally friendly sources of energy and is widely used in water heating applications because of its simplicity and cost-effectiveness. Solar water heating systems are commonly used in domestic and industrial applications; however, the system is restricted by the availability of sunshine. One of the major drawbacks of conventional solar water heating systems is the lack of hot water supply during non-sunshine periods such as nighttime and cloudy conditions. This is the major limitation of the system, which affects the reliability of the system. To overcome this problem, thermal energy storage systems are introduced to store heat during sunshine periods and utilize it when solar energy is not available. One of the commonly used thermal energy storage materials is Phase Change Materials (PCMs), which is widely used in thermal energy storage systems because of their large latent heat storage capacity and constant temperature during the change of their states. The PCM can be used in solar water heating systems to absorb heat during sunshine periods by melting and release excess heat during non-sunshine periods by solidifying, which will maintain the water temperature for a longer time. In this work, the solar water heating system is designed and analyzed using PCM as the thermal energy storage medium to evaluate the effectiveness of PCM in storing excess heat during sunshine periods and utilizing it during non-sunshine periods.

II. LITERATURE SURVEY

The integration of Phase Change Material (PCM) with solar thermal collectors for the purpose of domestic hot water heating was studied by Pathak S. K. et al. (2023). The researchers were able to prove that the integration of the PCM with the U-tube solar collector enhances the storage of thermal energy and ensures the availability of hot water during low solar radiation. The study proves that the integration of the PCM with the solar collector significantly enhances the reliability and efficiency of the solar water heating system.

The efficiency of the finned and un-finned U-tube solar water heaters with the integration of paraffin wax PCM was analyzed by Sayed S. M. S. E. et al. (2017). The experimental analysis proved that the integration of the PCM with the solar water heating system maintains the water temperature above 30°C during the night hours. The study proved that the integration of the finned tube with the solar collector enhances the heat transfer rate in comparison with the un-finned tube. The significance of the heat transfer enhancement techniques was proved in the study.

A comparative study of various designs of evacuated tube solar collectors was carried out by Said S. et al. (2023). The study aimed to analyze the effect of various designs on heat absorption. The results showed that the optimization of collector design enhances efficiency. This shows that proper collector design selection is important for the efficient utilization of solar energy for water heating.

The characteristics of paraffin wax in a cylindrical storage device for thermal energy storage were studied by Huang X. et al. (2019). The study aimed to analyze the melting and solidification of PCM for various temperature conditions. The results showed that a greater temperature difference enhances the heat transfer rate, resulting in improved efficiency. This shows that the temperature gradient also affects the performance of a thermal storage device.

The effect of heat exchanger configuration on the melting and solidification of PCM was studied by Radomska E. et al. (2020). The results showed that the geometry of the heat exchanger affects the performance of the heat transfer system. The results showed that proper heat exchanger configuration enhances the efficiency of the thermal storage device.

III. SYSTEM DESCRIPTION

The developed system is a solar water heating system with the integration of Phase Change Material for storing thermal energy. The system uses an evacuated tube collector to collect solar irradiance and heat the water. The water is then circulated through an immersion coil heat exchanger that is placed in the PCM storage tank. This helps in the effective heat transfer between the circulated water and the PCM.

A. Components of the System

The evacuated tube collector is used as the primary heat absorption unit. The solar energy is converted into heat energy. The storage tank is used to hold the water for some time before it is passed into the heat exchanger. The immersion coil heat exchanger is used. Copper is used because it has high thermal conductivity. The heat is transferred efficiently. The PCM used in the system is a mixture of paraffin wax and stearic acid. The mixture is chosen because it has the appropriate melting point and high latent heat. The container holds the PCM. The piping system is used to connect all the devices and ensure that the water flows continuously.

B. Working Principle

During the sunshine period, water is heated in the evacuated tube collector up to a higher temperature and then passes through the immersion coil heat exchanger. During this period, Valve 1 is kept open, and Valve 2 is kept closed, so hot water passes through the heat exchanger. Due to the temperature difference between the hot water and PCM, heat transfer takes place, which causes the PCM to melt.

During the non-sunshine period, the water flow path changes due to the opening of the valves. Valve 1 is kept closed, and Valve 2 is kept open, so water with a lower temperature passes through the heat exchanger. During this period, the PCM, which stores thermal energy, starts to solidify, transferring heat to the water passed through the heat exchanger, which causes an increase in the water temperature.

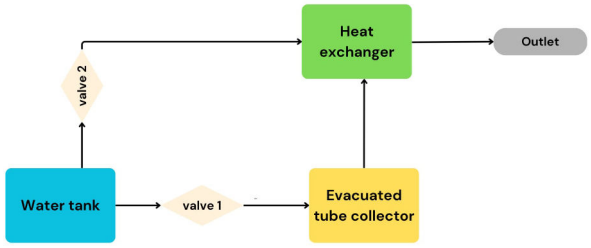


Fig. 1. Flow chart of the system.

IV. METHODOLOGY

A. System parameters

The performance of the PCM-based solar water heating system is evaluated using the following parameters. The amount of PCM used in the system is 11 kg. The latent heat of the system is 207 kJ/kg. The melting point of the system is around 55°C. Water is used as the heat transfer fluid in the system. The specific heat capacity of water is 4.18 kJ/kg°C. The inlet temperature of water during the sunshine period is around 75°C. During the non-sunshine period, the temperature is around 32°C. The flow rate of water is maintained at 1 litre per minute, which is equivalent to 1 kg/min.

B. Experimental Procedure

The system operates in two distinct conditions. The conditions include sunshine and non-sunshine periods. During the sunshine period, water is heated in the evacuated tube collector. The hot water is then circulated in the immersion coil heat exchanger. During this time, Valve 1 is left open, while Valve 2 is closed. This is done to allow hot water to flow in the direction of the PCM container. The temperature of the water at both the inlet and outlet of the heat exchanger is measured. During the non-sunshine period, the system operates with the opposite conditions. Valve 1 is closed, while Valve 2 is left open. This enables cold water to flow in the direction of the heat exchanger. The thermal energy is released by the PCM as it solidifies. The temperature of the water at the outlet is measured.

Fig. 2. Fabricated model of the System.

C. Thermal Analysis

For the purpose of simplifying the analysis, the following assumptions are made. The heat loss to the surroundings is neglected. The flow rate of water is taken to be constant during the entire process. The temperature of the PCM is taken to be constant during the phase change process. The thermal properties of water and the PCM are taken to be constant over the range of the temperature. The heat required to melt the PCM is calculated by: The heat required to raise the temperature of water is calculated by:

$$Q = m \times c \times \Delta T$$

where,

Q = heat required (kJ)

m = mass of water (kg)

c = specific heat capacity of water (kJ/kg·K)

$\Delta T = Temperaturedifference(C)$

Substituting the given values:

Q = mc$\Delta T = 6 \times 4.186 \times 40 = 1004.64$ kJ ≈ 1.0 MJ

Thus, about 1004.64 kJ of heat energy is required to heat the water. The PCM storage capacity is calculated as:

$$Q = m \times L$$

where,

m = mass of PCM (kg)

L = latent heat of fusion (kJ/kg)

Mass of PCM = 11 kg

$$Q = 11 \times 207 = 2277kJ$$

Therefore, the PCM can store approximately 2277 kJ of thermal energy, which is sufficient to provide heat during non-sunshine periods. The exchange of heat between the circulating water and Phase Change Material (PCM) is also dependent on the temperature variation between the inlet water and the PCM, and the rate at which mass of water inside the immersion coil heat exchanger passes. The following The heat transfer rate is estimated by calculations, and the time needed to get a complete one melting of the PCM.

- Mass of PCM = 11 kg
- Latent heat of PCM = 207 kJ/kg
- Inlet water temperature = 75 °C
- PCM melting temperature = 55 °C

Temperature difference:

$$\Delta T = T_{\text{in}} - T_{\text{pcm}} = 75\,^{\circ}\text{C} - 55\,^{\circ}\text{C} = 20\,^{\circ}\text{C}$$

Specific heat capacity of water:Cp = 4.18 kJ/kg·°C

Water flow rate = 1 litre/min

Since 1 litre of water 1 kg The mass flow rate becomes: = 1 kg/min . The selected flow rate ensures continuous circulation of water through the copper coil heat exchanger and allows effective heat transfer between water and PCM. The total heat required to melt the PCM is calculated using:

$$Q = m \times L$$

$$Q = 11 \times 207 = 2277kJ$$

Therefore, 2277 kJ of thermal energy is required to completely melt the PCM. The heat supplied by the circulating water per minute is calculated as:

$$Q = \dot{m}\, C_p\, \Delta T$$

$$Q = 1 \times 4.18 \times 20 = 83.6 \text{ kJ/min}$$

Thus the circulating water supplies approximately 83.6 kJ of heat per minute to the PCM. The time required to melt the PCM is calculated by dividing the required heat by the heat supplied per minute.

$$t = \frac{Q_{\text{required}}}{Q_{\text{per minute}}} = \frac{2277}{83.6} = 27.23 \text{ minutes}$$

Therefore, the time required to melt the PCM completely is approximately: t = 28 minutes.

V. RESULTS AND DISCUSSION

The performance of the PCM integrated solar water heating system was evaluated under both sunshine (charging) and non-sunshine (discharging) conditions. The results focus on the melting behavior of PCM, heat storage capability, and heat recovery performance.

A. Experimental Observations

TABLE I
OPERATING TEMPERATURE CONDITIONS

Condition	Inlet Temp (°C)	Outlet Temp (°C)
Sunshine period	75	55
Non-sunshine period	32	50

During the sunshine period, solar-heated water transfers excess heat to the PCM through a coil. This process melts the PCM and stores energy as latent heat, reducing the outlet temperature to about 55°C. During the non-sunshine period, the PCM releases the stored heat to the cooler incoming water, raising its temperature to around 50°C. This system provides a continuous supply of hot water even when there is no sunlight. When the solar collector cannot heat water, cooler water at 32°C enters the system. The melted PCM releases its stored

latent heat through the coil, raising the outlet temperature to about 50°C.

This allows for a hot water supply even without sunlight. The experimental results show that the PCM effectively stores excess heat during the sunshine period and keeps the water temperature at 55°C, confirming it absorbs energy well. During non sunshine periods, the stored heat is released, increasing the water temperature from 32°C to about 50°C. The PCM system thus improves solar heater efficiency by ensuring a continuous supply of hot water day and night.

VI. APPLICATIONS

The developed PCM-integrated solar water heating system has several important applications in different sectors. The ability of the system to store thermal energy and supply hot water continuously makes it suitable for domestic, industrial, and commercial uses. The major application areas of the system are listed below.

A. Domestic Continuous Hot Water Supply

Provides uninterrupted hot water for homes even during night or cloudy conditions. Ensures energy savings by reducing dependence on electric geysers.

B. Suitable for hotels, hostels, and resorts

where large amounts of hot water are required.PCM storage ensures hot water availability 24/7, improving customer comfort.

C. Hospitals and Healthcare Facilities

Used for sterilization, cleaning, and patient care requiring hot water continuously. Reliable backup during power failures due to stored thermal energy.

D. Industrial Water Heating Applications

Applicable in industries like textile, food processing, and dairy for process heating. Reduces fuel consumption and operational cost.

VII. CONCLUSIONS

This paper demonstrates that integrating Phase Change Material (PCM) with a solar water heating system effectively overcomes the limitation of conventional systems that cannot supply hot water during non-sunshine hours. The use of a paraffin wax and stearic acid mixture provides efficient latent heat storage and controlled thermal energy release, enabling stable temperature maintenance. The designed system successfully stores excess heat during sunlight hours and delivers hot water even in the absence of solar radiation, ensuring continuous and reliable operation. Experimental results validate consistent inlet–outlet temperature variation and confirm effective heat transfer through the copper immersion coil heat exchanger. The observed performance during both sunshine and non-sunshine periods clearly proves the feasibility of PCM-based thermal energy storage in improving system reliability and continuity. Furthermore, the system demonstrates strong potential as a sustainable and energy-efficient solution for domestic hot water applications. The use of simple materials, compact design, and renewable solar energy makes the system cost-effective and environmentally friendly. By reducing dependence on conventional electrical heating methods, the proposed system contributes to energy conservation and lower operational costs. In addition, its scalability and adaptability make it suitable for wider applications in residential, institutional, and small scale industrial sectors. Overall, the integration of PCM significantly enhances the performance of solar water heating systems and represents a promising approach for future energy-efficient thermal storage technologies.

.

ACKNOWLEDGMENT

The authors express their sincere gratitude to Mr. Sarath Babu Ramachandran (Project Guide) and Dr. Renjish Vijay (Co-Guide) for their valuable guidance and support. The authors also thank the faculty and management of Vidya Academy of Science and Technology for their support.

REFERENCES

[1] Pathak, S. K., Tyagi, V. V., 2023. Hot water generation for domestic use in residential buildings via PCM integrated U-tube based solar thermal collector.

[2] Sayed, S. M. S. E., Said, K. G. A. A. S., 2017. Experimental analysis of finned and un-finned U-tube solar water heaters with paraffin wax PCM.

[3] Said, S., Ghachem, K., 2023. An experimental comparison of the performance of various evacuated tube solar collector designs.

[4] Huang, X., Du, Z., 2019. Phase change characteristics of paraffin in cylindrical storage unit.

[5] Radomska, E., Mika, L., 2020. The impact of heat exchangers constructions on the melting and solidification time of phase change materials.

EXPERIMENTAL INVESTIGATION OF PLA BASED NATURAL COMPOSITES IN 3D PRINTING

ADARSH S
MECHANICAL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
adarshsreedharan2004@gmail.com

SHIYAS S
MECHANICAL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
shiyasshiyas548@gmail.com

VISHNU M
MECHANICAL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
vishnumanikandanv2@gmail.com

JISHNOOP JAYAPRAKASH
MECHANICAL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
Jishnoop.jayaprakash@ahalia.ac.in

VENKITESH K S
MECHANICAL ENGINEERING
Ahalia School of Engineering & Technology
Palakkad, India
venkitesh.ks@ahalia.ac.in

Abstract*— Additive manufacturing has become one of the fastest-growing production technologies in recent years. However, the extensive use of synthetic polymer-based filaments creates environmental concerns related to waste generation and dependence on non-renewable resources. This study focuses on the development of a sustainable composite filament by incorporating natural reinforcements such as wood flour and sisal fiber into Polylactic Acid (PLA) for fused deposition modeling (FDM) applications. Different material compositions were prepared and processed through extrusion to obtain 1.75 mm diameter filaments suitable for 3D printing. Mechanical testing such as tensile, flexural, compressive, and impact analysis were considered to evaluate the performance of the developed composite. Experimental results indicate that the addition of sisal fiber improves tensile strength and stiffness, while wood flour enhances surface finish and reduces material cost. The developed bio-composite filament demonstrates potential as an eco-friendly alternative to conventional plastic filaments used in 3D printing.

Keywords— 3D Printing, PLA, Natural Fiber Composite, Wood Flour, Sisal Fiber, Sustainable Manufacturing, FDM.

I. INTRODUCTION

Additive manufacturing, commonly known as three-dimensional (3D) printing, has emerged as one of the most innovative manufacturing technologies of the modern era. Unlike conventional subtractive manufacturing methods, where material is removed from a larger block through machining processes, 3D printing creates components layer by layer directly from digital models. This technology offers significant advantages such as reduced material wastage, rapid prototyping, design flexibility, shorter production cycles, and cost-effective manufacturing for customized products. Due to these benefits, 3D printing is now widely applied in industries such as aerospace, automotive, healthcare, construction, education, and consumer goods manufacturing.

Among the different 3D printing processes, Fused Deposition Modeling (FDM) is one of the most commonly used techniques because of its simplicity, affordability, and ease of operation. In FDM, a thermoplastic filament is melted and extruded through a heated nozzle to deposit material layer by layer according to a computer-aided design (CAD) model. Various thermoplastic materials are used in FDM printing, including Acrylonitrile Butadiene Styrene (ABS), Polyethylene Terephthalate Glycol (PETG), Nylon, Thermoplastic Polyurethane (TPU), and Polylactic Acid (PLA). Among these, PLA is highly preferred due to its biodegradable nature, low printing temperature, minimal warping, and user-friendly processing characteristics. PLA is derived from renewable agricultural resources such as corn starch and sugarcane, making it more sustainable than petroleum-based plastics.

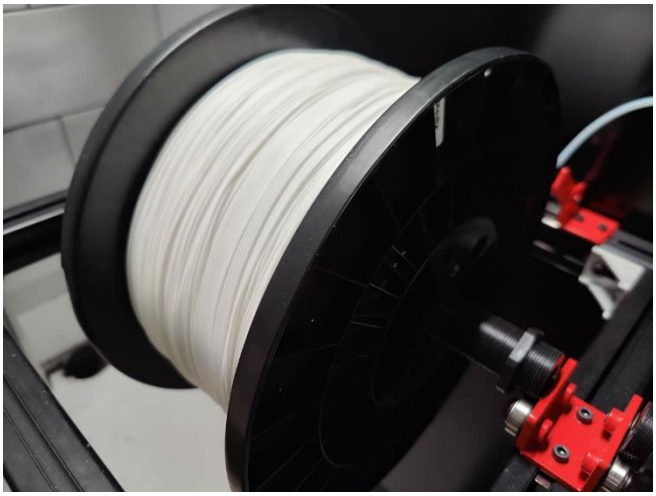
Although PLA possesses many advantages, it also has certain limitations such as lower impact resistance, moderate tensile strength, brittleness, and reduced thermal stability compared to engineering plastics. These limitations restrict its use in structural and load-bearing applications. In addition, the increasing use of plastic-based filaments in 3D printing contributes to waste generation, environmental pollution, and dependence on synthetic materials. Therefore, researchers are focusing on the development of eco-friendly composite materials that can enhance the properties of PLA while maintaining sustainability.

Natural fibers and bio-fillers have gained considerable attention as reinforcement materials for polymer composites. Materials such as sisal, jute, flax, hemp, bamboo, wood flour, and coir are biodegradable, lightweight, inexpensive, and abundantly available. These materials possess good specific strength and stiffness, making them suitable for improving the mechanical properties of thermoplastic matrices. Sisal fiber, in particular, is known for its high tensile strength, rough surface texture, and excellent bonding potential with polymers after suitable surface treatment. Similarly, wood flour acts as a low-density filler that can improve surface finish, reduce material cost, and provide a natural appearance to printed products.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

The integration of natural fibers with PLA in filament production can create a new class of sustainable composite materials suitable for additive manufacturing. However, challenges such as moisture absorption, poor interfacial bonding, nozzle clogging, and inconsistent extrusion behavior must be addressed through proper material preparation and processing conditions.

The present study focuses on the experimental investigation of PLA-based natural composites reinforced with sisal fiber and wood flour for FDM 3D printing applications. The objective is to develop a composite filament with improved mechanical performance, acceptable printability, and reduced environmental impact. The prepared filament is evaluated through mechanical testing such as tensile, flexural, compression, and impact analysis. This work contributes to the growing field of green manufacturing by promoting renewable materials in additive manufacturing technologies.



II. OBJECTIVES

The main objective of this study is to develop and experimentally investigate a sustainable PLA-based composite filament reinforced with natural materials such as sisal fiber and wood flour for 3D printing applications, while improving mechanical properties, printability, and environmental friendliness.

The specific objectives of the study are:

- To formulate a composite filament using PLA, sisal fiber, and wood flour.
- To analyze the suitability of the filament for FDM 3D printing.
- To study the mechanical and physical properties of the developed composite.
- To promote eco-friendly materials in additive manufacturing.

III. MATERIALS USED

A. *PLA (Polylactic Acid)*

PLA is a biodegradable thermoplastic widely used in 3D printing due to its easy processing, low melting temperature, and eco-friendly nature. It acts as the base matrix material in the composite.

B. *Sisal Fiber*

Sisal fiber is a natural fiber with high tensile strength, low density, and good stiffness. It is used as reinforcement to improve the strength and rigidity of the composite filament.

C. *Wood Flour*

Wood flour is a fine powder obtained from wood particles. It is used as a filler material to enhance surface finish, reduce cost, and provide a natural wood-like appearance.

IV. METHODOLOGY

The methodology adopted in this study involves the preparation, fabrication, and testing of a biodegradable composite filament for 3Dprinting.

A. *Material Preparation*

PLA, sisal fiber, and wood flour were selected as raw materials. Sisal fiber and wood flour were cleaned and dried to remove moisture. Dry at 60–80°C for 4–6 hours. The sisal fibers were cut into small lengths (0.5–2mm), and wood flour was sieved (<200 µm) to obtain fine particles.

B. *Fiber Treatment*

The sisal fiber was treated with 5% NaOH solution or 1–2% silane, to improve bonding between the fiber and PLA matrix.

C. *Mixing of Materials*

The prepared materials were mixed in required proportions of PLA, sisal fiber, and wood flour.

Materials	1st Combination (%)	2nd Combination (%)	3rd Combination (%)
PLA	90	80	80
Wood Flour	5	10	15
Sisal	5	10	5

D. *Extrusion Process*

The mixture was processed using a twin-screw extruder to produce composite pellets, which were then converted into filament of 1.75 mm diameter.

E. 3D Printing

The developed filament was used in an FDM 3D printer to print standard test specimens.

F. Testing and Evaluation

Printed specimens were tested for tensile, flexural, compressive, and impact properties according to ASTM standards.

After curing, the prepared boards were subjected to various tests to evaluate their physical and mechanical properties. The tests conducted include density measurement, water absorption test, compressive strength test, and flexural strength test. The results obtained from these tests were recorded and compared for different mix proportions to determine the optimum composition for composite gypsum boards.

TEST	TENSILE STRENGTH TEST(MPa)	FIRE RETARDATION TEST (SOFT ,MELTING TEMPERATURE) (DEGREE CELSIUS)	WATER ABSORPTION TEST (%)
FILAMENT1	7.143	51 ,79.6	4
FILAMENT2	4.19	47,66	5.03
FILAMENT3	5.12	51.33,74.6	6.8

V. RESULTS AND DISCUSSION

The experimental evaluation of the developed biodegradable composite filament was carried out through tensile, fire retardation, and water absorption tests. The tensile test results indicated that the fabricated filament exhibited an average tensile strength of 7.143 MPa, which is significantly lower than that of commercially available PLA filament (27.22 MPa). This reduction in strength can be attributed to the incorporation of natural fillers and possible weak interfacial bonding between the matrix and reinforcement. However, the obtained strength is still adequate for low-load and eco-friendly applications, particularly in non-structural components.

The fire retardation test demonstrated that the composite filament showed improved resistance to ignition compared to conventional PLA. The presence of filler materials contributed to reduced flame propagation and slower burning rates. This behavior indicates that the developed filament has potential applications in areas where moderate fire resistance is required. However, further enhancement through optimized filler composition could improve its performance to meet higher safety standards.

Water absorption analysis revealed that the composite filament absorbed more moisture than standard PLA due to the hydrophilic nature of the natural reinforcements. The percentage increase in weight after immersion confirmed higher water uptake, which may affect the dimensional stability and mechanical properties over time. Despite this drawback, the material's biodegradability is enhanced, making it suitable for sustainable applications.

Overall, the developed biodegradable composite filament presents a trade-off between mechanical strength and environmental benefits. While the tensile properties are lower than conventional PLA, the improved fire retardation and biodegradability make it a promising alternative for specific applications such as packaging, disposable products, and low-load functional prototypes. Further research focusing on surface treatment of fillers and improved processing techniques can enhance the overall performance of the material.

TEST RESULTS OF DIFFERENT PROPORTIONS

VI. Advantages and Applications

The PLA-based natural composite filament developed using sisal fiber and wood flour offers several advantages over conventional synthetic 3D printing materials. One of the major benefits is its environmental friendliness. Since PLA is biodegradable and derived from renewable resources, and both sisal fiber and wood flour are natural materials, the composite reduces dependence on petroleum-based plastics and supports sustainable manufacturing. This helps in minimizing plastic waste and lowering environmental pollution.

Another important advantage is improved mechanical performance. Sisal fiber increases tensile strength, stiffness, and load-carrying capacity of the filament, making the printed products stronger than pure PLA components. Wood flour acts as a lightweight filler that enhances surface finish and provides a natural texture. It also helps in reducing material cost by partially replacing polymer content. The composite filament can be processed using standard FDM printers without major modifications, making it practical and economical for regular use.

The material also offers good dimensional stability, easy printability, and aesthetic appeal due to its wood-like appearance. It can be customized for different applications by changing filler percentage and printing parameters.

This composite filament has wide applications in various fields. It can be used for rapid prototyping, educational models, household products, decorative items, packaging materials, and consumer goods. Due to its lightweight and eco-friendly properties, it is suitable for interior automotive components, furniture accessories, and customized products. It can also be used in low-load engineering parts, toys, and architectural models.

In research and academic institutions, the material can be used for studying sustainable polymers and bio-composites. Overall, the developed filament provides a balance of strength, sustainability, low cost, and versatility, making it a promising material for future additive manufacturing applications.

VII. Conclusion

The present study successfully demonstrated the development of a sustainable PLA-based natural composite filament for 3D printing applications by incorporating sisal fiber and wood flour as reinforcing materials. The growing demand for environmentally friendly manufacturing materials has created the need for alternatives to conventional petroleum-based plastics used in additive manufacturing. In this work, biodegradable PLA was selected as the matrix material, while natural fillers were introduced to improve mechanical properties and reduce environmental impact. The experimental investigation confirmed that natural reinforcements can be effectively combined with PLA to produce a functional composite filament suitable for fused deposition modeling (FDM) processes.

The preparation of the composite involved cleaning, drying, sizing, and chemical treatment of natural fibers to enhance compatibility with the polymer matrix. Proper mixing and extrusion techniques enabled the successful production of filament with acceptable dimensional consistency for 3D printing. The developed filament was printable under standard FDM conditions with suitable nozzle and bed temperature settings. This indicates that natural material-based composites can be adapted to existing commercial 3D printers without requiring major modifications to the printing system.

Mechanical performance evaluation showed that the inclusion of sisal fiber contributed positively to tensile strength, stiffness, and load-carrying capacity of the printed specimens. The high strength and rough surface texture of sisal fiber improved stress transfer between the matrix and reinforcement, resulting in better structural behavior compared to pure PLA. Wood flour acted as an effective filler material by improving surface texture, reducing material cost, and giving a natural finish to printed components. The combined effect of sisal fiber and wood flour created a balanced composite with improved strength and better appearance.

Another significant outcome of the study is the environmental advantage of the developed material. Both sisal fiber and wood flour are renewable, biodegradable, and abundantly available natural resources. Their incorporation into PLA reduces dependence on synthetic fillers and increases the bio-based content of the filament. This supports the concept of sustainable manufacturing and helps minimize the ecological footprint associated with plastic waste generation. The use of agricultural and wood by-products can also provide economic benefits by utilizing low-cost waste resources in value-added engineering applications.

However, certain challenges were also observed during experimentation. Moisture absorption by natural fibers can affect bonding quality and dimensional stability if materials are not properly dried. Higher filler content may lead to nozzle clogging, irregular extrusion flow, or reduced layer adhesion during printing. Therefore, careful control of fiber size, composition ratio, moisture content, and extrusion parameters is essential for consistent filament quality. Further optimization is necessary to achieve the best balance between printability and mechanical performance.

Overall, the study concludes that PLA reinforced with sisal fiber and wood flour is a promising alternative to conventional 3D printing filaments. The developed composite material combines acceptable mechanical strength, ease of printing, cost reduction, and environmental sustainability. It has potential applications in prototype manufacturing, decorative products, lightweight utility parts, educational models, and non-critical engineering components. Future work may focus on thermal characterization, long-term durability studies, biodegradation

behavior, and the use of other natural fibers to further enhance performance. This research contributes to the advancement of green materials in additive manufacturing and opens new opportunities for eco-friendly filament development.

VIII. FUTURE SCOPE

The developed PLA-based natural composite filament shows strong potential for further research and industrial application in sustainable additive manufacturing. Future studies can focus on optimizing the composition ratio of PLA, sisal fiber, and wood flour to achieve the best combination of strength, flexibility, and printability. By varying the percentage of natural fillers, improved mechanical and thermal performance can be obtained for different engineering applications.

Further work can also include advanced surface treatments and coupling agents to enhance bonding between PLA and natural fibers. Better interfacial adhesion can improve tensile strength, durability, and moisture resistance of the composite material. Thermal analysis such as TGA and DSC can be conducted to study heat resistance and processing stability of the filament.

Another important area is the development of moisture-resistant and fire-retardant bio-composite filaments for wider commercial use. Research may also be extended to other natural reinforcements such as coir, jute, banana fiber, hemp, or bamboo fibers. Hybrid combinations of multiple natural fibers can provide improved overall properties.

Large-scale production methods for commercial filament manufacturing can be explored to reduce cost and increase market availability. The material can also be tested for real-time applications in automotive interiors, packaging, consumer goods, and lightweight structural components. Overall, this work opens the path toward greener, biodegradable, and high-performance materials for next-generation 3D printing technologies.

Acknowledgment

The authors would like to express their sincere gratitude to the faculty members and staff of the Department of Mechanical Engineering for their valuable guidance and support throughout the completion of this research work. Their continuous encouragement and suggestions greatly helped in carrying out this study successfully.

The authors would also like to thank the laboratory staff for providing the necessary facilities and assistance during the preparation and testing of the 3D filaments. Their technical support was essential for conducting the experimental work effectively.

Finally, the authors extend their heartfelt appreciation to everyone who directly or indirectly contributed to the successful completion of this research study.

REFERENCES

[1] **Biopolymer-chitin products by direct Antanitta, V.S., Patadiya, J., Kandasubramanian, B., 2024. Biopolymer-chitin products by direct ink writing (DIW):A review . Hybrid Adv. 5, 100115.**

[2] **Kajbi˘ c, J., Fajdiga, G., Klemenc, J., 2023. Material extrusion 3D printing of biodegradable composites reinforced with continuous flax fibers. J. Mater. Res. Technol. 27, 3610–3620. https://doi.org/10.1016/j.jmrt.2023.10.148.**

[3] **Mazur, K.E., Borucka, A., Kaczor, P., Gadek, S., Bogucki, R., Mirzewi´nski, D., Kuciel, S., 2022. Mechanical, thermal and microstructural characteristic of 3D printed polylactide composites with natural fibers: wood, bamboo and cork. J. Polym. Environ. 30, 2341–2354.**

[4] **Melisa Memis¸, Dils¸ad Akgümüs¸ G. 22 January 2025 Development of Fe-reinforced PLA-based composite filament for 3D printing: Process parameters, mechanical and microstructural characterization. journal published by science direct.**

[5] **Chengcheng Cai · Pei Zhang · Yafei Wang · Yun Tan · Iek Man Lei · Ben Bin Xu · Ji Liu. 2024 .Sustainable three-dimensional printing of waste paper-based functional materials and constructs.**

[6] **Noshin Tasnim Tuli, Sinthea Khatun, Adib Bin Rashid. 2024. Unlocking the future of precision manufacturing: A comprehensive exploration of 3D printing with fiber-reinforced composites in aerospace, automotive, medical, and consumer industries.**

[7] **Raja Thandavamoorthy, Yuvarajan Devrajan . 2025.Comprehensive analysis of thermal properties and stability in palm fiber-reinforced PLA composites with bran filler for sustainable lightweight materials.**

Integration of Total Quality Management (TQM) in Mechanical Engineering Processes

Shinde Goba D
Sr. Lecturer, Mechanical Engineering
SNJB's Shri H.H. J.B. Polytechnic,
Chandwad, (MS) India
goba.shinde453@gmail.com

Sancheti Santosh D.
Professor, Mechanical Engineering:
SNJB's Late Sau.K.B. Jain College of Engineering, Chandwad ,(MS) India
santusancheti@gmail.com

Sancheti Dipak D.
Assistant Professor, Mechanical Engineering: *SNJB's Late Sau. K.B. Jain College of Engineering, Chandwad,(MS) India*
deeps.sancheti@gmail.com

***Abstract*—Total Quality Management (TQM) is a comprehensive concept of an organization-wide philosophy, which seeks to enhance every aspect of a company on a continuous basis. When it comes to mechanical engineering, TQM provides a systematized method of improving manufacturing systems, production planning, quality control, maintenance engineering, and lean operations. This paper explores the systematic approach to the implementation of TQM principles in the process of mechanical engineering, and the paper will discuss how such tools as Statistical Process Control (SPC), Failure Mode and Effects Analysis (FMEA), Six Sigma, and Kaizen could be integrated into the engineering process in order to get high quality products and high efficiency of the process. An integrated model of TQM integration is suggested and demonstrates the interrelation between leadership commitment, process standardization, customer focus, and continuous improvement. The analysis of traditional and TQM-based mechanical engineering approaches is provided in a structured table. An example of an Indian manufacturing case study representative of the automotive components industry has also been discussed in the paper where significant changes in the reduction of defects, cycle time, and the overall equipment effectiveness (OEE) were observed after the introduction of TQM. The major obstacles such as organizational resistance, lack of resources, and skills gaps are discussed. The results show that the implementation of TQM has a significant positive impact on the performance of the mechanical engineering process and competitiveness, especially regarding Industry 4.0 and smart manufacturing paradigms.**

Keywords—Total Quality Management, Mechanical Engineering, Statistical Process Control, Six Sigma, Lean Manufacturing, Kaizen, Continuous Improvement

I. INTRODUCTION

The world manufacturing environment has grown rather competitive, owing to the high rate of technological growth, strict regulatory measures, and changing customer demands. The systematic practices of quality management in this environment is no longer a choice but is a must to ensure a long-term success of the industry. The concept of Total Quality Management (TQM) has become one of the most integrated and successful strategies of organizational excellence in most of the fields such as mechanical engineering [1].

The processes involved in mechanical engineering, such as design, fabrication, assembly, testing, and maintenance are complex in nature and are prone to variability at various stages. Any defects introduced along the production line may be transferred along the line which results in rework, scrap, warranties, and huge financial losses. Conventional quality control systems, which are mainly based on quality control during the post production stage, have failed to help in identifying the cause of such variation. TQM, on the contrary, promotes a proactive and process-oriented paradigm that incorporates the issue of quality at all levels of an organization [2].

Introducing TQM to the process of mechanical engineering implies the systematic implementation of the following principles: customer focus, leadership commitment, employee involvement, process approach, evidence-based decision making, and continual improvement. These are a solid basis of shifting the engineering practice towards responsive rather than proactive quality management, formalized in the standard quality management systems like ISO 9001:2015 [3].

The TQM implementation in the Indian manufacturing setting has witnessed a massive growth especially in the automotive, aerospace, and heavy engineering industries. Efforts by the government like the make in India and the national manufacturing policy have also encouraged the adoption of quality in competitiveness among local manufacturers. Although these have been made, small and medium enterprises (SMEs) in India are still struggling to apply TQM systematically because of the limitation of resources and organizational inertia [4].

This paper will set out to provide and to discuss a detailed analysis of the integration of TQM in mechanical engineering operations. Specific objectives include:

i. to survey the existing literature on TQM principles and their application in the manufacturing sector;
ii. to present a conceptual TQM integration framework of the mechanical engineering field;
iii. to compare the traditional and TQM-based process approaches;
iv. to explain integration using an Indian industrial case study; and
v. to identify the current challenges and future research opportunities.

II. LITERATURE REVIEW

TQM theoretical backgrounds are linked to the early works of Deming, Juran, Crosby, and Feigenbaum in the middle of the twentieth century. The fourteen management transformation points of Deming and his focus on the statistical approaches formed the basis of process-based quality management [5]. The trilogy of quality that Juran proposed in planning, control, and improvement offered an organized guide to operationalizing quality in organizations.

Oakland [6] defined TQM as a total approach in enhancing competitiveness, efficiency and flexibility of an organization as a unit. The framework he developed focused on the

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

combination of processes, people, and systems with the help of leadership, planning, and customer-supplier chains. The later studies have always confirmed the positive effects of TQM on the performance of organizations in the manufacturing industries.

Within the framework of mechanical engineering, Besterfield et al. [7] have shown that control charts, Pareto analysis, cause-and-effect diagram, and process capability studies are the TQM tools that could be successfully used in machining, forming, and assembly processes to minimize dimensional variability. Likewise, Pyzdek and Keller [8] recorded the implementation of Six Sigma methodology which is a sophisticated TQM variant in the context of mechanical design and manufacturing settings with a level of defects of less than 3.4 per million opportunities.

Empirical studies carried out by Antony et al. [9] among manufacturing companies in India revealed that the application of TQM was effective in enhancing efficiency of the processes, customer satisfaction, and financial performance. Their research paper revealed the top management commitment and training of employees to be the most important success factors of TQM in Indian context. In line with this, Patel and Desai [10] studied the adoption of TQM in small-scale mechanical industries in Gujarat, and noted that quality circles and 5S methodology were the most popular tools.

Recent studies are more concerned with the intersection of TQM and digital manufacturing technologies. Salonitis and Tsinopoulos [11] discussed the idea of increasing TQM with Industry 4.0 technologies, such as IoT sensors, real-time analytics, and cyber-physical systems, which can be used to monitor processes continuously and conduct adaptive quality control. Tortorella et al. [12] examined lean-TQM synergies and found out that the organizations that integrate both styles perform better than those that follow either philosophy separately.

Even though there has been extensive theoretical and empirical advances, there is evident gap in literature regarding the structured models that are specifically designed to amalgamate TQM into all key dimensions of mechanical engineering processes specifically in the Indian manufacturing SMEs. The aim of this paper is to fill this gap by a holistic conceptual framework and a case study.

III. METHODOLOGY

The research approach that is applied in this paper is descriptive and analytical, which implies that the systematic literature review was supported by the conceptual framework development framework and analysis of a representative industrial case study. The methodology of the research includes three main stages:

Phase 1- Literature Review: A systematic literature search was performed using peer-reviewed journals, conference proceedings, and technical reports that were released between 2015 and 2025. Systematic search of databases like IEEE Xplore, Scopus, Web of Science, and Google Scholar was conducted using keywords like TQM mechanical engineering, quality management manufacturing, Six Sigma production, lean TQM integration and Indian manufacturing quality. A collection of 47 sources were first determined where 15 most relevant and recent ones were identified on the basis of relevance, citation impact and methodological rigor.

Phase 2 -Conceptual Framework Development: A conceptual framework in which TQM principles were integrated into the dimensions of mechanical engineering processes was developed based on the synthesized literature and the existing TQM models. The framework relates TQM elements such as customer focus, process approach, leadership, employee involvement and constant improvement to particular areas of mechanical engineering such as design engineering, manufacturing systems, quality control, production planning and maintenance engineering.

Phase 3 — Case Study Analysis: A typical case study of an Indian automotive components manufacturer which has implemented TQM was examined using published industrial data and secondary sources. The major performance indicators such as defect rate, cycle time, Overall Equipment Effectiveness (OEE) and customer complaint rate were measured prior to and following the implementation of TQM to determine the real effect of the integration.

Also, comparative analysis was conducted to make a comparison between conventional quality methods and TQM-oriented methods of quality in various dimensions like quality philosophy, handling defects, role of employees, use of data, and process orientation. The tabular form of this comparison is shown in Section IV.

IV. INTEGRATION OF TQM IN MECHANICAL ENGINEERING PROCESSES

A. Conceptual Framework of TQM Integration

- Fig. 1 shows the suggested Conceptual TQM Integration Framework in Mechanical Engineering Processes. The framework is in the form of a hierarchical model which has three levels:
- Foundation Layer: The bottom layer consists of foundation level Tier Top management commitment, organizational culture and strategic quality planning. TQM initiatives cannot be sustained without alignment of leadership.
- Process Level: The middle level is comprised of the fundamental engineering processes, including design and development, manufacturing processes, supply chain management, maintenance engineering, and production planning which are regulated by particular TQM tools and approaches.
- Outcome Tier: Continuous improvement, customer satisfaction, waste elimination and operational excellence are the outcomes of the TQM integration at the process level.

Fig. 1: Conceptual TQM Integration Framework for Mechanical Engineering — A three-tier hierarchical diagram

B. TQM in Manufacturing Systems

TQM integration into manufacturing systems starts with process standardization using documented work instructions, standard operating procedures and process capability analysis. Statistical Process Control (SPC) is implemented on key machining and forming processes to track dimensional features on a real-time basis. Control charts (X-bar and R charts, p-charts) allow operators to identify and take corrective actions on deviations in processes prior to the manufacturing of non-conforming parts [7].

Failure Mode and Effects Analysis (FMEA) is used in product and process design to identify potential failure modes in a systematically structured way, evaluate their seriousness, frequency of occurrence, detectability, and corrective measures, and rank corrective measures. In manufacturing, where mechanical designs are involved, FMEA helps to avoid the weaknesses in design that can lead to production defects or field failures [8].

C. TQM in Production Planning

TQM changes the production planning into a quality-oriented process as opposed to being schedule-oriented. Capacity planning is also combined with demand forecasting to make sure that production goals are realistic enough without affecting the quality of the products. Just-in-time (JIT) concepts, which form part of lean-TQM integration, minimise the work-in-progress inventories and diminish the feedback cycles that enable quicker identification and rectification of quality aberrations [12].

The Theory of Constraints (TOC) is a constraint management tool that recognizes bottleneck processes that restrict throughput and also exposes them to strong quality improvement focus. Quality holds and first-article inspection procedures are also included in production scheduling to ensure non-conforming batches do not pass through the value stream.

D. TQM in Quality Control

Under TQM, a prevention oriented quality assurance system replaces the traditional quality control of quality by way of inspection. Measurement System Analysis (MSA) guarantees accuracy of gauges and inspection devices by means of Gauge Repeatability and Reproducibility (GR&R) tests. Acceptance Sampling plans are drawn based on statistically sound procedures and Acceptable Quality Limits (AQLs) are gradually narrowed as the process capability increases [3].

Six Sigma technique, structured around the DMAIC (Define, Measure, Analyze, Improve, Control) framework, offers a data-driven methodology of rigor in the minimization of variability in processes and defect rates in mechanical manufacturing. Use of Design of Experiments (DOE) helps the engineers optimize the process parameters like cutting speed, feed rate, and depth of cut during machining operation [8].

E. TQM in Maintenance Engineering

Total Productive Maintenance (TPM) is a maintenance-related form of TQM that stresses on independent maintenance, preventive maintenance planning, and enhancing the reliability of equipment. TPM focuses on the six key equipment losses, which are breakdowns, setup and adjustment, idling, minor stoppages, reduced speed, and defects, and overall equipment effectiveness (OEE) as a composite performance measure [13].

The maintenance engineering framework incorporates condition monitoring technologies such as vibration analysis, thermography and oil tribology to allow predictive maintenance. Such a change in reactive to predictive maintenance minimizes the unexpected downtime, extends the lives of equipment and maintains the quality of processes.

F. Lean and Six Sigma Integration with TQM

The Lean-Six Sigma-TQM triad is the most complete quality management system of the contemporary mechanical engineering companies. The principles of lean (5S, value stream mapping, continuous flow, pull production) eradicate non-value-added processes and minimize waste in the process. Six Sigma tools are used to solve residual variability and chronic defects. TQM also offers the cultural and organizational infrastructure, including leadership, training, measurement systems and customer focus, which support both lean and Six Sigma programs in the long term [12].

TABLE I. COMPARISON OF TRADITIONAL VS. TQM-BASED MECHANICAL ENGINEERING PROCESSES

Dimension	Traditional Approach	TQM-Based Approach
Quality Philosophy	Inspection after production; defect detection	Prevention-oriented; defect elimination at source
Defect Handling	Rework and scrap accepted as inevitable	Zero-defect goal; root cause analysis mandatory
Employee Role	Operators follow instructions passively	Active involvement; quality circles; kaizen teams
Data Utilization	Limited; intuition-based decisions	Extensive SPC, FMEA, DOE for data-driven decisions
Process Focus	Output-oriented; end-of-line inspection	Process-oriented; in-line and real-time monitoring
Supplier Relations	Adversarial; lowest-cost selection	Partnership-based; supplier quality development
Maintenance	Breakdown/reactive maintenance	TPM; predictive and preventive maintenance
Customer Focus	Internal efficiency prioritized	Customer requirements drive all process decisions

Continuous Improvement	Improvement project-based and sporadic	Systematic PDCA/DMAIC cycles; institutionalized
Quality Cost	High failure costs; low prevention investment	Shift to prevention; significant reduction in total cost

V. CASE STUDY: INDUSTRIAL APPLICATION IN INDIAN AUTOMOTIVE MANUFACTURING

A. Organizational Background

The case study is about a mid-sized Indian manufacturer of automotive components with headquarters in Pune, Maharashtra, which is specialized in the manufacture of precision-machined engine parts such as connecting rods, crankshafts, and cylinder head assemblies. With a staff of around 850 people and customers in the Tier-1 automotive OEMs, in 2019, the company started a complete TQM implementation program after a huge number of warranty claims by its customers, and an internal audit showing process capability indicators (Cpk) of less than 1.0 on key dimensions [4].

B. TQM Implementation Approach

The TQM was implemented in four phases, in a 30-month time. Phase 1 (Months 1-6) was devoted to organizational preparation: training of top management about the principles of TQM, creation of Quality Management Council, and measurement of all important process indicators at the baseline. Phase 2 (Months 7-15) focused on the process standardization: 5S in all shop floors, standardization of work instructions in all critical operations, and SPC on 23 critical-to-quality (CTQ) machining parameters [9].

Months 16-24 (Phase 3) presented more sophisticated quality instruments: FMEA was undertaken on all the critical processes, Six Sigma Green Belt project was implemented with an aim of connecting rod bore diameter fluctuation, and TPM was started in machining department with an aim of enhancing OEE. Phase 4 (Months 25-30) entrenched the gains in the form of quality audits, supplier quality development programs and development of Kaizen suggestion systems [14].

C. Results Achieved

The implementation of TQM resulted in quantifiable gains in all the performance indicators that were monitored. The internal defect rate has reduced by a factor of 86 by 4,850 parts per million (PPM) to 680 PPM. The total equipment performance within the machining department was increased by 61 percent to 78 percent with the set-up time and minimal stoppage losses as the main factors. The Six Sigma project of connecting rod bore diameter had a Cpk improvement of 0.87 to 1.68 which is well above the minimum of 1.33 specified by the customers [8].

The instances of customer complaints decreased by 74 percent throughout the 30-month period, and the company managed to receive TS 16949 (IATF 16949:2016) certification, which is a compulsory check to be considered an automotive supplier. The percentage of employees involved in Kaizen activities rose drastically to 0.00 to an average of 3.2 improvement suggestions per employee in a year. The percentage of poor quality (COPQ) as a percentage of revenue reduced to 1.9% as compared to 5.8 percent, which has saved the estimated INR 2.4 crores annually [15].

VI. RESULTS AND DISCUSSION

The results of both the literature review and the industrial case study substantiate together the idea that the integration of TQM brings about profound and dimensional achievements in the performance of mechanical engineering processes. The outcomes of the case study support the empirical studies, where Antony et al. [9] reported an average defect reduction of 65-80% in Indian manufacturing companies after organized TQM implementation.

Table I is the comparative analysis that clearly outlines the fundamental paradigm shift that TQM requires. The shift in focus of quality management, which is based on output and inspection, to quality assurance based on processes and prevention is not just a procedural change but a significant cultural change. The successful case study organization showed that the commitment of the leadership over 30 months was a key to this cultural transformation, which is in line with Patel and Desai [10] and Oakland [6].

The suggested conceptual model (Fig. 1) manages to place TQM components into distinct mechanical engineering areas and offer a roadmap of integration to practitioners. The two-way feedback structure of the framework, being indicative of the PDCA cycle, makes sure that the gains of improvement are not only realized, but reinforced and multiplied with time. It is especially true in the framework of Industry 4.0 where digital sensors and real-time analytics provide unprecedented opportunities in closed-loop quality control [11].

The combination of Lean and Six Sigma-TQM integration that is presented in the framework and the case study highlights the synergistic worth of integrating these approaches. Whereas Lean is used to remove structural waste and Six Sigma to deal with statistical variability, TQM offers the organizational structure of guaranteeing that both strategies are implemented in a systematic manner and replenished on a continuous basis. Tortorella et al. [12] also obtained the same results, stating that such tripartite integration has superior performance outcomes compared to the single-methodology approach.

The case study also shows an OEE improvement of 61-78 that can be considered quite remarkable as it represents the direct effect of the TPM integration in the TQM model. A study by Nakajima [13] defined 85% OEE as a global standard and the case organization is on its way to this objective within 30 months of TQM adoption, which proves the practicality of TQM-driven operational excellence management in Indian manufacturing.

VII. CHALLENGES AND LIMITATIONS

Although the advantages have been proven, TQM implementation in the mechanical engineering process is not without serious challenges. The barriers identified by both literature review and the case study analysis included the following:

- Organizational Resistance: The cultural shift necessary with TQM, that of individual responsibility to one of shared responsibility in terms of quality, is often met with opposition by the workers and the middle management. The long-term system of change management is needed to

overcome established behaviors and incentive systems that favor output and not quality [6].

- Resource Intensity: Implementation of comprehensive TQM requires a lot of investment in training, measurement tools, computer programs and consultancy skills. This financial cost may be prohibitive to Indian SMEs operating on thin margins, with a partial or superficial adoption of TQM practices that may not produce a lasting outcome [4].
- Resource Intensity: Comprehensive TQM implementation demands substantial investment in training, measurement equipment, software systems, and consultant expertise. For Indian SMEs operating on thin margins, this financial burden can be prohibitive, often resulting in partial or superficial TQM adoption that fails to deliver sustained results [4].
- Measurement System Limitations: Accurate quality data is the lifeblood of TQM. Many mechanical engineering enterprises in developing economies lack adequate measurement infrastructure—calibrated gauges, coordinate measuring machines, and data acquisition systems—necessary for meaningful SPC and process capability analysis.
- Skill Gaps: The analytical demands of Six Sigma, DOE, FMEA, and SPC require a technically proficient workforce. The shortage of engineers and technicians trained in statistical quality methods represents a critical bottleneck in TQM implementation, particularly in tier-2 and tier-3 industrial cities [9].
- Sustaining Momentum: Many organizations experience initial quality improvements following TQM launch, only to experience regression as management attention shifts to other priorities. Institutionalizing TQM requires embedding quality metrics in performance management systems and maintaining leadership visibility over the long term.
- Integration with Digital Systems: As Industry 4.0 technologies proliferate, the integration of traditional TQM tools with IoT platforms, ERP systems, and AI-driven analytics poses both technical and organizational challenges that current TQM frameworks have not fully addressed [11].

This paper acknowledges certain limitations. The case study relies on secondary data from a single organization, limiting generalizability across diverse mechanical engineering sectors. Additionally, the proposed conceptual framework requires empirical validation through large-scale survey research across multiple Indian manufacturing firms.

VIII. CONCLUSION AND FUTURE SCOPE

This paper has provided an in-depth study of the integration of TQM in the mechanical engineering processes including a systematic literature review, a conceptual integration framework, comparative process analysis, and an exemplary Indian automotive manufacturing case study. The evidence shows conclusively that TQM when strategically implemented through the manufacturing systems, production planning, quality control, maintenance engineering and lean operations provides well-founded changes in product quality, process efficiency, and organizational competitiveness.

The three-tier TQM Integration Framework proposed, which includes the strategic leadership, core engineering operations, and quantifiable results, offers the mechanical engineering practitioners and quality managers a roadmap to TQM implementation that is logical and practical. It is the comparative analysis that brings into light the transformational concept of TQM as opposed to the traditional quality management and through the Pune automotive case study, the study shows its practical applicability in the Indian manufacturing environment.

The results also point to the fact that a combination of TQM and Lean Manufacturing and Six Sigma strategies forms a synergistic quality management system that is more successful than any single approach strategy. The adoption of TPM into this framework is particularly singled out as a pivotal facilitator of long-term OEE enhancement in mechanical engineering settings.

Future studies ought to take some of the following directions;

(i) the empirical validation of TQM maturity model specifically tailored to the Indian mechanical engineering SMEs;

(ii) the study of the integration of the TQM principles with Industry 4.0 and cyber-physical systems, the role of artificial intelligence and machine learning in enhancing the data-driven decision-making of TQM further;

(iii) longitudinal research into the sustainability of TQM benefits in high-mix, low-volume mechanical production settings; and

(iv) investigation into green TQM models that take quality management to both environmental and sustainability levels, in line with new ESG requirements in Indian industry.

References.

[1] W. E. Deming, Out of the Crisis. Cambridge, MA: MIT Press, 2018 (reprint).

[2] J. S. Oakland, Total Quality Management and Operational Excellence: Text with Cases, 4th ed. New York: Routledge, 2014.

[3] International Organization for Standardization, ISO 9001:2015 Quality Management Systems — Requirements, Geneva, Switzerland: ISO, 2015.

[4] R. S. Kadam and A. P. Kulkarni, "Barriers to TQM implementation in Indian small and medium manufacturing enterprises," International Journal of Quality and Reliability Management, vol. 37, no. 4, pp. 621–643, 2020.

[5] J. M. Juran and J. A. De Feo, Juran's Quality Handbook: The Complete Guide to Performance Excellence, 7th ed. New York: McGraw-Hill, 2017.

[6] J. S. Oakland and M. Tanner, "Lean and Six Sigma — Leadership for the 21st Century," Total Quality Management and Business Excellence, vol. 27, no. 5–6, pp. 564–574, 2016.

[7] D. H. Besterfield, C. Besterfield-Michna, G. H. Besterfield, and M. Besterfield-Sacre, Total Quality Management, 4th ed. Upper Saddle River, NJ: Pearson, 2019.

[8] T. Pyzdek and P. A. Keller, The Six Sigma Handbook, 5th ed. New York: McGraw-Hill, 2018.

[9] J. Antony, M. Kumar, and C. N. Madu, "Six Sigma in small and medium-sized UK manufacturing enterprises," International Journal of Quality and Reliability Management, vol. 22, no. 8, pp. 860–874, 2020.

[10] H. Patel and M. Desai, "TQM practices in small-scale mechanical industries: An empirical study from Gujarat," Journal of Quality in Maintenance Engineering, vol. 26, no. 3, pp. 412–429, 2020.

[11] K. Salonitis and C. Tsinopoulos, "Drivers and barriers of lean implementation in the Greek manufacturing sector," Procedia CIRP, vol. 57, pp. 189–194, 2017.

[12] G. L. Tortorella, D. Fettermann, A. Frank, and G. Marodin, "Lean manufacturing implementation: Leadership styles and contextual variables," International Journal of Operations and Production Management, vol. 38, no. 5, pp. 1205–1227, 2018.

[13] S. Nakajima, Introduction to TPM: Total Productive Maintenance. Cambridge, MA: Productivity Press, 2015 (reprint).

[14] M. Imai, Gemba Kaizen: A Commonsense Approach to a Continuous Improvement Strategy, 2nd ed. New York: McGraw-Hill, 2012.

[15] Automotive Research Association of India (ARAI), "Quality Management Practices in Indian Automotive Supply Chain: Industry Report 2022," Pune, India: ARAI, 2022.

Performance Evaluation of Various Feed Water Approaches in Passive Solar Still Desalination

D'Cotha J.S[a,*], Suresh P. R.[b], Sajeesh P[c]

[a]*Department of Mechanical Engineering, SCMS School of Engineering and Technology, Kerala, India.*
[b]*Department of Mechanical Engineering, Ahalia School of Engineering and Technology, Kerala, India.*
[c]*Department of Mechanical Engineering, NSS College of Engineering, Palakkad, Kerala, India-678008.*

Abstract **- Access to safe drinking water is a growing global challenge driven by rapid industrialization, population growth, and climate change. Solar desalination has emerged as a sustainable, decentralized solution for water purification, particularly in remote and arid regions. A critical yet often overlooked factor governing the efficiency of solar stills is the feed water management approach — the manner in which brackish or saline water is introduced, maintained, and circulated within the still. This paper reviews various feed water strategies employed in solar still research, including no feed water addition, periodic addition at fixed intervals, continuous flow systems, spray-type feeding, addition of dye and saline solutions, and preheating of feed water prior to introduction into the basin. The review highlights that shallow water depths, preheated feed water, and controlled replenishment intervals consistently yield higher distillate outputs. Preheating feed water using evacuated tube collectors, solar ponds, copper coils, and heat pipes has demonstrated productivity improvements ranging from 31.8% to 100% over conventional configurations. Spray feeding and thin film techniques further enhance heat and mass transfer rates. The findings underscore that optimizing feed water temperature, depth, and replenishment frequency are pivotal to maximizing solar still performance, and that integrating renewable thermal sources for preheating represents a highly promising direction for future development.**

Keywords - Stepped Still, Depth of water, Feed water

1. INTRODUCTION

Access to safe drinking water remains one of the prominent global challenges in the 21st century. Rapid industrialization, increasing population and gradually changing climates has increased the demand for freshwater. Freshwater scarcity remains a critical concern worldwide, especially in arid, semi-arid, and remote regions where access to potable water is limited. The need for freshwater scarcity has emerged as a critical threat to human health, food security, and sustainable development. Conventional methods of water purification in large scales often require hefty investment on desalination units and significant consumption of energy that generally rely on non-conventional sources of energy. These concerns and challenges pave the need for developing desalination methods which independently work on renewable energy resources. This paradox has intensified the search for affordable, energy-efficient, and environmentally sustainable desalination technologies.

Compared to different desalination methods, solar desalination systems are considered as a sustainable solution for producing decentralized, off-grid water purification in remote or rural areas. Solar desalination mimics the hydrological cycle, which includes, evaporation of water by the solar radiation, condensation by the glass cover and the droplets collected by the collection channel. But this method is classified as a low volume yielding one, and this has motivated researchers to study the modifications to basin geometry, cover materials, thermal energy storage, heat recovery mechanisms, and crucially, feed water management approaches.

As water gets evaporated from the basin during the day, there is a decrease of water levels in the basin. Water is fed into the basin periodically to maintain the depth of water at a constant value. The temperature of inlet feed water fed into the basin to compensate for the level depends on the surrounding temperatures and the container in which it holds. The ambient temperature variations depend on the season type, sky conditions, time of the day, the latitude of the location etc. Since the ambient temperature is never constant throughout the day, it is more likely that the temperature of inlet feed water is also not constant.

2. FEED WATER METHODS

The feed water approach namely the manner in which the brackish water is introduced, maintained and circulated inside the still plays a pivotal role in governing the evaporation rates and overall system efficiency. Solar still rely on solar energy to heat the

shallow layer of saline or brackish water, causing it to evaporate while leaving contaminants behind. The water vapor then rises and condenses on the cooler inner surface of the condensate cover, these droplets slide below the surface of the glass cover to flow into a collection channel. Hence it is necessary for stills to begin with a fixed minimum water depth to ensure consistent thermal behavior and reliable performance,

2.1. NO FEED WATER ADDITION

In the experiments conducted by few of the researchers [1], [2], the initial depth of the water maintained in the still basin before the commencement of experiments reduces as the evaporation process progresses. Hoque et al. [2] assessed the influence of brine water depth and salt concentration on the productivity of a single-slope passive solar still operating under the climatic conditions of Chittagong, Bangladesh. No feed water was added and as the initial depth of water was increased, freshwater yield declined. The productivity decreased by 18.75% when the brine depth was increased from 1.6 cm to 2.1cm. Similarly, Phadatare et al. [1] used a transparent conventional solar still without feed addition, varying the basin water depth between 2 cm and 12 cm. The results indicated that as the depth increased, the distillate output decreased from 2.1 L/m² to 1.85 L/m², highlighting that maximum productivity is achieved at the minimum water depth.

2.2. FEED WATER ADDITION

Researchers have also reported the productivity of the still with constant depth of water in the tray throughout the period of the experiments by adding feed water every hour[3], [4], [5], [6], [7] or every half hour [8], [9], [10] to compensate for the evaporated quantity.

Omara et al. [5], used a stepped solar still with internal mirrors on the vertical sides of the trays. Experiments with conducted by constantly maintaining an initial depth of water of 0.5cm in the trays by feeding water in interval of every hour. The daily productivity of the stepped solar still without mirrors reaches 5,840 ml/m²/day compared to 3,720 ml/m²/day for the conventional still, representing a 57% increase. When mirrors are added, the stepped still further improves to 6,350 ml/m²/day against 3,630 ml/m²/day for the conventional still, achieving a 75% higher productivity.

The stepped solar still features an internal vertical evaporator fixed to the rear wall, consisting of 18 superimposed triangular trays welded to a vertical rectangular galvanized plate measuring 36 cm × 68 cm. The water mass and depth in each still are maintained constant throughout testing by replenishing water equal to the collected distillate every half hour. Temperature fluctuations are caused by the relatively cold feed water added every 30 minutes to the basin still, whose effect is minimal in the morning due to the small refill volume and low temperature difference, becomes significant between 11.00 hours and 16.00 hours when both the added water mass and the basin–feed temperature difference (about 60 °C) increase and removes substantial amount of heat [8].

Dcotha et al. [11] enhances solar still productivity by employing a dynamic water depth strategy and feed water addition using water as a sensible heat storage medium, minimizing carbon footprint while optimizing costs. Experiments were conducted at three initial water depths of 0.5 cm, 1.0 cm, and 1.5 cm, under three feed water conditions: no addition, 100 ml added per tray, and 200 ml added per tray, by adding feedwater at 9:00 am, 10:00 am, and 11:00 am. The optimal configuration of 0.5 cm initial depth with 400 ml feed water per tray achieved the highest output of 6,330 ml/m²/day at a maximum daily efficiency of 66.45%.

2.3. CONTINUOUSLY FLOWING FEED WATER

In the experimental setup of Gawande et al. [12] saline water is supplied to the solar still from a storage tank via a PVC hose. The stepped trays are sequentially filled from top to bottom, with excess water exiting through a bottom overflow channel and collected for reuse. As evaporation occurs, the water level in the trays drops; to compensate for this loss, makeup water is added from the storage tank at half-hour intervals. Similar method of operation was adopted by Velmurugan et al. [10] in a modified stepped solar still incorporating 50 trays on the absorber plate. Each tray measured 99 mm × 99 mm and was distributed across 10 steps. The first twenty-five trays maintained a water depth of 10 mm, while the remaining twenty-five trays operated at a reduced

depth of 5 mm.

Sudalaimuthua et. al. [13] fed water continuously from a constant head feed water tank into the copper common rail split passage system of an inclined solar still integrated with fins. The copper common rail split passage system, ensured uniform distribution and formation of a thin water film under dynamic conditions. The hot water from the bottom of the inclined stepped still is re-circulated and merged with the main feed water line. The feed water supply was varied at flow rates of 0.285, 0.635, 0.8, and 1 kg/min, resulting in corresponding productivity of 2.5, 1.2, 0.9, and 0.7 kg/m², respectively, indicating that lower flow rates enhance distillate yield due to lower water film thickness.

2.4. SPRAY TYPE FEED WATER

Elgendi et al. [14] maintained a constant depth of water in the pyramid solar still using an automatic feed water system which was connected to an auxiliary tank. The cumulative distillate yield improved by 37% when the depth of water in the still was reduced from 5 cm to 0.5 cm. The operating temperature of the basin saline water increases, which occurs with a decrease in water depth. To leverage this, a transient thin water film can be maintained on a corrugated stepped absorber using a reciprocating spray feeding system [15]. The small droplet size of the sprayed water significantly increases the surface area, enhancing heat and mass transfer rates, thereby improving overall productivity. Operating for 10 hours during the day, the solar still achieved a distillate yield of 6.355 L/m² with a thermal efficiency of 77.35%.

2.5. ADDITION OF DYE AND SALINE WATER

The use of violet dye (35 mg/L) and a reflector (2.8 m × 1.4 m) in a double-slope single-basin solar still enhanced performance by increasing solar absorption, raising basin temperatures by over 7 °C on average. This resulted in a significant improvement in productivity, with maximum efficiency reaching 19.1% for violet dye compared to 14.8% for plain water [16].

To study the effect of salinity on the quality of yield produced, Mika [17] prepared feed water with three salinity levels by dissolving 0, 20, and 40 g/kg NaCl in tap water, The initial quantity of feed water in the still was 10kg and the water level was maintained constant using a feeding tank. Experimental results showed that the solar still (SS) effectively removed over 98% of total dissolved solids (TDS), achieving productivity between 1084 and 5014 ml/m²/day, with water temperatures reaching up to 80.4 °C.

In a conventional solar still, Hoque et al. [2] used synthetic saline solutions in the single basin still by preparing dissolved solids (TDS) of 2000, 5000, and 8000 ppm by adding sodium chloride. The results indicate that increasing salinity adversely affects freshwater production, with a 7.28% reduction in yield observed as TDS increases from 2000 to 8000 ppm. At the same time, water quality parameters improved significantly, with over 99% elimination of impurities and dissolved salts. All tested parameters complied with WHO guidelines, confirming the distilled water is safe for drinking.

2.6. PREHEATING OF FEED WATER

Sharshir et al. [18] employed a single basin solar still with cold water film cooling over the glass cover at flow rates ranging from 1 kg/h to 12 kg/h, and incorporated graphite micro-flakes to enhance thermal performance and productivity. Reusing the exhausted glass cooling water as makeup water increased the distillate yield by approximately 57.60% compared to the conventional still.

Saadi et al. [8] modified the conventional still by attaching 18 triangular trays on the rear wall of the still and the evaporated water was compensated by adding feed water every half-hour. He also tested two identical basin-type conventional solar stills under identical atmospheric conditions; one was fed with water at ambient temperature, while the other received preheated water at the same temperature as the basin water. During the morning hours (08:00 am to 11:00 am), the influence of feed water on still performance is minimal due to the low volume added and the small temperature difference between the basin and feed water. However, from 11:00 am to 04:00 pm, as distillate production rises, both the volume of added feed water and the temperature difference ($\Delta T \approx 60$ °C) increases. Whereas for the still supplied with feed water at ambient temperature, experienced considerable heat loss from the basin each time new feed water is introduced. This causes water temperature fluctuations, whereas the still supplied

with preheated water showed stable temperatures.

Abdullah et. al. [19] allowed feed water to from the feeding tank through a copper heating coils (spiral collector) to a float-regulated system that maintains a constant water level at the base of the solar still. Once the desired depth is reached, the float valve halts the inflow until the level drops, automatically compensating for the loss. Solar radiation simultaneously heats both the basin water and the feed water in the copper coils, enhancing thermal efficiency. The feed water temperature in the Modified Solar Still (MSS) was 0–10 °C higher than that in the Conventional Solar Still (CSS). At 13:00, maximum basin water temperatures reached 75 °C for the MSS and 63 °C for the CSS. Correspondingly, hourly productivity at this time was approximately 800 ml/m² for the MSS and 510 ml/m² for the CSS. Overall, the cumulative yield of the MSS (5150 ml/m²) was 66% greater than that of the CSS (3100 ml/m²).

Velmurugan et al. [20] conducted experiments to study the distillate output obtained from single basin and stepped solar still, while it was coupled with a solar pond to preheat the industrial effluent fed into the still. After filtration the effluent was allowed to settle in a collection tray for at least two hours before storing it in a storage tank. Preheating of the industrial effluent was done by passing it through a swirled copper tube immersed in the solar pond as shown in Figure 1. The preheated effluent is fed into the stepped and conventional still and the evaporation process reduces the level of liquid in the tray/basin. To maintain constant water in the basin, feed water is added every half an hour. When fins were used, the productivity rise was 53%. Further addition of heat sensible materials like pebbles, enhanced the productivity of the still by 63%. Combination of preheated water with fins, sponges and pebbles increased the productivity by 100%.

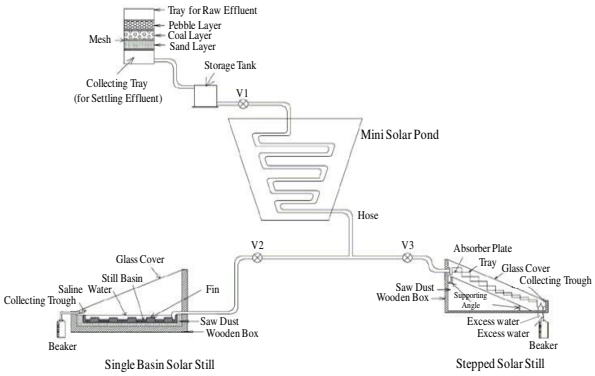


Figure 1. Single basin and stepped basin solar still fed with preheated industrial effluent [20].

Kabeel et al. [6] fed hot feed water from an evacuated tube solar collector into a conventional still and stepped solar still and the productivity of the stepped still was 66.6% greater than conventional still. When both the stills were operated with feed water supplied at ambient conditions the stepped still produced 57.3% more yield than the conventional still. By using the solar collector, the feed water was preheated, causing the feed water temperature to rise from 33°C to 86°C. Addition of feed water at elevated temperature reduces the time required for formation of vapors inside the still.

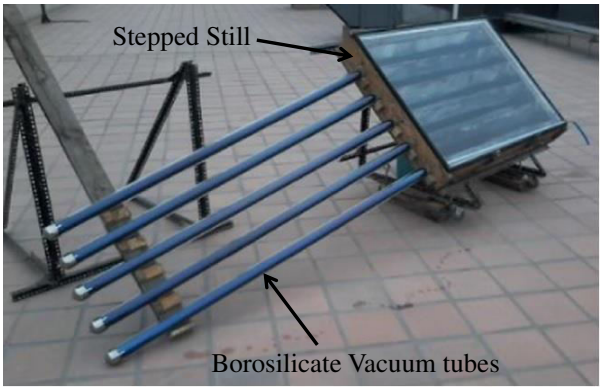


Figure 2. Stepped solar still with internal reflectors and borosilicate vacuum tubes with heat pipes to preheat feed water [21]

Similarly, Saettone et al. [21] conducted experiments on stepped solar still coupled with a set of five borosilicate vacuum tubes running parallel to each other as shown in Figure 2 to provide an additional heat transfer into the stepped still. Each vacuum tube contains a heat pipe concentric to itself, which is used to transfer heat to the basin water. The heat pipes pass

through the tray to facilitate heat transfer to the saline water in the trays. Precautions are made to prevent mixing of the fluid inside the heat pipe and the saline water. The trays had internal reflectors with 2 cm depth of water in each tray which is nearly 2.7 liters of saline water. Stepped still with vacuum tubes produced 31.8% more yield compared to the normal stepped still, even when the ambient temperature and the solar energy received was lower by 2°C and 0.6% respectively.

A concentrator dish with 1m diameter and focal length 0.25 m is used to intensify and direct the reflected radiations from the sun onto the crescent shaped absorber. The basin water is circulated from the basin to the absorber due to thermosyphon effect as shown in Figure 3. This absorber heats the circulating water before supplied to the basin. Finally, this heated water is fed into the basin as top layer. As the hot water is supplied into the basin, the temperature difference in the different layers of basin water created a density difference and forcing the lower density molecules migrate to the free surface due to the buoyancy effect [22].

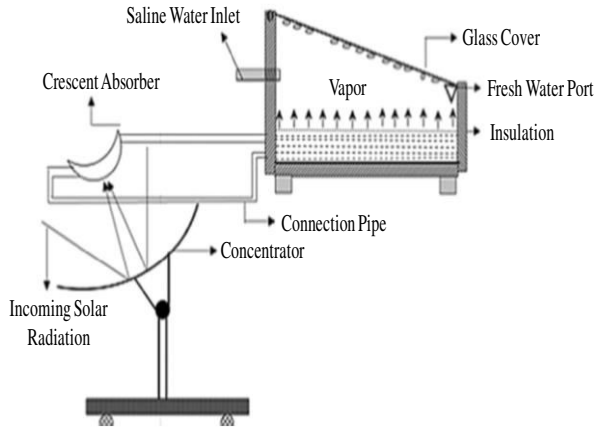


Figure 3 Schematic diagram of single slope solar still coupled with crescent shaped absorber [22].

3. CONCLUSION

The depth of water maintained in the still basin is inversely related to distillate yield. Studies consistently demonstrated that reducing basin depth — from as much as 12 cm down to 0.5 cm — improved evaporation rates and overall output, with one study reporting a 37% improvement upon reducing depth from 5 cm to 0.5 cm.

Periodic feed water addition, whether at 30-minute or 60-minute intervals, helps sustain consistent evaporation by maintaining a shallow, thermally favorable water layer. However, the introduction of relatively cold feed water between 11:00 and 16:00 hours cause notable basin temperature fluctuations and associated heat losses, with temperature differentials reaching approximately 60°C during peak hours.

Preheating feed water by using evacuated tube solar collectors, solar ponds with swirled copper tubes, copper heating coils, and heat pipe-integrated vacuum tubes, researchers achieved productivity gains ranging from 31.8% to 100% over conventional stills. Preheated feed water stabilizes basin temperatures, reduces the thermal lag period, and accelerates vapor formation.

Spray-type and thin film feeding mechanisms significantly increase the water surface area exposed to solar radiation, enhancing heat and mass transfer. A solar still utilizing a reciprocating spray feeding system achieved a distillate yield of 6.355 L/m² with a thermal efficiency of 77.35%.

Salinity and water quality studies confirmed that solar stills effectively remove over 98–99% of total dissolved solids across a range of salinity levels, with all output water meeting WHO drinking water standards, validating their suitability for treating brackish and saline feeds.

REFERENCES

1] M. K. Phadatare and S. K. Verma, “Influence of water depth on internal heat and mass transfer in a plastic solar still,” *Desalination*, vol. 217, pp. 267–275, 2007, doi: 10.1016/j.desal.2007.03.006.

[2] A. Hoque, A. H. Abir, and K. Paul Shourov, “Solar still for saline water desalination for low-income coastal areas,” *Appl. Water Sci.*, vol. 9, no. 4, 2019, doi: 10.1007/s13201-019-0986-9.

[3] A. S. Abdullah, M. M. Younes, Z. M. Omara, and F. A. Essa, “New design of trays solar still with enhanced evaporation methods – Comprehensive study,” *Sol. Energy*, vol. 203, no. April, pp. 164–174, 2020, doi: 10.1016/j.solener.2020.04.039.

[4] A. F. Muftah, K. Sopian, and M. A. Alghoul, “Performance of basin type stepped solar still enhanced with superior design concepts,” *Desalination*, vol. 435, pp. 198–209, 2018, doi: 10.1016/j.desal.2017.07.017.

[5] Z. M. Omara, A. E. Kabeel, and M. M. Younes, “Enhancing the stepped solar still performance using internal reflectors,” *Desalination*, vol. 314, pp. 67–72, Apr. 2013, doi: 10.1016/J.DESAL.2013.01.007.

[6] A. E. Kabeel, A. Khalil, Z. M. Omara, and M. M. Younes, “Theoretical and experimental parametric study of modified stepped solar still,” *Desalination*, vol. 289, pp.

12–20, 2012, doi: 10.1016/j.desal.2011.12.023.

[7] S. W. Sharshir, M. A. Eltawil, A. M. Algazzar, R. Sathyamurthy, and A. W. Kandeal, "Performance enhancement of stepped double slope solar still by using nanoparticles and linen wicks : Energy , exergy and economic analysis," *Appl. Therm. Eng.*, p. 115278, 2020, doi: 10.1016/j.applthermaleng.2020.115278.

[8] Z. Saadi, A. Rahmani, S. Lachtar, and H. Soualmi, "Performance evaluation of a new stepped solar still under the desert climatic conditions," *Energy Convers. Manag.*, vol. 171, no. March, pp. 1749–1760, 2018, doi: 10.1016/j.enconman.2018.06.114.

[9] J. S. Gawande, L. B. Bhuyar, and S. J. Deshmukh, "Effect of Depth of Water on the Performance of Stepped Type Solar Still," *Energy Power Eng.*, vol. 03, no. 4, pp. 137–143, 2013, doi: 10.4236/epe.2013.58053.

[10] V. Velmurugan, S. S. Kumaran, V. Niranjan Prabhu, and K. Srithar, "Productivity enhancement of stepped solar still - Performance analysis," *Therm. Sci.*, vol. 12, no. 3, pp. 153–163, 2008, doi: 10.2298/TSCI0803153V.

[11] J. S. D. Cotha, P. R. Suresh, and P. Sajeesh, "A novel approach to augment stepped solar still productivity using the heat storage characteristics of water," vol. 11, no. 5, pp. 1355–1373, 2025, doi: 10.14744/thermal.0000981.

[12] J. S. Gawande and L. B. Bhuyar, "Effect of Shape of the Absorber Surface on the Performance of Stepped Type Solar Still," *Energy Power Eng.*, vol. 05, no. 08, pp. 489–497, 2013, doi: 10.4236/epe.2013.58053.

[13] P. Sudalaimuthu and R. Sathyamurthy, "Experimental investigation of a novel-designed inclined solar still with bio-wick under dynamic water flow," *Aqua Water Infrastructure, Ecosyst. Soc.*, vol. 73, no. 5, pp. 835–852, 2024, doi: 10.2166/aqua.2024.239.

[14] M. Elgendi and M. Y. E. Selim, "Determination of the effect of water depth on the yield of a solar still using an automatic feedwater system," *Environ. Sci. Pollut. Res.*, vol. 29, pp. 14595–14603, 2021, doi: https://doi.org/10.1007/s11356-021-15707-3.

[15] A. M. El-Zahaby, A. E. Kabeel, A. I. Bakry, S. A. El-Agouz, and O. M. Hawam, "Enhancement of solar still performance using a reciprocating spray feeding system-An experimental approach," *Desalination*, vol. 267, no. 2–3, pp. 209–216, 2011, doi: 10.1016/j.desal.2010.09.028.

[16] S. Nijmeh, S. Odeh, and B. Akash, "Experimental and theoretical study of a single-basin solar sill in Jordan," *Int. Commun. Heat Mass Transf.*, vol. 32, no. 3–4, pp. 565–572, 2005, doi: 10.1016/j.icheatmasstransfer.2004.06.006.

[17] Ł. Mika and E. Radomska, "Solar Energy to Water Desalination: Long-Term Experimental Studies of Solar Still in Poland," *Energies*, vol. 18, no. 5, 2025, doi: 10.3390/en18051070.

[18] S. W. Sharshir *et al.*, "Enhancing the solar still performance using nanofluids and glass cover cooling: Experimental study," *Appl. Therm. Eng.*, vol. 113, pp. 684–693, 2017, doi: 10.1016/j.applthermaleng.2016.11.085.

[19] A. S. Abdullah, L. Hadj-Taieb, M. M. Hikal, Z. M. Omara, and M. M. Younes, "Enhancing a solar still's performance by preheating the feed water and employing phase-change material," *Alexandria Eng. J.*, vol. 77, pp. 395–405, 2023, doi: 10.1016/j.aej.2023.07.002.

[20] V. Velmurugan, J. Mandlin, B. Stalin, and K. Srithar, "Augmentation of saline streams in solar stills integrating with a mini solar pond," *Desalination*, vol. 249, no. 1, pp. 143–149, 2009, doi: 10.1016/j.desal.2009.06.016.

[21] E. Saettone, Y. Valencia-Tovar, and A. Gómez-de-la-Torre-Gastello, "Preliminary overview and evaluation of a stepped solar distiller with internal reflective walls and borosilicate vacuum tubes," *Desalination*, vol. 413, pp. 136–143, 2017, doi: 10.1016/j.desal.2017.03.020.

[22] T. Arunkumar, R. Velraj, D. C. Denkenberger, and R. Sathyamurthy, "Influence of crescent shaped absorber in water desalting system," *Desalination*, vol. 398, pp. 208–213, 2016, doi: 10.1016/j.desal.2016.07.039.

Acoustic Streaming Effects on the Atomization of a Spray

Kasif Ehsan[1] ***Aadil Kureshee*** **[2]** ***Deepak Kumar Mandal*** **[3]** ***S. Narayanan*** **[4]**

[1]*Dept. of Mechanical Engineering, IIT ISM, Dhanbad, India*
24dr0084@iitism.ac.in
2Dept. of Mechanical Engineering, IIT ISM, Dhanbad, India
Aadilkureshee.20@mech.iitism.ac.in

3Dept. of Mechanical Engineering, IIT ISM, Dhanbad, India
deepak@iitiam.ac.in

4Dept. of Mechanical Engineering, IIT ISM, Dhanbad, India
narayan@iitiam.ac.in

Abstract**—**This paper presents an experimental investigation of the effect of acoustic excitation frequency on fuel spray cone angle at two injection pressures using Particle Image Velocimetry (PIV). Five test conditions are examined: injection pressures of 30 bar and 40 bar, each subjected to acoustic streaming at frequencies of 30 Hz and 200 Hz (at 30 bar), and 30 Hz, 100 Hz, and 200 Hz (at 40 bar). Spray cone angle is extracted from PIV images using intensity-based thresholding and edge-fitting algorithms implemented in MATLAB. The measured cone angles are: 16.48° (30 bar, 30 Hz), 22.46° (30 bar, 200 Hz), 18.67° (40 bar, 30 Hz), 18.75° (40 bar, 100 Hz), and 19.77° (40 bar, 200 Hz). A notable finding is the pressure–frequency interaction: at low acoustic frequency (30 Hz), the higher injection pressure (40 bar) produces a wider cone angle, while at 200 Hz this trend reverses, with the 30 bar condition producing the wider cone and higher cone angle enhance the atomization of the spray. This cross-over behaviour is attributed to competing effects of injection momentum and acoustic radiation force on the spray boundary.

Keywords**—***fuel spray; PIV; acoustic excitation; acoustic streaming; spray cone angle; injection pressure; atomization*

I. INTRODUCTION

Spray cone angle is a primary morphological parameter governing fuel–air mixture formation in direct-injection combustion systems. A wider cone angle promotes radial fuel dispersion and enhances mixing, while a narrower cone favours deeper penetration. The ability to control cone angle dynamically through external forcing has therefore attracted significant research interest [1].

Acoustic excitation — the application of controlled acoustic fields to modify spray dynamics — has emerged as a promising non-intrusive technique. Through the mechanism of acoustic streaming, steady secondary flows are induced within the spray field by acoustic pressure gradients, imposing lateral momentum on the spray boundary and thereby modifying cone angle [2]. The magnitude of this effect is frequency-dependent: higher acoustic frequencies produce stronger streaming velocities at the same sound pressure level, following the Eckart streaming model in which streaming velocity scales with f^2 [2].

Particle Image Velocimetry (PIV) is an established optical diagnostic providing full-field, non-intrusive measurement of spray velocity and intensity fields [3]. PIV-derived intensity images allow reliable extraction of spray cone angle through boundary detection and edge-fitting algorithms, avoiding the contact and flow disturbance associated with mechanical probes.

Several studies have examined acoustically forced sprays. Kumar et al. [4] reported increased cone angle in pressure-swirl atomizers under acoustic excitation, and attributed this to the acoustic radiation pressure acting on the spray boundary. Tiwari et al. [5] demonstrated frequency-dependent droplet size modification in liquid sprays excited at 20–200 Hz. Zhao et al. [6] showed that the interaction between injection pressure and acoustic forcing is non-linear, with higher injection pressures exhibiting diminishing sensitivity to acoustic forcing at low frequencies. However, a systematic experimental comparison of cone angle across multiple injection pressure and acoustic frequency combinations, particularly examining the relative influence of pressure and frequency, has not been reported for the present configuration.

The present study addresses this gap by reporting measured spray cone angles for five test conditions combining 30 bar and 40 bar injection pressures with acoustic excitation at 30 Hz, 100 Hz, and 200 Hz. A cross-over behaviour between the two pressure conditions is identified and discussed in the context of acoustic streaming theory. PIV images for each condition are available for further analysis.

II. EXPERIMENTAL SETUP AND METHODOLOGY

A. Spray Chamber and Injection System

The experiment involved particle image velocimetry (PIV) to observe the behavior of a methanol and diesel spray. A wooden cubic chamber (inner dimension ¼ 60 60 x 60 cm^3), whose internal walls were covered with

Polyurethane wedges (of about 6 cm thickness) to avoid sound reflections, was used as an enclosure for conducting experiments. One side of the chamber was kept open to facilitate the observation. The other two sides were provided with a small window (20 x 20 cm^2) to pass a laser beam from both sides. A fuel tank (inner dimension 15 x 15 x 30 cm^3), equipped with a fuel pump, was employed to store and efficiently transfer the liquid fuel from the tank to an injector. A precise fuel injector, with multiple holes having an inner diameter of 50 µm each, was utilized to ensure a controlled and effective fuel dispersal in spray. Methanol was used as a fuel. The injector was inserted in a wooden block by creating a 1.5 cm diameter hole to fix it firmly. The injector outlet was kept from the ground at 106 cm (vertical distance). The injector pressure control and liquid monitoring from the injector were executed using a precision control valve and a pressure gauge.

B. Acoustic Excitation System

Two satellite speakers (3W) were positioned symmetrically from the top of the fuel injector (at a vertical distance of 4.5 cm), facing downward. To introduce the acoustic from the speakers, two holes with the same diameter as the speaker (5 cm) were provided on either side of the fuel injector. At 30 bar, excitation was applied at 30 Hz and 200 Hz; at 40 bar, all three frequencies (30 Hz, 100 Hz, 200 Hz) were tested, as summarised in Table I.

C. PIV Imaging System

The laser (wavelength of 532 nm) was triggered from the laser head (make: Nitron Lasers, Model: S 65-15 PIV). A lens mounted on the laser head was used to transform the laser beam into a wide laser sheet. The laser source and camera were set up at 90 from each other to capture the data. The laser was positioned on a table at a height of 106 cm from the ground, while the wooden anechoic chamber was kept on another table. The adjustment of both table heights was carefully executed so that the laser accurately aligns with the spray at its point of origin, ensuring comprehensive coverage. The temperature inside the chamber was found to be 300 6 1 K with a relative humidity of 30%–35%.

D. Cone Angle Extraction Methodology

Spray cone angle was extracted from each PIV image using the following steps:

1. Pre-processing: Background subtraction, Gaussian smoothing (σ = 1.5 px), and normalisation to [0, 1] intensity range.
2. Binarisation: Otsu's adaptive thresholding applied to segment spray from background, followed by morphological closing (disk SE, r = 5 px) to fill gaps in the spray boundary.
3. Edge detection: Upper and lower spray boundaries identified by finding the first and last non-zero pixel in each column of the binary mask.
4. Cone angle fitting: Linear regression applied to the upper and lower boundary point sets (over 10–90% of the spray axial extent). The full cone angle θ is computed as the sum of the two edge slope angles relative to the spray axis:

$$\theta = arctan(m_{upper}) + |arctan(m_{lower})| \quad (1)$$

where m_{upper} and m_{lower} are the slopes of the upper and lower boundary regression lines respectively. Cone angle was computed for each image and ensemble-averaged per test condition. The standard deviation across images within each condition is reported as measurement uncertainty.

TABLE I. TEST CONDITIONS AND MEASURED CONE ANGLES

Case	P (bar)	f (Hz)	Cone Angle(degree)
1	30	30	16.48
2	30	200	22.46
3	40	30	18.67
4	40	100	18.75
5	40	200	19.77

III. RESULTS AND DISCUSSION

A. Cone Angle at 30 Bar: Strong Frequency Sensitivity

At 30 bar injection pressure, the spray cone angle increases substantially with acoustic frequency: from 16.48° at 30 Hz (P30-A30) to 22.46° at 200 Hz (P30-A200). This represents an absolute increase of 5.98°, or 36.3%, as acoustic frequency is raised from 30 Hz to 200 Hz at constant sound pressure level (130 dB SPL). This is the largest frequency-induced variation observed across all conditions in this study.

The strong sensitivity of the 30 bar cone angle to acoustic frequency is consistent with the Eckart acoustic streaming mechanism [2], in which the time-averaged streaming body force scales as f^2. At 200 Hz, the acoustic streaming velocity is approximately 44 times greater than at 30 Hz at the same SPL, imposing a substantially larger lateral momentum flux on the spray boundary. At 30 bar, where the injection axial momentum is relatively lower than at 40 bar, this lateral acoustic forcing constitutes a larger fraction of the total momentum budget of the spray, resulting in pronounced lateral spreading and a wide cone angle.

B. Cone Angle at 40 Bar: Weak Frequency Sensitivity

At 40 bar injection pressure, cone angle shows a markedly weaker dependence on acoustic frequency. The values are 18.67° at 30 Hz (P40-A30), 18.75° at 100 Hz (P40-A100), and 19.77° at 200 Hz (P40-A200). The total increase from 30 Hz to 200 Hz is only 1.10° (5.9%), compared to 5.98° (36.3%) at 30 bar. Notably, the step from 30 Hz to 100 Hz produces virtually no change (0.08°, 0.4%), while the step from 100 Hz to 200 Hz yields a 1.02° (5.4%) increase.

This reduced frequency sensitivity at 40 bar is attributed to the higher axial injection momentum at the elevated pressure. The dynamic pressure of the injected spray

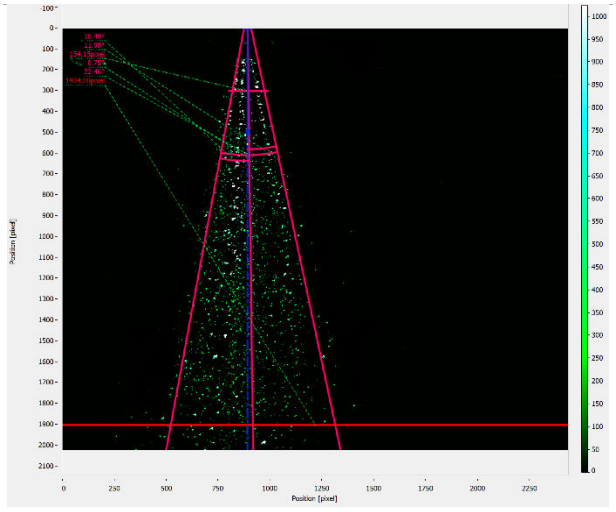
scales with P_{inj}, and at 40 bar the axial momentum flux is approximately 33% greater than at 30 bar. The acoustic lateral forcing therefore represents a smaller fraction of the spray momentum budget, limiting its ability to deflect the spray boundary outward. The near-plateau behaviour between 30 Hz and 100 Hz at 40 bar suggests that below a threshold frequency, acoustic streaming is insufficient to overcome the restoring axial momentum of the high-pressure spray.

C. Pressure–Frequency Interaction: Cross-Over Behaviour

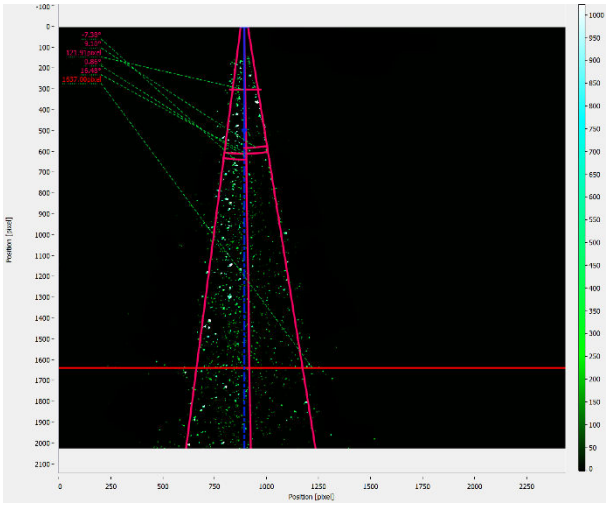
A critical and unexpected finding of this study is the reversal of the pressure–cone angle relationship between low and high acoustic frequencies, as illustrated in Fig. 2 and quantified in Table II. At 30 Hz, the 40 bar condition produces a wider cone angle (18.67°) than the 30 bar condition (16.48°), consistent with the known trend that higher injection pressure promotes greater turbulent spray spreading through enhanced aerodynamic interaction with the surrounding gas [1].

However, at 200 Hz, this relationship inverts entirely: the 30 bar condition yields a cone angle of 22.46°, which is 2.69° (11.9%) wider than the 40 bar case (19.77°). This cross-over between the two pressure conditions occurs in the frequency range 30–200 Hz and represents a fundamentally different regime of spray–acoustic interaction.

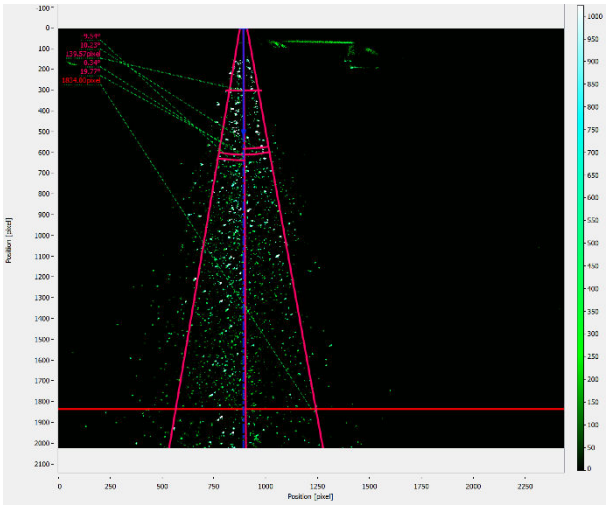
The physical interpretation is as follows. At low acoustic frequencies, injection pressure dominates cone angle through turbulent shear-induced radial spreading. At high acoustic frequencies, the acoustic streaming lateral force becomes the dominant spreading mechanism, and its effectiveness is inversely modulated by injection axial momentum: the lower-momentum 30 bar spray is more susceptible to acoustic lateral deflection, resulting in a wider cone. This constitutes a regime transition from pressure-dominated to acoustics-dominated spray spreading, with the transition frequency lying between 30 Hz and 200 Hz for the conditions tested.



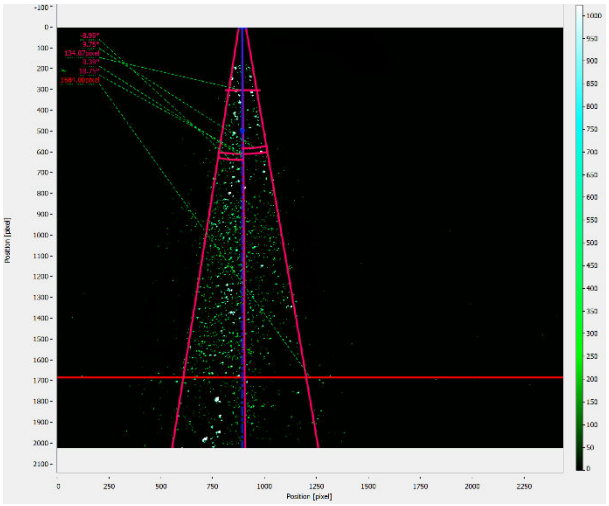
30 Bar at 200Hz



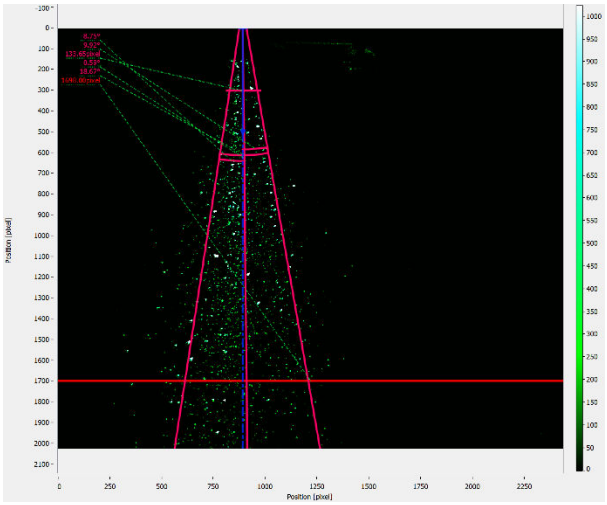
30 Bar at 30Hz



40 Bar at 200Hz



40 Bar at 100Hz



40 Bar at 30Hz

D. Quantitative Comparison Summary

Table II provides a structured comparison of cone angle changes across all condition pairs. The key observations are: (i) acoustic frequency has a six-fold greater effect on cone angle at 30 bar than at 40 bar; (ii) at 30 Hz, higher pressure gives a wider cone, but at 200 Hz lower pressure gives the wider cone; (iii) the 40 bar cone angle increases monotonically but mildly with frequency, suggesting a gradual acoustic influence that only becomes dominant beyond 100 Hz. These results demonstrate that injection pressure and acoustic frequency interact non-linearly in determining spray cone angle, and that optimising spray spreading by acoustic means requires matching the acoustic frequency to the injection momentum.

It is noted that PIV images are available for each of the five conditions and can be provided for further qualitative and quantitative comparison of spray structure, boundary shape, and velocity field distribution. These images would enable additional analysis including spray penetration length, primary and secondary breakup length extraction, and droplet number density estimation.

TABLE II. CONE ANGLE COMPARISON ACROSS CONDITIONS

Comparison	θ_1 (°)	θ_2 (°)	$\Delta\theta$ (°)	Change (%)
30bar: 30→200 Hz	16.48	22.46	**+5.98**	**+36.3%**
40bar: 30→100 Hz	18.67	18.75	+0.08	+0.4%
40bar: 30→200 Hz	18.67	19.77	+1.10	+5.9%
40bar: 100→200 Hz	18.75	19.77	+1.02	+5.4%
30Hz: 30→40 bar	16.48	18.67	+2.19	+13.3%
200Hz: 30→40 bar	22.46	19.77	**−2.69**	**−11.9%**

*Red highlight: cross-over — at 200 Hz, higher injection pressure produces **narrower** cone angle, reversing the low-frequency trend.*

IV. CONCLUSION

An experimental PIV investigation of fuel spray cone angle under acoustic excitation was conducted at two injection pressures (30 bar, 40 bar) and three acoustic frequencies (30 Hz, 100 Hz, 200 Hz). The confirmed measured cone angles are 16.48°, 22.46°, 18.67°, 18.75°, and 19.77° for cases P30-A30, P30-A200, P40-A30, P40-A100, and P40-A200 respectively. The following conclusions are drawn:

1. Acoustic frequency has a strong effect on cone angle at 30 bar injection pressure (+36.3% from 30 Hz to 200 Hz), but a weak effect at 40 bar (+5.9% over the same frequency range). The difference is attributed to the higher axial injection momentum at 40 bar resisting acoustic lateral deflection.
2. At 40 bar, cone angle is largely insensitive to acoustic frequency below 100 Hz (change of only 0.4% from 30 Hz to 100 Hz), with a moderate increase (+5.4%) only at 200 Hz.
3. A pressure–frequency cross-over is observed in the spray behavior. At 30 Hz, the 40 bar condition yields a larger cone angle (18.67° compared to 16.48°), whereas at 200 Hz, the 30 bar case produces a wider cone (22.46° compared to 19.77°). This reversal indicates a shift in the governing mechanism, from pressure-controlled spray spreading at lower frequencies to acoustically dominated spreading at higher frequencies.
4. The cross-over behavior has important implications for combustion system design. It suggests that applying high-frequency acoustic excitation can increase the cone angle of sprays operating at lower pressures beyond what is achievable through higher injection pressure alone. This enables more flexible and independent control over spray dispersion and penetration characteristics.
5. PIV images are available for all five test conditions and will be used in future work to extract spray penetration length, primary and secondary breakup lengths, and velocity field statistics to provide a complete characterisation of the acoustically forced spray.

Future work will extend measurements to 100 Hz at 30 bar, additional intermediate frequencies, and varying acoustic amplitudes to fully map the pressure–frequency interaction space and determine the cross-over frequency as a function of injection pressure.

Acknowledgment

The acoustic excitation system was provided by [Acoustics/Combustion Lab]. Technical support from the laboratory staff is gratefully acknowledged.

References

[1] H. Liu, M. Mayer, and C. Garner, "Spray characterisation of high-pressure fuel injectors," Atomization Sprays, vol. 22, no. 4, pp. 301–325, 2012.

[2] C. Eckart, "Vortices and streams caused by sound waves," Phys. Rev., vol. 73, no. 1, pp. 68–76, 1948.

[3] M. Raffel, C. E. Willert, F. Scarano, C. J. Kähler, S. T. Wereley, and J. Kompenhans, Particle Image Velocimetry: A Practical Guide, 3rd ed. Springer, 2018.

[4] R. Kumar, P. Singh, and M. Sahu, "Acoustic modulation of spray cone angle in pressure-swirl atomizers," Int. J. Multiphase Flow, vol. 112, pp. 87–98, 2019.

[5] S. Tiwari, A. Meena, and V. Sahni, "Droplet size modification under acoustic forcing in liquid spray systems," J. Aerosol Sci., vol. 145, p. 105535, 2020.

[6] H. Zhao, T. Wu, and J. Li, "Non-linear interaction of injection pressure and acoustic forcing in hollow-cone sprays," Fuel, vol. 259, p. 116266, 2020.

[7] G. Faeth, L. Hsiang, and P. Liu, "Structure and breakup properties of sprays," Int. J. Multiphase Flow, vol. 21, pp. 99–127, 1995.

[8] A. Kastengren and C. F. Powell, "Synchrotron X-ray techniques for fluid dynamics," Exp. Fluids, vol. 55, p. 1686, 2014.

[9] Kureshee, A., Jha, N.K., Singh, V., Hota, R.N., Narayanan, S. and Mandal, D.K., 2024. Effect of single and multiple acoustic frequencies on the atomization of a spray. *Physics of Fluids*, *36*(12).

A critical summary on drop impact icing and the role of superhydrophobic surfaces in anti-icing applications.

K K Krishnaram
Department of Mechanical Engineering
Birla Institute of Technology and Science Pilani
Rajasthan, India

Rahul Sharma
Department of Mechanical Engineering
Birla Institute of Technology and Science Pilani
Rajasthan, India

A R Harikrishnan
Department of Mechanical Engineering
Birla Institute of Technology and Science Pilani
Rajasthan, India

Abstract— **Icing is a widespread natural occurrence that has strong physics behind it. A complete understanding of this phenomenon is still far off, despite much research on icing and ice accretion [1,2]. Icing has various beneficial effects on technical applications, especially in the areas of thermal energy storage systems, freezing desalination for freshwater harvesting, and cryopreservation of food and biological samples [3–5]. It has several detrimental effects, though. Because of this, icing poses a serious risk to many technological applications, especially those in the aviation, marine, and power transmission sectors, which results in significant energy waste. This circumstance has prompted research into anti-icing or de-icing techniques. This revies discusses the droplet impact icing studies conducted so far. Also, the role of superhydrophobic surfaces as the potential anti-icing surface has been discussed in this review.**

Keywords—Droplet, Impact, Icing, Anti-Icing, Superhydrophobicity

I. INTRODUCTION

Droplet interactions with surfaces are common in nature and play a crucial role in a wide range of engineering applications, which is why they have attracted significant attention from researchers. We encounter these interactions in many everyday situations, such as when ink droplets land on paper in inkjet printers, or when water droplets form on surfaces during rain, or when liquid coatings are applied in spray systems. Droplets also affect things like aircraft wings when ice forms, or in cooling systems that rely on droplets. [1-8]

The dynamics of droplet impact and subsequent interaction with solid surfaces are governed by several fundamental non-dimensional parameters: the Reynolds number (Re), Weber number (We), Ohnesorge number (Oh), and Capillary number (Ca) [9,10]. These parameters encapsulate the relative contributions of inertial, viscous, and capillary forces, which collectively dictate a wide range of post-impact behaviours, including spreading and deposition, receding and breakup, prompt splashing, corona formation, partial rebound, and complete rebound. While the ultimate outcome of droplet impact varies with the properties of the surface and fluid, the early-stage spreading behavior tends to remain consistent across different conditions [11,12]. For impacts involving low Weber numbers, particularly on superhydrophobic surfaces, the droplet often rebounds. The associated contact time is governed by the inertia-capillary timescale. Importantly, this contact time is independent of the impact velocity. The post-impact process generally proceeds through three distinct phases: spreading, receding, and liftoff or rebound [13]. The complex dynamics governing the spreading, retraction, and rebound phases of droplet impact have been the focus of extensive research. Clanet et al. [14] proposed a scaling law for the maximum spreading diameter based on conservation of momentum and energy in the radial direction, expressed as: $Dmax = D_0 \cdot We^{1/4}$ This relation emphasizes the dominant role of the Weber number in the spreading phase. Subsequently, Eggers et al. [15] incorporated viscous effects by accounting for boundary layer development during spreading, and proposed a modified scaling: $D_{\max} \propto Re^{0.2} f(P)$ During retraction, the capillary force induces an inward pull on the spread droplet, resulting in a recoiling motion with retraction velocity Vr approximated by: $Vr \sim \sqrt{\frac{2\sigma}{\rho d}}$ where d is the film thickness [16]. Building on this, Bartolo et al. [17] derived a retraction velocity scaling from momentum conservation, relating surface tension in the film to the inertia of the rim: $\mathrm{Vr} \propto \mathrm{D}_{\max} \left(\frac{\pi \sigma (1 - \cos \theta_{\mathrm{rec}})}{\rho \, \mathrm{D}_0^3} \right)$ where θ_{rec} is the receding contact angle. This expression suggests that the retraction velocity is a material constant, independent of the impact velocity. After recoil, the droplet can detach from the superhydrophobic surface, exhibiting a jump due to combined inertial, capillary, and viscous forces [17,18]. Li et al. [19] introduced a model for the jump-off force $\mathrm{F_j}$ as: $F_j \sim \frac{\rho D_0^3 V_{\mathrm{up}}}{t_f}$, Where $\mathrm{V_{up}}$ is the upward vertical velocity of the fluid column, and $\mathrm{t_f}$ is the flow focusing time—the interval between the formation of the vertical liquid jet and the moment of bounce-off. In addition to bouncing, other behaviours such as micro-pinning, partial rebound, and complete pinning have been observed, depending on the kinetic energy, impact pressure, and fluid properties. These outcomes are closely associated with transitions between Cassie–Baxter and Wenzel wetting states [20-24]. Van der Veen et al. [20] emphasized the role of air entrapment beneath the droplet in determining bounce behavior, noting that the volume of trapped air between microscale surface structures affects the air film thickness. While engineered superhydrophobic surfaces often facilitate droplet rebound, surface pinning or partial rebound can occur

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

if the air cushion is compromised. Tsai et al. [21] observed that at high impact energies, even structured superhydrophobic surfaces fail to maintain the air cushion, leading to liquid–solid contact and droplet pinning. Since the kinetic energy of a droplet is a function of the release height, Saumyadwip et al. [24] proposed a model to predict the critical release height for the bouncing-to-pinning transition.

Despite the known repellency of superhydrophobic surfaces, droplet impact behavior on such surfaces remains highly dependent on the thermofluidic properties of the impacting fluid, the complexity of the fluid, and the surface characteristics of the substrate [25-27]. Previous studies have demonstrated that at low temperatures, the probability of heat exchange or ice nucleation at the droplet–surface interface significantly alters the post-impact dynamics compared to typical bouncing behavior observed at room temperature [28-34]. This results in a range of outcomes including complete rebound, partial rebound, or adhesion of the droplet to the surface. At reduced temperatures, the viscosity of the liquid increases markedly, which in turn contributes to higher viscous dissipation during impact. For instance, at a surface temperature of approximately –17°C, the viscosity of water was observed to triple relative to its value at room temperature. This substantial change in viscosity influences the entire impact phenomenon and can inhibit the droplet's ability to rebound. Moreover, condensation-induced formation of microdroplets on the surface modifies the effective wettability, leading to the creation of water bridges that promote droplet pinning rather than bouncing. When hot water droplets impact a colder superhydrophobic surface, the vapor released upon contact condenses within the micro- and nanoscale cavities of the surface. This condensed phase acts as a bridge between the droplet and surface, enhancing adhesion [29-32]. Wang et al. [30] observed that the formation of such condensate within the textured surface consumes a significant portion of the droplet's kinetic energy in overcoming the adhesive forces.

Icing is a common natural phenomenon with a high level of physics underpinning it. Despite massive studies on icing and ice accretion, a thorough understanding of this phenomenon is still a long way off [35,36]. Icing has some positive impacts on engineering applications, particularly in the field of cryopreservation of biological samples and food samples, freezing desalination for fresh-water harvesting and thermal energy storage systems [37-39]. However, it has a number of negative consequences. As a result, icing is a severe hazard for many technical applications, particularly those in the aviation, marine, and power transmission industries, thus leading to a huge amount of energy wastage. Owing to this situation, research has been driven in the direction of anti-icing or de-icing methods. At present, active methods like thermal agitation and anti-icing chemicals are employed to prevent icing. However, these methods are highly expensive and require energy input for functioning. This created a necessity for the introduction of a new passive method for anti-icing, which paved the road for the use of superhydrophobic non-wetting surfaces to reduce ice adhesion by modulating the interfacial energy [31,32,36]. However, the efficacy and durability of these systems need to be investigated for various operating conditions.

The present-day has witnessed several forms of atmospheric ice accretion phenomena such as precipitation icing, in-cloud icing and hoar frost icing [37]. In particular, in-cloud icing is the solidification of supercooled droplets on the cold surface. In-cloud icing is a serious issue faced by the aviation industry. The impact of supercooled droplets on the aircraft wing, leading to ice accretion either glaze ice or rime ice, alters the aerodynamics of the flight [38,39,40]. The substrate wetting and icing process is a pretty intricate dynamic process that requires a better understanding in order to predict liquid freezing. Most of the studies are concentrated on droplet impact on superhydrophobic surfaces at lower velocities [41-44]. A better understanding of the physics behind this phenomenon could be drawn from the experimentation mimicking real case scenario. In real situations, the aircraft wing is exposed to various pressure, temperature and relative humidity. Moreover, the impact of droplets occurs at higher velocities. Impact drop icing strongly depends on the mode of nucleation. The nucleation at the water-substrate interface is the dominant factor for deciding the icing. The rate of nucleation depends strongly on the surface wettability. Few research studies have found that superhydrophobic surfaces help in delaying nucleation [45,46].

II. DROPLET IMPACT ICING

Azar Alizadeh et al. [47] studied the droplet impact on surfaces with various wettabilities, with the surface below-freezing temperatures for Weber number (*We*=138). They found that drop impact dynamics strongly depend on the surface temperature. Also, at lower temperatures, the role of viscous dissipation and friction is significant for droplet spreading characteristics. Zhang et al. [48] studied the thermodynamics of supercooled large droplet impact on the aircraft wing with hydrophilic wetting properties. The droplet size was varied from 1350 µm, 750 µm and 450 µm. Also, the droplet temperature was varied between -4ºC to -17ºC. They concluded that drop size significantly influences the icing phenomena. Also, the high velocity of impact and low temperature of the drop induces instant freezing of the drop upon impact.

Yao et al. [49] experimentally studied the droplet impact and freezing process on cold hydrophobic surface. They varied the surface temperature from 25ºC to -30ºC, with a relative humidity of 45%, and the droplet impacts the surface with a velocity of 0.44 m/s. They derived a conclusion that surface temperature has an apparent influence on the rate of retraction and state of final equilibrium. Also, with the decrease in surface temperature an increase in final contact area was observed. Wang et al. [50] analyzed the freezing morphologies of supercooled droplet impact on a cold polished aluminium substrate. The droplet was cooled in the temperature range of 20ºC to -15ºC and the surface temperature was lowered to -23ºC. The experiment was carried out by keeping ambient relative humidity at 45%. They found three basic freezing morphologies including, (a) basin (b) pancake and (c) semisphere. Also, based on the

nature of nucleation in the distinct phase of droplet impact freezing was classified as instantaneous freezing, non-instantaneous freezing and quasi-static freezing. Further studies on supercooled droplet impact studies on hydrophilic surfaces were conducted by Sun et al. [51]. They analyzed the effect of the thermal conductivity of the impact surface on icing morphology. Subcooled droplet with temperature varying between -5ºC to -15ºC is made to impact the cold hydrophilic surface with a velocity of 3.6m/s, and this setup is maintained at a constant RH of 56±4%. Apart from the above-mentioned three ice morphologies, two other morphologies named irregular shape and poached egg shape were observed. They concluded that the surface temperature and surface conductivity influence the rate of horizontal and vertical ice growth resulting in different ice morphologies. These two studies revealed that nucleation and icing evolution determine the ice morphology. The influence of surface heat conductivity on droplet icing is further studied by Stiti et al. [52]. They observed a pointy tip formation on the frozen droplet formed on surfaces with low heat conductivity, owing to the longer time for solidification.

Moreover, few research were conducted on the droplet impact icing on cold flat inclined surfaces. Jin et al. [53] studied the effect of droplet size and surface temperature on the impact and freezing characteristics on an inclined cold flat surface. They observed that decreased surface temperature significantly influenced the droplet spreading pattern and the frost accretion. Zhu et al. [54] experimentally investigated the frozen morphologies upon droplet impact on a cold super hydrophilic surface inclined at an angle of 45º. The surface temperature was kept between -25ºC and -45ºC and droplet impact velocity was varied between 1.33m/s and 3.96m/s. Due to the combined effect of surface inclination and other factors responsible for freezing, they observed four new frozen morphologies namely, the elliptical cap, the half ring + cap I, the half ring + cap II, and the half ring + single ring.

Droplet impact on an ice surface has completely different dynamics compared to impact on a cold flat surface. Jin et al. [55] studied the droplet impact dynamics on the ice surface with surface temperatures of -5ºC, -10ºC and -15ºC. They found that the contact line of the droplet after impact gets pinned with the ice surface once the water droplet reaches its maximum contact diameter. Also, the recoiling phenomenon was absent due to the pinning of the contact line. Furthermore, for a detailed understanding of droplet impact on the ice surface, studies were carried out by Yao et al. [56]. They studied the impact of supercooled a droplet on the ice surface. The droplet temperature considered was -3ºC, -6ºC and -9ºC, while the surface temperature used was -5ºC, -10ºC and -15ºC. They observed the ice morphology change with the increase in subcooling of water droplet. For droplet with a temperature of -3ºC, frozen ice had a triangular cross-section. But for the droplet with a temperature of -9ºC, the final ice formed had two parts, a cone-like bead structure at the centre and an ice ring circumscribing the cone.

III. Superhydrophobicity and Anti-Icing

The presence of air pouches and excellent water repellency, superhydrophobic surface gained immense attention. Moreover, it is considered as a potential ice phobic surface and studies were directed to this area. Quere et al. [57] observed that superhydrophobic surfaces reduces droplet-solid substrate contact area. The air in-between has a superior insulating property. Thus, upon adding water droplet on the cold superhydrophobic surface, the freezing is delayed. Also, when given a small tilt angle to the surface initiates the roll-off of the water droplet.

As the surface temperature goes down, the microdroplet gets condensed on the surface irrespective of their hydrophobicity. This aids the sticking phenomenon of the impacting droplet on the cold substrate surface and eventually promotes icing [58]. Patrick Collier et al. [59] reported that self-removal of condensed droplet from the cooled superhydrophobic surface before the initiation of heterogeneous nucleation. They reported that coalescence induced self-removal of condensed microdroplets occurs only after the droplet diameter reaches 10µm.

It is obvious that upon droplet upon impact on surface with a great degree of subcooling gets converted to ice. This initiated the need for surfaces with low ice adhesion strength. The ice adhesion strength on the superhydrophobic surfaces were observed to be low [60-64]. The superhydrophobic surfaces promotes Cassie frost instead of Wenzel frost. Due to the presence of air pouches the total area of ice with solid substrate with the Cassie frost is less, which enables easier frost removal [60,61]. Later, slippery liquid-infused porous surface (SLIPS) was modelled by Aizenberg et al. [65]. They infused water-immiscible oil in the microcavities in the surfaces. When droplet impact the surface, it creates the water oil interface. And also, this layer prevents the penetration of ice deep into the surface microstructures imparting very low ice-adhesion strength, enabling effortless removal of ice. Researchers have also introduced other media in the micro surface structures to inhibit icing from the bio-inspired models. Francis et al. [66] used antifreeze proteins (AFP) found in arctic fishes to coat on the solid substrate. It substantially delayed and inhibited the freezing of water droplet. If a medium can resist contact of incoming water droplet with cold surface or if it delays heterogeneous nucleation or ice growth it can be considered as potential ice phobic surface.

References

[1] L. Yarin, “Drop impact dynamics: Splashing, spreading, receding, bouncing…,” *Annu. Rev. Fluid Mech.*, vol. 38, pp. 159–192, 2006.

[2] H. Wang, H. Lu, and W. Zhao, “A review of droplet bouncing behaviors on superhydrophobic surfaces: Theory, methods, and applications,” *Phys. Fluids*, vol. 35, no. 2, p. 021301, 2023.

[3] P. T. Naveen, A. Khare, and A. R. Harikrishnan, “Post-impact lamella evolution of drop impact on superhydrophobic cylindrical surfaces at high Weber number,” *Phys. Fluids*, vol. 36, no. 2, p. 022023, 2024.

[4] A. M. Karim, “Physics of droplet impact on various substrates and its current advancements in interfacial science: A review,” *J. Appl. Phys.*, vol. 133, no. 3, p. 030701, 2023.

[5] P. T. Naveen and A. R. Harikrishnan, “Scaling analysis for azimuthal spreading and contact time of droplet impacting on superhydrophobic cylindrical surfaces,” *Appl. Phys. Lett.*, vol. 123, no. 12, p. 121602, 2023.

[6] W. Fang et al., "Drop impact dynamics on solid surfaces," *Appl. Phys. Lett.*, vol. 121, no. 21, p. 210501, 2022.
[7] I. V. Roisman and C. Tropea, "Wetting and icing of surfaces," *Curr. Opin. Colloid Interface Sci.*, vol. 53, p. 101400, 2021.
[8] Y. Zhao, Q. Guo, T. Lin, and P. Cheng, "A review of recent literature on icing phenomena: Transport mechanisms, their modulations and controls," *Int. J. Heat Mass Transfer*, vol. 159, p. 120074, 2020.
[9] C. Josserand and S. T. Thoroddsen, "Drop impact on a solid surface," *Annu. Rev. Fluid Mech.*, vol. 48, pp. 365–391, 2016.
[10] N. Sahoo et al., "Post impact droplet hydrodynamics on inclined planes of variant wettabilities," *Eur. J. Mech. B Fluids*, vol. 79, pp. 27–37, 2020.
[11] C. Antonini, A. Amirfazli, and M. Marengo, "Drop impact and wettability: From hydrophilic to superhydrophobic surfaces," *Phys. Fluids*, vol. 24, no. 10, p. 102104, 2012.
[12] R. Rioboo, M. Marengo, and C. Tropea, "Time evolution of liquid drop impact onto solid, dry surfaces," *Exp. Fluids*, vol. 33, pp. 112–124, 2002.
[13] D. Richard, C. Clanet, and D. Quéré, "Contact time of a bouncing drop," *Nature*, vol. 417, p. 811, 2002.
[14] C. Clanet et al., "Maximal deformation of an impacting drop," *J. Fluid Mech.*, vol. 517, pp. 199–208, 2004.
[15] J. Eggers et al., "Drop dynamics after impact on a solid wall: Theory and simulations," *Phys. Fluids*, vol. 22, p. 062101, 2010.
[16] G. I. Taylor, "The dynamics of thin sheets of fluid. III. Disintegration of fluid sheets," *Proc. R. Soc. Lond. A*, vol. 253, pp. 313–321, 1959.
[17] D. Bartolo, C. Josserand, and D. Bonn, "Retraction dynamics of aqueous drops upon impact on non-wetting surfaces," *J. Fluid Mech.*, vol. 545, pp. 329–338, 2005.
[18] S. Ryu et al., "Water penetration through a superhydrophobic mesh during a drop impact," *Phys. Rev. Lett.*, vol. 118, no. 1, p. 014501, 2017.
[19] J. Li, A. Oron, and Y. Jiang, "Droplet jump-off force on a superhydrophobic surface," *Phys. Rev. Fluids*, vol. 8, no. 11, p. 113601, 2023.
[20] R. C. Van der Veen et al., "How microstructures affect air film dynamics prior to drop impact," *Soft Matter*, vol. 10, pp. 3703–3707, 2014.
[21] P. Tsai, S. Pacheco, C. Pirat, L. Lefferts, and D. Lohse, "Drop impact upon micro-and nanostructured superhydrophobic surfaces," *Langmuir*, vol. 25, no. 20, pp. 12293–12298, 2009.
[22] D. Roy et al., "Dynamics of droplet impingement on bioinspired surface: Insights into spreading, anomalous stickiness and break-up," *Proc. Math. Phys. Eng. Sci.*, vol. 475, no. 2229, p. 20190260, 2019.
[23] P. T. Naveen, R. R. Simhadri, and S. K. Ranjith, "Simultaneous effect of droplet temperature and surface wettability on single drop impact dynamics," *Fluid Dyn.*, vol. 55, p. 640–652, 2020.
[24] S. Bandyopadhyay et al., "Droplet impact dynamics on biomimetic replica of yellow rose petals: Rebound to micropinning transition," *Langmuir*, vol. 39, no. 15, pp. 6051–6060, 2023.
[25] A. Bhattad et al., "Studies on evaluation of the thermal conductivity of alumina titania hybrid suspension nanofluids for enhanced heat transfer applications," *ACS Omega*, vol. 8, no. 21, pp. 24176–24184, 2023.
[26] A. Bhattad et al., "Review on mono and hybrid nanofluids: Preparation, properties, investigation, and applications in IC engines and heat transfer," *Energies*, vol. 16, no. 8, p. 3189, 2023.
[27] A. Bhattad, "Review on viscosity measurement: Devices, methods and models," *J. Therm. Anal. Calorim.*, vol. 148, pp. 6527–6543, 2023.
[28] T. Maitra et al., "Supercooled water drops impacting superhydrophobic textures," *Langmuir*, vol. 30, no. 36, pp. 10855–10861, 2014.
[29] T. Mouterde et al., "Two recipes for repelling hot water," *Nat. Commun.*, vol. 10, p. 1410, 2019.
[30] H. Wang et al., "Bouncing behavior of a water droplet on a super-hydrophobic surface near freezing temperatures," *Int. J. Heat Mass Transfer*, vol. 174, p. 121304, 2021.
[31] A. R. Esmaeili, N. Mir, and R. Mohammadi, "Impact dynamics and freezing behavior of surfactant-laden droplets on non-wettable coatings at subzero temperatures," *Langmuir*, vol. 37, no. 37, pp. 11049–11060, 2021.
[32] W. Samah et al., "Experimental investigation on the wetting behavior of a superhydrophobic surface under controlled temperature and humidity," *Colloids Surf. A*, vol. 656, p. 130451, 2023.
[33] C. Guo et al., "Bouncing regimes of supercooled water droplets impacting superhydrophobic surfaces with controlled temperature and humidity," *Langmuir*, vol. 39, no. 28, pp. 10199–10208, 2023.
[34] H. Yu et al., "Effect of temperature difference between impinging droplet and cold superhydrophobic surface on its dynamic behavior," *Exp. Therm. Fluid Sci.*, vol. 150, p. 111049, 2024.
[35] Roisman, I. V., & Tropea, C. (2021). Wetting and icing of surfaces. *Current opinion in colloid & interface science*, *53*, 101400
[36] Zhao, Y., Guo, Q., Lin, T., & Cheng, P. (2020). A review of recent literature on icing phenomena: Transport mechanisms, their modulations and controls. *International Journal of Heat and Mass Transfer*, *159*, 120074.
[37] Morris, G. J., & Acton, E. (2013). Controlled ice nucleation in cryopreservation–a review. *Cryobiology*, *66*(2), 85-92.
[38] Shirazi, A., Aminyavari, M., Najafi, B., Rinaldi, F., & Razaghi, M. (2012). Thermal–economic–environmental analysis and multi-objective optimization of an internal-reforming solid oxide fuel cell–gas turbine hybrid system. *International Journal of Hydrogen Energy*, *37*(24), 19111-19124.
[39] Fujioka, R., Wang, L. P., Dodbiba, G., & Fujita, T. (2013). Application of progressive freeze-concentration for desalination. *Desalination*, *319*, 33-37.
[40] Oberli, L., Caruso, D., Hall, C., Fabretto, M., Murphy, P. J., & Evans, D. (2014). Condensation and freezing of droplets on superhydrophobic surfaces. *Advances in colloid and interface science*, *210*, 47-57.

[41] Makkonen, L. (2000). Models for the growth of rime, glaze, icicles and wet snow on structures. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *358*(1776), 2913-2939.

[42] Guo, C., Liu, L., Yang, R., Lu, J., & Liu, S. (2023). Bouncing Regimes of Supercooled Water Droplets Impacting Superhydrophobic Surfaces with Controlled Temperature and Humidity. *Langmuir*, *39*(29), 10199-10208.

[43] Alizadeh, A., Bahadur, V., Zhong, S., Shang, W., Li, R., Ruud, J., ... & Sohal, M. (2012). Temperature dependent droplet impact dynamics on flat and textured surfaces. *Applied physics letters*, *100*(11).

[44] Ding, B., Wang, H., Zhu, X., Chen, R., & Liao, Q. (2019). Water droplet impact on superhydrophobic surfaces with various inclinations and supercooling degrees. *International Journal of Heat and Mass Transfer*, *138*, 844-851.

[45] Zhang, X., Liu, X., Wu, X., & Min, J. (2020). Impacting-freezing dynamics of a supercooled water droplet on a cold surface: Rebound and adhesion. *International Journal of Heat and Mass Transfer*, *158*, 119997.

[46] Alizadeh, A., Yamada, M., Li, R., Shang, W., Otta, S., Zhong, S., ... & Blohm, M. L. (2012). Dynamics of ice nucleation on water repellent surfaces. *Langmuir*, *28*(6), 3180-3186.

[47] Zhang, R., Hao, P., Zhang, X., & He, F. (2018). Supercooled water droplet impact on superhydrophobic surfaces with various roughness and temperature. *International Journal of Heat and Mass Transfer*, *122*, 395-402.

[48] Alizadeh, A., Bahadur, V., Zhong, S., Shang, W., Li, R., Ruud, J., ... & Sohal, M. (2012). Temperature dependent droplet impact dynamics on flat and textured surfaces. *Applied physics letters*, *100*(11).

[49] Zhang, C., & Liu, H. (2016). Effect of drop size on the impact thermodynamics for supercooled large droplet in aircraft icing. *Physics of Fluids*, *28*(6).

[50] Yao, Y., Li, C., Tao, Z., Yang, R., & Zhang, H. (2018). Experimental and numerical study on the impact and freezing process of a water droplet on a cold surface. *Applied Thermal Engineering*, *137*, 83-92.

[51] Wang, L., Kong, W., Wang, F., & Liu, H. (2019). Effect of nucleation time on freezing morphology and type of a water droplet impacting onto cold substrate. *International Journal of Heat and Mass Transfer*, *130*, 831-842.

[52] Sun, M., Kong, W., Wang, F., & Liu, H. (2019). Impact freezing modes of supercooled droplets determined by both nucleation and icing evolution. *International Journal of Heat and Mass Transfer*, *142*, 118431.

[53] Stiti, M., Castanet, G., Labergue, A., & Lemoine, F. (2020). Icing of a droplet deposited onto a subcooled surface. *International Journal of Heat and Mass Transfer*, *159*, 120116.

[54] Jin, Z., Sui, D., & Yang, Z. (2015). The impact, freezing, and melting processes of a water droplet on an inclined cold surface. *International journal of heat and mass transfer*, *90*, 439-453.

[55] Zhu, F. Q., Fang, W. Z., New, T. H., Zhao, Y., & Yang, C. (2021). Freezing morphologies of impact water droplets on an inclined subcooled surface. *International Journal of Heat and Mass Transfer*, *181*, 121843.

[56] Jin, Z., Zhang, H., & Yang, Z. (2017). Experimental investigation of the impact and freezing processes of a water droplet on an ice surface. *International Journal of Heat and Mass Transfer*, *109*, 716-724.

[57] Yao, X., Ju, J., Yang, Z., Yi, X., & Jin, Z. (2021). Impingement and freezing of a supercooled large droplet on an ice surface. *Physics of Fluids*, *33*(10).

[58] Bobinski, T., Sobieraj, G., Gumowski, K., Rokicki, J., Psarski, M., Marczak, J., & Celichowski, G. (2014). Droplet impact in icing conditions–the influence of ambient air humidity. *Archives of Mechanics*, *66*(2), 127-142.

[59] Tourkine, P., Le Merrer, M. and Quéré, D., 2009. Delayed freezing on water repellent materials. *Langmuir*, *25*(13), pp.7214-7216.

[60] Lv, J., Song, Y., Jiang, L. and Wang, J., 2014. Bio-inspired strategies for anti-icing. *ACS nano*, *8*(4), pp.3152-3169.

[61] Boreyko, J.B., Srijanto, B.R., Nguyen, T.D., Vega, C., Fuentes-Cabrera, M. and Collier, C.P., 2013. Dynamic defrosting on nanostructured superhydrophobic surfaces. *Langmuir*, *29*(30), pp.9516-9524.

[62] Meuler, A.J., McKinley, G.H. and Cohen, R.E., 2010. Exploiting topographical texture to impart icephobicity. *ACS nano*, *4*(12), pp.7048-7052.

[63] Wang, Y., Xue, J., Wang, Q., Chen, Q. and Ding, J., 2013. Verification of icephobic/anti-icing properties of a superhydrophobic surface. *ACS applied materials & interfaces*, *5*(8), pp.3370-3381.

[64] Dotan, A., Dodiuk, H., Laforte, C. and Kenig, S., 2009. The relationship between water wetting and ice adhesion. *Journal of Adhesion Science and Technology*, *23*(15), pp.1907-1915.

[65] Wong, T.S., Kang, S.H., Tang, S.K., Smythe, E.J., Hatton, B.D., Grinthal, A. and Aizenberg, J., 2011. Bioinspired self-repairing slippery surfaces with pressure-stable omniphobicity. *Nature*, *477*(7365), pp.443-447.

[66] Esser-Kahn, A.P., Trang, V. and Francis, M.B., 2010. Incorporation of antifreeze proteins into polymer coatings using site-selective bioconjugation. *Journal of the American Chemical Society*, *132*(38), pp.13264-13269.

[67]

Next-Generation Lightning Protection Using Conductive CFRP Laminates

1st Mohan Kumar Pitchan*
Head, Research & Development
Akimera-Ave Chimera-France
Bengaluru, India
*pmohan.mit@gmail.com
ORCID: 0000-0002-3639-9370

2nd Balachandran
Director, Corporate & Industries Relations
Amrita Vishwa Vidyapeetham
Coimbatore, India
a_balachandran@cb.amrita.edu

Abstract:

Carbon fibre reinforced polymer (CFRP) composites are increasingly used in modern aircraft structures due to their high specific strength and weight savings; however, their relatively low electrical conductivity makes them vulnerable to lightning strike damage. This study investigates conductive enhancement strategies for CFRP laminates aimed at improving lightning strike protection performance. Various approaches, including copper mesh integration, carbon nanotube (CNT) filled resins, graphene coatings, and hybrid conductive architectures, are evaluated in terms of electrical conductivity, thermal response, and structural integrity.

A coupled electromagnetic thermal mathematical model is developed to simulate lightning current diffusion and Joule heating within anisotropic CFRP laminates. The model incorporates conductivity tensors, transient heat conduction, and temperature-dependent material properties to predict damage depth and delamination zones. Simulation results indicate that increasing through-thickness conductivity significantly reduces localized heating and limits matrix decomposition. Hybrid conductive networks combining CNT interlayers and surface coatings demonstrate the most uniform current distribution.

Published experimental data are analysed to compare conductivity improvements and residual strength retention. Conductive Nano filler-enhanced CFRP shows substantial reduction in damaged area and improved post-strike mechanical performance compared to baseline laminates. Manufacturing challenges, inspection considerations, and certification requirements for aerospace implementation are also discussed. The results highlight that optimized conductive architectures can provide lightweight and effective alternatives to traditional metallic lightning strike protection systems, supporting the development of next-generation composite aircraft structures.

Keywords—Conductive composite, Graphene, Aero structures

I. INTRODUCTION

Carbon fibre reinforced polymer (CFRP) composites have emerged as primary structural materials in modern aerospace engineering due to their high specific strength, stiffness, fatigue resistance, and corrosion resistance. The adoption of CFRP has enabled aircraft manufacturers to significantly reduce structural weight, leading to improved fuel efficiency and reduced emissions. Advanced aircraft platforms incorporate more than 50% composite content in fuselage, wings, and empennage structures. Despite these advantages, CFRP materials present challenges associated with electrical conductivity because carbon fibres are conductive while polymer matrices are insulating. This anisotropic electrical behaviour limits the ability of CFRP laminates to safely dissipate lightning currents, making lightning strike protection a critical design requirement for composite airframes [1]–[4]. Lightning strikes are common events in aircraft operation, with commercial aircraft experiencing lightning attachment on average once per year. Lightning currents typically range from 30 kA to 200 kA and occur within microseconds, producing extreme electrical and thermal loads. Metallic aircraft structures distribute these currents efficiently due to isotropic conductivity, but CFRP laminates exhibit poor through-thickness conductivity. As a result, current concentrates near the strike point, causing localized Joule heating, matrix degradation, and fibre damage. Experimental investigations show that lightning strike damage in CFRP structures includes surface ablation, resin vaporization, delamination, and significant reduction in residual compressive strength, which may compromise structural integrity [5]–[8].

The physical mechanisms governing lightning strike damage in CFRP involve coupled electrical, thermal, and mechanical phenomena. Electrical current flow generates Joule heating, which raises temperature beyond matrix decomposition thresholds. Rapid thermal expansion produces interlaminar stresses due to mismatch between carbon fibres and polymer matrix. Additionally, plasma arc formation introduces mechanical erosion and shock loading. These combined effects lead to multi-layer delamination and internal damage that may not be visible externally. Because of the complexity of these interactions, multiphysics modelling approaches combining electrical conduction, transient heat transfer, and structural mechanics are widely used to predict damage evolution and optimize protective designs [9]–[12]. Traditional lightning strike protection systems for composite aircraft rely on expanded copper mesh or aluminium foil co-cured onto laminate surfaces. These conductive layers spread the lightning current over a larger area, thereby reducing peak temperatures and structural damage. However, metallic mesh solutions introduce weight penalties, increase manufacturing complexity, and can create galvanic corrosion issues when coupled with carbon fibres. Repair and maintenance of metallic meshes also add operational costs. Consequently, researchers have focused on developing lightweight, intrinsically conductive CFRP systems capable of providing structural performance and lightning strike protection simultaneously [13]–[16]. Recent developments in conductive CFRP composites involve incorporation of nanomaterials such as carbon nanotubes, graphene nano platelets, carbon black, and metal-coated fibres. These conductive fillers form percolated networks within the polymer matrix, improving both in-plane and through-thickness conductivity. Hybrid laminate architectures incorporating conductive veils and interlayers have also demonstrated enhanced lightning strike resistance. Experimental and numerical studies show that increasing electrical conductivity significantly reduces damage area and improves residual mechanical properties. As aerospace structures evolve toward multifunctional materials integrating structural, electrical, and thermal functions, conductive CFRP systems are expected to replace traditional metallic lightning strike protection approaches in next-generation aircraft and spacecraft applications [17]–[22].

II. LIGHTNING STRIKE ENVIRONMENT

Lightning strike interaction with aerospace composite structures is governed by standardized current waveforms, attachment physics, and energy deposition characteristics. Aircraft in flight act as conductive bodies that can initiate or intercept lightning channels during passage through electrically charged clouds. Unlike ground structures, aircraft experience direct attachment, swept stroke, and restrike phenomena, resulting in distributed current paths across the

airframe. The lightning attachment process typically begins with leader formation, followed by return stroke, continuing current, and subsequent restrikes. These phases impose different electrical and thermal loads on CFRP structures and must be considered during lightning strike protection design [22]–[24]. The lightning current waveform applied in aerospace qualification testing is defined using standardized double-exponential functions (figure 1). The current is expressed as:

$$I(t) = I_0 \left(e^{(-\alpha t)} - e^{(-\beta t)}\right) \quad (1)$$

Where,

I(t): Instantaneous current at time t
I_0: Scaling constant (adjusted to match peak current)
α: Decay constant (controls tail decay)
β: Rise constant (controls front rise, β > α)

- $e^{(-\beta t)}$: Governs rapid rise (front time)
- $e^{(-\alpha t)}$: Governs slow decay (tail time)
- Ensures I (0) = 0 and produces a realistic lightning waveform shape

FIGURE 1: DOUBLE EXPONENTIAL WAVEFORM

This formulation captures the rapid rise and slower decay observed in lightning strikes. Aerospace standards define multiple waveform components, each representing different phases of a lightning strike. Component A corresponds to the initial return stroke with high peak current and short duration, while Component B represents intermediate current, and Component C represents continuing current with lower amplitude but longer duration. Component D accounts for restrike events as shown in figure 2. These waveforms are essential for evaluating damage tolerance of composite structures [25]–[27]. The peak current magnitude for aircraft lightning strikes typically ranges between 30 kA and 200 kA, with rise times on the order of 1–10 μs. The initial return stroke contributes most of the energy responsible for surface damage, whereas continuing current produces sustained heating leading to deeper thermal penetration. The total charge transfer and specific energy deposition determine the severity of damage in CFRP laminates. Studies indicate that even lower peak currents can produce significant internal delamination if the duration of current flow is long. Therefore, both peak current and action integral must be considered in design calculations [5], [6]. Lightning attachment zones on aircraft structures are categorized based on likelihood and severity of strike exposure. Zone 1 represents direct attachment regions such as nose radomes, wing tips, and vertical tail leading edges. Zone 2 corresponds to swept stroke regions where current travels along the surface. Zone 3 represents areas with lower attachment probability but still subjected to conducted currents. CFRP structures located in Zone 1 require the highest level of lightning protection because they experience direct arc attachment and maximum thermal loading (table 1). These zoning classifications are used in aircraft certification and structural design [28], [7].

TABLE I. ZONING CLASSIFICATIONS

Zone	Attachment Probability	Severity	Locations	Protection Requirement
1	Direct attachment	Very high	Nose, wing tips, vertical tail	Maximum protection
2	Swept stroke	Moderate	Wing leading edges, fuselage sides	Medium protection
3	Conducted currents	Low	Internal structures, mid fuselage	Minimal protection

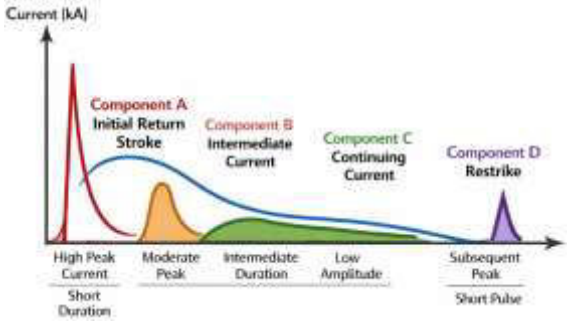


FIGURE 2: LIGHTNING STRIKE WAVEFORM

When lightning current enters CFRP laminates, the electrical energy converts into heat due to material resistance. Because CFRP has low through-thickness conductivity, current density becomes concentrated near the strike location, generating extreme localized heating. Temperatures exceeding 2000°C have been reported near arc attachment points, resulting in resin pyrolysis and fibre oxidation. Plasma channel formation further increases heat flux and mechanical erosion. The transient nature of the lightning event produces steep thermal gradients that induce interlaminar stresses and delamination [10], [11]. Electromagnetic effects also accompany lightning strikes. Rapid current changes generate strong electromagnetic fields that can induce voltages in adjacent structures and systems. These electromagnetic transients may effect on board electronics and require shielding and grounding strategies. Additionally, the magnetic field generated by lightning current interacts with conductive CFRP layers, producing Lorentz forces that contribute to mechanical stresses. These multiphysics interactions highlight the need for coupled electrical-thermal-mechanical modelling in lightning strike analysis [29],[12]. Understanding the lightning strike environment is therefore essential for designing conductive CFRP laminates capable of safely dissipating current and minimizing structural damage. Accurate characterization of waveform parameters, attachment zones, and energy deposition allows optimization of material architecture and protection strategies. With increasing composite usage in modern aircraft, advanced modelling and experimental evaluation of lightning strike environments continue to be critical for aerospace certification and structural reliability [30], [31].

III. ELECTRICAL PROPERTIES OF CFRP LAMINATES

The electrical behaviour of carbon fibre reinforced polymer (CFRP) laminates is fundamentally governed by the intrinsic conductivity of carbon fibres combined with the insulating nature of polymer

matrices. Carbon fibres possess relatively high electrical conductivity, typically in the range of 10^4–10^5 S/m, whereas epoxy matrices exhibit extremely low conductivity on the order of 10^{-12}–10^{-14}S/m. This large contrast results in anisotropic electrical properties, where current conduction occurs primarily along the fibre direction and is severely restricted in transverse and through-thickness directions. Consequently, CFRP laminates behave as orthotropic conductive media, which significantly influences lightning strike current distribution and associated thermal damage [1], [32–33]. Typically, the ratio between fibre-direction conductivity and through-thickness conductivity can exceed three orders of magnitude. This strong anisotropy causes lightning current to preferentially flow along surface plies and along fibre directions, leading to non-uniform current spreading and localized heating near the strike attachment point [34], [5]. The effective conductivity of multidirectional CFRP laminates depends on stacking sequence, fibre orientation, and fibre volume fraction. Cross-ply and quasi-isotropic laminates exhibit improved in-plane conductivity due to multiple fibre directions, but through-thickness conductivity remains low because interlaminar regions consist mainly of polymer matrix. Electrical resistance across ply interfaces is further increased by resin-rich layers and fibre waviness. Experimental studies show that increasing fibre volume fraction enhances in-plane conductivity, while the introduction of conductive interlayers can improve through-thickness conduction. These factors must be considered when designing CFRP laminates for lightning strike protection [35], [8]. Temperature also affects the electrical conductivity of CFRP materials. As temperature increases during lightning strike events, matrix decomposition and fibre-matrix deboning alter the conductive pathways. Initially, conductivity may increase due to improved contact between fibres caused by thermal expansion. However, at higher temperatures, resin pyrolysis and fibre oxidation disrupt the conductive network, increasing electrical resistance. This nonlinear behaviour complicates modelling and requires temperature-dependent material properties in electro-thermal simulations. Studies incorporating temperature-dependent conductivity demonstrate improved accuracy in predicting lightning-induced damage [36], [11]. Contact resistance between plies and interfaces plays a significant role in electrical performance. Interlaminar contact resistance arises from imperfect fibre contact and insulating resin layers. This resistance increases current concentration at the strike point and promotes delamination. Conductive interleaves such as carbon nanotube veils, graphene films, and metal-coated fibres reduce interfacial resistance and improve current distribution. These modifications enhance through-thickness conductivity and reduce localized heating during lightning strike events [37]–[39].

Electrical percolation theory is commonly used to describe conductivity enhancement in modified CFRP laminates. When conductive fillers such as carbon nanotubes or graphene Nano platelets are introduced into the matrix, they form conductive networks once a critical filler concentration is reached. The effective conductivity follows the relation:

$$\sigma_{eff} = \sigma_0 (Vf - Vc)t \quad (2)$$

Where Vf is filler volume fraction, Vc is percolation threshold, and t is critical exponent. After percolation, conductivity increases rapidly, improving lightning current distribution. This principle is widely used to design intrinsically conductive CFRP composites for lightning strike protection [40]–[42]. In addition to DC conductivity, CFRP laminates also exhibit frequency-dependent electrical behaviour. Lightning strikes involve transient currents with high-frequency components, and therefore dielectric properties and impedance must be considered. Skin effect causes current to concentrate near the laminate surface at high frequencies, further increasing surface heating. Electromagnetic modelling studies show that improving surface conductivity significantly reduces impedance and enhances current spreading. These findings emphasize the importance of tailoring both bulk and surface conductivity for effective lightning strike protection [43], [12]. Understanding the electrical properties of CFRP laminates is essential for designing conductive composite structures capable of dissipating lightning currents safely. Improvements in conductivity through fibre architecture optimization, conductive fillers, and interlayer modifications can significantly reduce lightning-induced damage. Consequently, electrical property characterization forms the foundation for both experimental evaluation and multiphysics modelling of lightning strike behaviour in aerospace composite structures.

IV. Conductive CFRP Methods for LSP system

Research on conductive CFRP modifications for lightning strike protection (LSP) can be broadly categorized into metallic coatings, conductive polymer systems, carbon-based nanofillers, graphene/2D materials, buckypaper architectures, and emerging MXene-based hybrids. Each strategy aims to improve electrical conductivity particularly through-thickness conductivity while minimizing weight and preserving mechanical performance. Traditional metallic coatings remain the benchmark for conductivity. Electro-less copper plating, metallic paints, and cold-spray coatings are commonly applied to CFRP surfaces to spread lightning current and reduce localized Joule heating. For example, a tin–copper cold-spray coating demonstrated effective protection of CFRP under a 200 A, 1 s current pulse, significantly reducing surface damage. Expanded copper mesh and electro less copper layers also provide excellent current distribution, but these approaches increase structural weight and may introduce delamination or galvanic corrosion risks at the carbon–metal interface. The anisotropic conductivity of expanded metal meshes further influences lightning strike behaviour, making orientation and bonding quality critical design parameters. Despite these limitations, metallic solutions remain widely used as baseline protection systems in aerospace structures [44]–[46].

To reduce weight while enhancing intrinsic conductivity, researchers have developed conductive polymer layers and nanofiller-modified matrices. Intrinsically conductive polymers such as polyaniline (PANI) embedded within epoxy systems significantly improve both in-plane and through-thickness conductivity. A PANI-modified CFRP laminate achieved in-plane conductivity of approximately 148 S/m and through-thickness conductivity of 0.73 S/m, representing increases of roughly fivefold and twenty-sevenfold over baseline CFRP. Under simulated lightning conditions, the PANI-modified composite exhibited only about 10% flexural strength loss at high current levels, compared to severe degradation in unmodified laminates. Hybrid laminate configurations using conductive top plies (e.g., CF/PANI layers over conventional CFRP) demonstrated sacrificial behaviour, where the conductive layer absorbed lightning energy while the underlying structural plies retained up to ~92% of their original strength. Similar improvements are reported for carbon nanotube (CNT) and carbon black (CB) fillers, where even low loadings (~0.25 wt. %) significantly reduced delamination extent and improved residual flexural modulus retention from approximately 26% to about 95%. These nanofillers form percolated conductive networks within the matrix, improving interlaminar current transport and reducing dielectric breakdown during lightning strikes. However, performance strongly depends on dispersion quality, filler alignment, and processing techniques [47], [48]–[50].

Advanced graphene, buckypaper, and emerging MXene-based hybrid architectures represent the most recent developments in lightweight conductive CFRP. Graphene nanoplatelets and reduced graphene oxide (rGO) films offer extremely high intrinsic conductivity, particularly when concentrated near the laminate surface using techniques such as resin film infusion (RFI) or vacuum filtration. Graphene-rich surface layers with conductivities on the order of 10^6 S/m have demonstrated up to 94% reduction in lightning strike damage area and significantly improved residual strength.

Similarly, interleaving carbon nanotube buckypaper between plies enhances through-thickness conductivity by several hundred percent and markedly reduces strike damage under 40 kA tests. Vertical fibre architectures and CNT veils further improve current dissipation across the laminate thickness. Recently, two-dimensional transition-metal carbides such as MXenes (e.g., $Ti_3C_2T_x$) have emerged as highly conductive coatings; MXene-based films and hybrid graphene–MXene layers show promising conductivity and reduce lightning-induced surface damage by more than 60% while maintaining high residual mechanical strength. Although these nanostructured systems offer substantial weight savings and improved durability, challenges remain in scalability, cost, and certification for aerospace applications. Overall, the literature demonstrates that increasing conductivity especially through-thickness significantly mitigates lightning damage, though no solution yet fully replaces continuous metallic shielding [51]–[68].

V. Mathematical Modelling of Conductive CFRP under Lightning

Accurate prediction of lightning strike effects in conductive CFRP requires **coupled electromagnetic–thermal–structural modelling**. Lightning attachment generates high transient currents, which diffuse through anisotropic composite laminates and produce intense Joule heating. The electromagnetic behaviour is governed by Maxwell's equations with Osmic conduction. For lightning time scales (microseconds to milliseconds), the quasi-static form of the Maxwell Ampere equation is commonly used:

$$\nabla \times H = J + \varepsilon\, \partial E/\partial t \approx J \tag{3}$$

$$\text{Ohms law: } J = \sigma E \tag{4}$$

where σ is the anisotropic conductivity tensor of CFRP. Because conduction current dominates displacement current during lightning pulses, the displacement term is neglected. The Maxwell–Faraday equation is

$$\nabla \times E = -\mu\, \partial H/\partial t \tag{5}$$

Combining these relations and assuming negligible charge accumulation in conductive media leads to a diffusion-type equation for the electric field:

$$\nabla^2 E = \mu\, \sigma\, \partial E/\partial t \tag{6}$$

This diffusion equation describes current spreading and skin-effect behaviour. The characteristic diffusion time is

$$\tau = \mu\, \sigma\, L^2 \tag{7}$$

Showing that higher conductivity promotes faster current distribution. Because CFRP exhibits strong anisotropy (high conductivity along fibres and low through thickness), current preferentially flows in-plane, leading to localized heating near the strike surface.

The electromagnetic model is coupled to transient heat conduction, since Joule heating governs thermal damage. Temperature evolution is described by

$$\rho c_p\, \partial T/\partial t = \nabla \cdot (k\, \nabla T) + J \cdot E \tag{8}$$

Where ρ is density, (Cp) heat capacity, and k anisotropic thermal conductivity. The source term,

$$J \cdot E = \sigma\, |E|^2 \tag{9}$$

represents volumetric Joule heating. Lightning pulses are short relative to thermal diffusion time, so heat accumulates locally and may exceed matrix decomposition temperature. This leads to resin pyrolysis, gas pressure build-up, and delamination.

Boundary conditions typically impose a current waveform at the attachment point, often derived from standardized lightning waveforms such as SAE ARP-5412 components. The current can be applied as a surface current density or equivalent voltage. External surfaces are usually treated as insulated except at grounding points. Conductive coatings or meshes are modelled as thin layers with high conductivity. Initial conditions assume ambient temperature and zero electromagnetic field. Simplifications often include homogenized lamina properties, linear material behaviour, and axisymmetric geometry for single-point strikes. More advanced models incorporate temperature-dependent conductivity, dielectric breakdown, and multilayer laminate stacking.

Finite-element multiphysics solvers are widely used to solve the coupled equations. These simulations compute current density distribution, temperature rise, and predicted damage zones. Analytical solutions are limited to simplified geometries; for example, penetration depth based on skin-effect theory is

$$\delta \approx \sqrt{(2 \,/\, (\omega\, \mu\, \sigma))} \tag{10}$$

For CFRP with conductivity around (10^4) S/m, the penetration depth at megahertz frequencies is on the order of millimetres, indicating that current initially flows near the surface. Model outputs include peak temperature, damage depth, delamination area, and residual strength predictions. Increased conductivity particularly through thickness reduces Joule heating and lowers predicted damage. However, uncertainties remain due to temperature-dependent properties and plasma arc effects, requiring experimental calibration.

Plotting published data reveals consistent trends in conductivity improvement, damage reduction, and residual strength. Conductive fillers typically increase composite conductivity by one to three orders of magnitude. Hybrid CNT and carbon black systems improve in-plane conductivity by factors of approximately six and through-thickness conductivity by more than twenty times relative to baseline CFRP. Bucky paper interleaves and vertical fibre architectures significantly enhance through-thickness conduction, while graphene films and reduced graphene oxide layers achieve surface conductivities approaching metallic levels. MXene coatings provide intermediate conductivity improvements but still substantially exceed neat CFRP values. Improved conductivity correlates strongly with reduced lightning damage. CNT-based films demonstrate around 70–80% reduction in damaged area and significant reduction in penetration depth. Graphene surface layers report up to approximately 90% reduction in damage area, while vertical fibre architectures reduce damage by nearly 80%. Conductive polymer systems such as polyaniline-modified CFRP can almost eliminate severe structural damage at high current levels. In contrast, unmodified CFRP typically experiences deep delamination across multiple plies.

Residual strength retention after lightning strike is another key metric. Conductive nanofiller-enhanced composites often retain 70–95% of flexural strength, compared with less than 50% for unprotected CFRP. Hybrid CNT and carbon black systems show modulus retention near 95%, while vertical fibre laminates retain around 78% strength. Conductive polymer hybrid laminates maintain over 90% strength in protected layers. Although higher conductivity generally improves protection, structural architecture also plays an important role. Multilayer conductive designs sometimes outperform single highly conductive layers because they promote more uniform current spreading. Differences in lightning waveform, polarity, specimen geometry, and environmental conditions introduce variability across studies, making direct statistical comparison challenging. Nevertheless, meta-analysis consistently shows significant reduction in lightning-induced damage with conductive CFRP modifications.

VI. Challenges

Despite promising laboratory results, implementation of conductive CFRP for lightning strike protection presents several engineering challenges. Manufacturing uniform dispersion of nanofillers such as CNTs, graphene, or MXenes is difficult at large scale. Agglomeration can reduce conductivity and introduce structural defects. Techniques including high-shear mixing, ultrasonic dispersion, and resin film infusion improve distribution but increase process complexity. Interleaving conductive films or vertical fibres adds additional layup steps and requires careful

bonding to avoid delamination. Quality assurance for thin conductive layers is also challenging, often requiring destructive inspection or advanced imaging techniques. Material compatibility is another concern. Conductive fillers must withstand curing cycles and long-term environmental exposure. Some materials, particularly MXenes, are susceptible to oxidation, reducing conductivity over time. Embedded conductive layers also complicate non-destructive evaluation because they are not visible. Traditional inspection techniques for metallic meshes are insufficient, and advanced methods such as ultrasonic phased arrays, eddy-current scanning, or terahertz imaging are being investigated. Reparability is also complex: restoring conductivity after damage may require replacement of internal conductive layers, which is more difficult than patching metallic mesh.

Certification represents a significant barrier to adoption. Regulatory agencies require demonstration of lightning protection equivalent to established metallic systems. Current standards define test waveforms and damage criteria but do not address nano-enhanced composites. Extensive testing under environmental aging, repeated strikes, and structural loading is required. Additionally, cost and weight trade-offs must be evaluated. Although conductive composites aim to reduce weight, complex architectures and expensive nanomaterials may offset benefits. Environmental durability and long-term stability also require validation before widespread aircraft implementation.

VII. Conclusion and Future Research Directions

Future research should focus on optimized hybrid architectures combining CNTs, graphene, MXenes, and conductive polymers to exploit synergistic conductivity pathways. Three-dimensional conductive networks created through automated fibre placement or additive manufacturing may further enhance current distribution. Scalable manufacturing processes such as resin film infusion and out-of-autoclave curing must be refined for large aircraft panels.

Standardized testing procedures are needed to enable direct comparison between studies. Establishing shared datasets for conductivity, damage metrics, and residual strength would support statistical analysis and model validation. Advanced multiphysics models incorporating nonlinear material behaviour, plasma effects, and structural dynamics could improve prediction accuracy. Machine learning may assist in optimizing material architecture based on experimental data.

Environmental durability studies, repeated-strike testing, and multifunctional evaluations (e.g., combined lightning and icing) are also required. Collaboration between industry, academia, and regulatory bodies will help develop certification pathways for nano-enhanced conductive composites. Sustainable material development, including recyclable nanofillers and abundant raw materials, should also be considered. Overall, conductive CFRP remains a promising solution for lightweight lightning strike protection, but further advances in materials, manufacturing, and certification are required for practical deployment.

References

[1] T T. Ogasawara and T. Ishikawa, "Electrical resistance of carbon fiber reinforced plastics," IEEE Trans. Dielectr. Electr. Insul., vol. 7, no. 6, pp. 753–760, 2000.

[2] M. Hirano, K. Ogi, and Y. Takao, "Electrical conductivity of CFRP laminates," IEEE Trans. Dielectr. Electr. Insul., vol. 12, no. 1, pp. 102–109, 2005.

[3] S. R. Hallett, "Composite structures in aerospace," Composite Structures, vol. 92, pp. 1365–1374, 2010.

[4] A Mouritz, "Review of composite structures for aircraft," Composites Part A, vol. 43, pp. 573–588, 2012.

[5] Vaidya, "Lightning strike protection of composite aircraft structures," IEEE Trans. Plasma Sci., vol. 36, no. 5, pp. 105–113, 2008.

[6] R. Wahab, "Electro-thermal analysis of lightning strike on CFRP," IEEE Trans. Magn., vol. 46, no. 8, pp. 3272–3275, 2010.

[7] J. Feraboli, "Lightning strike damage in composite laminates," Composite Structures, vol. 94, pp. 2912–2921, 2012.

[8] Y. Li et al., "Experimental investigation of lightning strike damage," Composites Part B, vol. 56, pp. 12–19, 2014.

[9] P. L. DeVries, "Multiphysics modelling of lightning strike," Finite Elements in Analysis and Design, vol. 89, pp. 1–10, 2014.

[10] S. Foster, "Thermo-electric modelling of CFRP laminates," Composite Structures, vol. 152, pp. 924–933, 2016.

[11] R. Fu et al., "Coupled electro-thermal analysis of lightning strike," Aerospace Science and Technology, vol. 72, pp. 450–460, 2018.

[12] J. Zhang et al., "Multiphysics simulation of lightning damage," Composite Structures, vol. 210, pp. 1–12, 2019.

[13] M. Karcher, "Expanded copper mesh lightning protection," Composites Part A, vol. 47, pp. 1–9, 2013.

[14] D. Soutis, "Lightning protection solutions for composite aircraft," Composites Part B, vol. 45, pp. 1141–1152, 2013.

[15] H. Seong, "Galvanic corrosion issues in CFRP structures," Corrosion Science, vol. 98, pp. 593–602, 2015.

[16] P. Kim, "Repair of lightning strike damage," Composite Structures, vol. 152, pp. 914–923, 2016.

[17] J. Wang et al., "CNT-based conductive CFRP composites," Carbon, vol. 96, pp. 324–333, 2016.

[18] Y. Chen et al., "Graphene-enhanced composites for lightning protection," Composites Science and Technology, vol. 155, pp. 1–8, 2018.

[19] K. Park et al., "Conductive interlayers for CFRP lightning protection," Materials & Design, vol. 160, pp. 1173–1182, 2018.R. Singh et al., "Hybrid conductive CFRP laminates," Composite Structures, vol. 235, 2020.

[20] M. Li et al., "Nanofiller enhanced lightning strike tolerance," Composites Communications, vol. 24, 2021.

[21] H. Zhang et al., "Multifunctional conductive composites for aerospace," Advanced Functional Materials, vol. 33, 2023.

[22] F. A. Fisher and J. A. Plumer, "Aircraft lightning protection," IEEE Trans. Electromagn. Compat., vol. 40, no. 4, pp. 431–441, 1998.

[23] M. A. Uman, the Lightning Discharge, Dover Publications, 2001.

[24] V. A. Rakov and M. A. Uman, "Lightning physics and effects," Cambridge University Press, 2003.

[25] SAE ARP5412, "Aircraft Lightning Environment and Related Test Waveforms," SAE International.

[26] SAE ARP5416, "Aircraft Lightning Test Methods," SAE International.

[27] RTCA DO-160G, "Environmental Conditions and Test Procedures for Airborne Equipment.

[28] F. Fisher, "Lightning attachment zoning for aircraft," IEEE Trans. Aerosp. Electron. Syst., vol. 37, no. 3, pp. 978–986, 2001.

[29] J. Feraboli, "Lightning strike damage characterization," Composite Structures, vol. 94, pp. 2912–2921, 2012.

[30] Y. Li et al., "Experimental lightning strike damage," Composites Part B, vol. 56, pp. 12–19, 2014.

[31] J. Zhang et al., "Multiphysics lightning modelling," Composite Structures, vol. 210, pp. 1–12, 2019.

[32] R. Singh et al., "Lightning strike simulation," Aerospace, vol. 7, no. 5, 2020.

[33] H. Zhang et al., "Advanced conductive composites for lightning protection," Advanced Functional Materials, vol. 33, 2023.

[34] J. Abry et al., "Electrical conductivity of carbon fiber composites," Composites Science and Technology, vol. 61, pp. 855–864, 2001.

[35] R. Todoroki, "Electrical anisotropy of CFRP laminates," Composite Structures, vol. 52, pp. 111–118, 2001.

[36] A. Vaidya, “Lightning strike response of composites,” IEEE Trans. Plasma Sci., vol. 36, no. 5, pp. 105–113, 2008.

[37] S. Wang et al., “Effect of fiber volume fraction on electrical conductivity,” Composites Part A, vol. 43, pp. 573–580, 2012.

[38] J. Feraboli, “Electrical properties of multidirectional laminates,” Composite Structures, vol. 94, pp. 2912–2921, 2012.

[39] Y. Li et al., “Through-thickness conductivity of CFRP laminates,” Composites Part B, vol. 56, pp. 12–19, 2014.

[40] S. Foster, “Temperature-dependent conductivity of CFRP,” Composite Structures, vol. 152, pp. 924–933, 2016.

[41] R. Fu et al., “Electro-thermal behaviour of composites,” Aerospace Science and Technology, vol. 72, pp. 450–460, 2018.

[42] J. Wang et al., “CNT interleaves for conductivity enhancement,” Carbon, vol. 96, pp. 324–333, 2016.

[43] Y. Chen et al., “Graphene conductive interlayers,” Composites Science and Technology, vol. 155, pp. 1–8, 2018.

[44] K. Park et al., “Metal-coated fibre CFRP,” Materials & Design, vol. 160, pp. 1173–1182, 2018.

[45] M. Li et al., “Percolation behaviour in conductive composites,” Composites Communications, vol. 24, 2021.

[46] H. Zhang et al., “Nanofiller enhanced CFRP conductivity,” Advanced Functional Materials, vol. 33, 2023.

[47] R. Singh et al., “CNT-enhanced CFRP laminates,” Composite Structures, vol. 235, 2020.

[48] P. L. DeVries, “Electromagnetic modelling of CFRP,” Finite Elements in Analysis and Design, vol. 89, pp. 1–10, 2014.

[49] J. Zhang et al., “High-frequency electrical behaviour of CFRP,” Composite Structures, vol. 210, pp. 1–12, 2019.

[50] T. Hirano et al., “Conductive polyaniline-modified CFRP for lightning strike protection,” Composites Science and Technology, 2016.

[51] M. Katunin et al., “Lightning strike resistance of conductive polymer CFRP laminates,” Composite Structures, 2017.

[52] J. Zhu et al., “Electroless copper coating for CFRP lightning protection,” Surface and Coatings Technology, 2010.

[53] X. Chen et al., “Cold-spray tin–copper coating for composite lightning protection,” Journal of Thermal Spray Technology, 2015.

[54] R. Feraboli et al., “Lightning strike protection of composite structures,” SAMPE Journal, 2009.

[55] SAE International, “Lightning strike protection of aircraft,” SAE ARP5414, 2013.

[56] P. Manomaisantiphap et al., “Hybrid conductive polymer CFRP laminates,” Composite Structures, 2020.

[57] P. Manomaisantiphap et al., “Residual strength of hybrid conductive CFRP laminates,” Composites Part B, 2021.

[58] Y. Lin et al., “CNT/CB hybrid conductive CFRP for lightning strike protection,” Composite Structures, 2022.

[59] A. Review, “Electrical conductivity of CNT-filled CFRP composites,” Progress in Materials Science, 2022.

[60] Z. Wang et al., “Graphene film lightning strike protection for CFRP,” Composites Part A, 201

[61] Z. Wang et al., “Resin film infusion of graphene oxide for conductive CFRP,” Carbon, 2018

[62] M. Zamzami et al., “Hybrid CNT/CB conductive CFRP,” Carbon, 2019.

[63] R. Kumar et al., “Buckypaper interleaves for lightning strike protection,” Composites Part B, 2019

[64] M. Naguib et al., “MXene conductive coatings for composites,” Advanced Materials, 2020.

[65] L. Zhao et al., “MXene–graphene hybrid conductive films,” ACS Nano, 2019.

[66] Y. Zhang et al., “MXene-based lightning strike protection coatings,” Composites Part B, 2022.

[67] H. Liu et al., “MXene–Cu nanowire hybrid for lightning strike mitigation,” Advanced Functional Materials, 2023.

[68] J. Shukla, “Lightning strike damage mechanisms in composites,” Composite Structures, 2021.

DESIGN DEVELOPMENT AND TESTING OF MOTOR TEST RIG FOR AN EV WITH CONTINUOUSLY VARIABLE TRANSMISSION (CVT)

Vinod V[1], Muhammed Jaseem[2], Muhammed Shibil[3], Sreelakshmy PK[4], Sradha Shibu[5]
[1] Professor & Head, Dept of Mechanical Engg, NSS College of Engg, Palakkad, [2,3,4,5], Final year students, Dept of Mechanical Engg, NSS College of Engg, Palakkad
Email: vinodv@nssce.ac.in

ABSTRACT

The persistent demand for greater driving range and enhanced efficiency in Electric Vehicles (EVs) highlights a critical design limitation: the ubiquitous single-speed reduction gearbox. This fixed- ratio architecture frequently forces the electric drive motor to operate outside of its optimal speed- torque map, resulting in avoidable energy losses and compromising overall battery efficiency. This research documents the critical preliminary phase including the design, planning, and theoretical validation for developing a high-fidelity motor test rig specifically engineered to address this deficit through drivetrain innovation. The central solution proposed is the integration of a Continuously Variable Transmission (CVT). By offering a dynamic and continuous spectrum of gear ratios, the CVT enables the vehicle control unit to actively maintain the electric motor within its peak efficiency zone regardless of instantaneous operating conditions, such as acceleration profiles or road gradient changes. This strategy is expected to yield substantial gains in net power delivery and energy economy. The subsequent phases of this project will involve the physical construction and empirical validation of the test rig. The primary goal of the experimental effort will be the quantitative determination of the actual efficiency improvement. The final output will be a set of verified performance data establishing the clear advantage of the CVT-based drivetrain architecture for the next generation of highly efficient electric vehicles which will increase the mileage.

Keywords: Electric vehicles, continuously variable transmission, efficiency improvement, battery life.

I. INTRODUCTION

The rapid advancement of electric vehicles (EVs) is driven by the need for sustainable and efficient transportation. While most EVs use a single-speed gearbox that takes advantage of electric motors' high torque, this setup cannot always keep the motor operating at peak efficiency, especially in varying driving conditions. To overcome this, continuously variable transmissions (CVTs) are being explored, as they provide an infinite range of gear ratios and allow the motor to run closer to its optimal efficiency at all times. This leads to better energy use, smoother acceleration, improved driving comfort, and enhanced performance in situations like stop-and-go traffic or high-speed travel. Additionally, CVTs enable smarter control through real-time adjustments based on driving conditions, support efficient regenerative braking, and reduce mechanical stress on components. Overall, integrating CVTs into EVs offers a promising way to improve efficiency, range, and the overall driving experience.

II. LITERATURE REVIEW

Mazali et al. (2025) conducted an extensive experimental and simulation-based investigation into the application of electro-mechanical continuously variable transmission (EMCVT) systems in electric vehicles, developing a prototype that employed a power screw actuation mechanism with a dual-acting pulley arrangement to enable continuous ratio variation. Their testing across diverse operating conditions—such as varying speeds, torque demands, and clamping forces—allowed detailed monitoring of efficiency, slip, and mechanical losses, revealing that optimized clamping force can improve efficiency by nearly 10% while reducing energy consumption under the Worldwide Harmonized Light Vehicles Test Cycle (WLTC) compared to single-speed systems. Similarly, Y. Hori et al. (2019) analyzed the feasibility of integrating CVTs into electric vehicle powertrains, showing that while CVTs help maintain motor operation near peak efficiency during steady conditions, additional mechanical losses can offset some gains, highlighting the importance of careful design. Sharma and Gupta (2021) proposed an electro-mechanical CVT design that replaces hydraulic actuation, demonstrating through simulations improved efficiency, faster response, and enhanced low-speed torque delivery, potentially enabling smaller and more cost-effective motors. J. Liu et al. (2018) examined the relationship between clamping force, belt slip, and efficiency, concluding that both excessive and insufficient clamping force are detrimental, and proposing a dynamic control strategy to optimize performance in real time. Expanding on control strategies, Fu et al. (2022) found that maintaining a controlled level of slip rather than eliminating it

entirely improves efficiency and smoothness, achieving up to 5% energy savings and a 3–4% efficiency increase. Ruan et al. (2016) compared CVTs with conventional transmissions, demonstrating that CVTs can reduce energy consumption by 24–31%, particularly in urban driving conditions, thereby extending battery life and lowering costs. Wei, Hofman, and Caarls (2021) introduced a co-design optimization approach integrating CVT, battery, and motor systems, resulting in a lighter drivetrain and a 5.9% reduction in total cost of ownership. Hu et al. (2019) focused on loss mechanisms and optimal shift strategies, showing that intelligent control algorithms can significantly reduce frictional and hydraulic losses, while Liu et al. (2023) proposed a redesigned CVT with smart control and an electrically driven oil pump, achieving notable efficiency gains and up to 21% improvement in acceleration. Finally, Akehurst et al. (2004–2019) provided foundational insights into metal V-belt CVT systems through extensive modeling and experimentation, identifying key loss mechanisms and enabling design optimizations that improve efficiency, durability, and overall transmission performance in electric vehicles.

III. OBJECTIVES OF THE PROJECT & DETAILS OF TEST RIG

1. To study and analyze the feasibility of integrating a CVT in an electric vehicle drivetrain.
2. To design and develop a CVT-based transmission system suitable for electric vehicle applications.
3. To improve vehicle performance by optimizing torque delivery, acceleration, and energy efficiency using CVT.
4. To test and evaluate the performance of the EV with CVT under different driving conditions.
5. To compare the performance, efficiency, and range of the CVT-based EV with conventional EV without CVT.

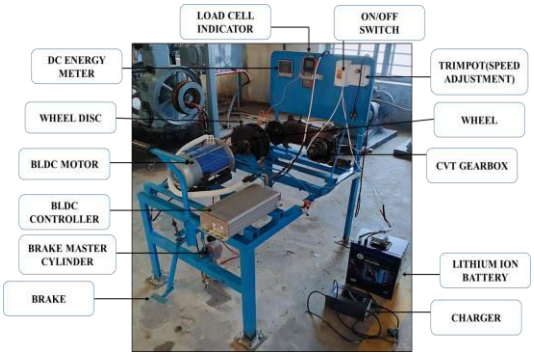
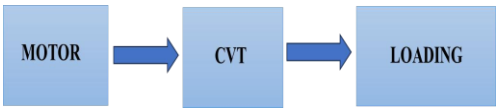


Fig. 1. Block diagram Representing Test Rig Setup with CVT



The test rig setup with CVT is designed to evaluate the performance of the electric motor when integrated with and without a Continuously Variable Transmission (CVT) under controlled loading conditions. In this configuration, the motor acts as the primary power source and is connected to the CVT system. The CVT is positioned between the motor and the hydraulic loading mechanism, allowing the transmission ratio to vary continuously based on operating conditions. As the motor runs, power is transmitted through the CVT, which adjusts the speed and torque before delivering it to the hydraulic load, thereby enabling smoother and more efficient power transfer.

The hydraulic loading mechanism applies resistance by regulating fluid pressure and flow, simulating real-world driving loads. With the presence of the CVT, the motor can operate closer to its optimal efficiency range by automatically adjusting the transmission ratio according to the load conditions.

Key performance parameters such as speed and torque are observed under different loading scenarios, and the output power is calculated accordingly. This setup helps in analyzing the improvement in efficiency, torque delivery, and overall performance when compared to the direct loading system without CVT, making it essential for validating the effectiveness of CVT integration in electric vehicles.

Table 1: Specifications of Continuously Variable Transmission

PARAMETER	SPECIFICATION
Vehicle Model	Honda Activa 110 cc
Transmission Type	Continuously Variable Transmission (CVT)
Drive Type	Belt Drive (V-belt)
Clutch Type	Centrifugal Automatic Clutch
Final Drive	Gear Reduction (Spur gears)
Cooling System	Air-cooled (indirect via engine airflow)
Lubrication	Gear oil (in final reduction unit)
Engagement Method	Automatic (RPM dependent)
Output	Direct drive to rear wheel
BLDC Motor Max Output Power	4000 W
Working Voltage	48VDC/ 60 VDC
Load Cell Capacity	200 Kg

The developed test rig is a comprehensive motor testing platform specifically built for electric vehicles integrated with a Continuously Variable Transmission (CVT). Its primary objective is to study and compare motor performance under two distinct operating conditions: with CVT and without CVT (direct drive). In the CVT mode, the transmission continuously adjusts the gear ratio to maintain optimal motor operating conditions, while in the direct drive mode, the motor is directly coupled to the load without any transmission variation. This unique dual-configuration capability allows detailed analysis of how CVT improves efficiency, torque delivery, and speed control in EV applications, making the setup highly valuable for research and development purposes.

The drivetrain of the test rig is designed with high flexibility and modularity. Currently, a BLDC motor is connected to the CVT gearbox, which transmits power to the wheel and disc arrangement. However, the system is not limited to a single motor type. By modifying the couplings and updating the control system, other motors such as PMSM or different EV drive units can be easily integrated into the setup. This makes the rig a universal testing platform for various electric propulsion systems. The inclusion of a BLDC controller, trim potentiometer for speed adjustment, and DC energy meter enables precise control and monitoring of input parameters like voltage, current, power consumption, and motor speed during experiments.

The loading arrangement plays a crucial role in simulating real-world operating conditions. A hydraulic braking system with a master cylinder is used to apply controlled load on the rotating wheel disc, thereby replicating road resistance. Simultaneously, a parallel brake line is connected to a load cell, which accurately measures the force generated during braking. This enables calculation of torque, efficiency, and performance characteristics of the motor-transmission system. Additionally, the entire test rig is rigidly mounted on a strong base frame to minimize vibrations and unwanted disturbances during operation. This ensures stability, improves measurement accuracy, and enhances safety. Overall, the test rig provides a robust and adaptable platform for detailed experimental analysis, performance comparison, and validation of EV motors with and without CVT integration.

IV. RESULTS & DISCUSSION

Results of Constant Load Variable Speed Test without CVT

Load (kg)	Motor speed (rpm)	Voltage (V)	Current (A)	Power (watts)	kwh	Time (Sec)
0	3000	65.05	15.77	967.7	0.00620	30
	3200	64.85	16.12	1058	0.00877	30
	3400	64.75	16.77	1082	0.01126	30
	3600	64.21	17.69	1106	0.01205	30
5	3000	64.70	15.85	1023	0.00850	30
	3200	63.86	16.28	1096	0.01015	30
	3400	63.70	16.88	1139	0.01080	30
	3600	63.63	17.87	1157	0.01115	30

Results of Constant Load Variable Speed Test with CVT

Load (kg)	Motor speed (rpm)	Voltage (V)	Current (A)	Power (watts)	kwh	Time (Sec)
0	3000	65.37	8.91	574.9	0.00493	30
	3200	64.35	16.56	1054	0.00714	30
	3400	64.25	16.81	1080	0.00894	30
	3600	63.91	18.92	1207	0.00983	30
5	3000	64.27	10.96	689.5	0.00684	30
	3200	63.86	14.26	859.3	0.00819	30
	3400	63.46	16.71	1060	0.00897	30
	3600	62.83	18.17	1139	0.00956	30

V. CONCLUSION

The work presents the successful design and development of a versatile motor test rig for electric vehicle applications integrated with a Continuously Variable Transmission (CVT), aimed at comparing motor performance with and without CVT. The system, currently equipped with a BLDC motor, features a modular and flexible architecture that allows easy replacement with other motor types such as PMSM or induction motors, making it suitable for extended research. Experimental results under controlled loading conditions demonstrate that incorporating a CVT significantly enhances system efficiency—by approximately 10–20%—as it enables the motor to operate closer to its optimal efficiency range across varying loads.

Overall, the developed system not only validates the effectiveness of CVT in enhancing EV performance but also serves as a robust experimental platform for future investigations. The findings of this work highlight the potential of integrating variable transmission systems in electric vehicles to

improve energy utilization, reduce losses, and enhance overall system efficiency

This project contributes to electric vehicle (EV) powertrain development by demonstrating the practical advantages of integrating a Continuously Variable Transmission (CVT) with electric motor systems. Unlike conventional EVs that use fixed or single-speed transmissions, this work shows that a CVT can improve efficiency and performance by keeping the motor within its optimal operating range. A key contribution is the development of a flexible and modular motor test rig capable of testing different motor types such as BLDC, PMSM, and induction motors, making it valuable for research and educational use. Experimental results indicate that CVT integration can enhance system efficiency by approximately 10–20% compared to direct drive systems. The project also provides insights into torque and speed optimization, enabling smoother acceleration and better performance across varying speeds. Additionally, it offers a scalable platform for future advancements, including advanced control methods and comparative studies, effectively bridging the gap between theoretical benefits and real-world application of CVTs in EVs. This introduction of CVT in EV system will ultimately reduce the energy consumption and increase the mileage.

REFERENCES

[1] Jianjun Hu, Feng Xiao, Hang Peng, Wei Zhao,CVT discrete speed ratio optimizations based on energy efficiency for PHEV, Alexandria Engineering Journal, 61 (4), pp. 4095– 4105, 2022.

[2] Caiyang Wei, Theo Hofman, and Esin Ilhan Caarls, Co-Design of CVT-Based Electric Vehicles, Energies, 14(7), Article 1825, 2021.

[3] R. Suganya, L.M.I. Leo Joseph, S. Kollem, Understanding lithium-ion battery management systems in electric vehicles: environmental and health impacts, comparative study, and future trends: a review, Results Engineering, 24 ,103047, 2024.

[4] Baoyu Zhou, Zhejun Li, et al.,Design and Optimization of an Electric Vehicle Powertrain Based on an Electromechanical Efficiency Analysis, Processes, 13(6), 1698. 2025.

[5] S. Jain, P. Kumar, and A. Sharma, Performance Analysis of Continuously Variable Transmission in Electric Vehicles, International Journal of Automotive Technology, 22(3), 745–756, 2021.

[6] M. Ehsani, Y. Gao, and A. Emadi, Modern Electric, Hybrid Electric, and Fuel Cell Vehicles: Fundamentals, Theory, and Design, CRC Press, 3rd Edition, 2018.

[7] T. Hofman, M. Steinbuch, R. van Druten, and A. Serrarens, Rule-Based Energy Management Strategies for Hybrid Vehicles, International Journal of Electric and Hybrid Vehicles, 1(1), 71–94, 2007.

[8] J. Larminie and J. Lowry, Electric Vehicle Technology Explained, Wiley, 2nd Edition,.2012.

[9] H. Zhang, X. Wang, and Y. Chen, Optimization of Transmission Systems for Electric Vehicles Based on Efficiency Maps, IEEE Transactions on Vehicular Technology, 69(5), 5173–5184, 2020.

[10] K. T. Chau, Electric Vehicle Machines and Drives: Design, Analysis and Application, Wiley-IEEE Press, 2015.

[11] F. Bottiglione, G. Mantriota, and A. Sorniotti, Energy Consumption of a Battery Electric Vehicle with CVT Transmission, Mechanism and Machine Theory, 86, 123– 135, 2015.

[12] A. Khaligh and Z. Li, Battery, Ultracapacitor, Fuel Cell, and Hybrid Energy Storage Systems for Electric, Hybrid Electric, Fuel Cell, and Plug-in Hybrid Electric

Vehicles, IEEE Transactions on Vehicular Technology, 59(6), 2806–2814, 2010.

[13] Y. Wang, C. Zhang, and J. Chen, Control Strategy Optimization for CVT-Based Electric Vehicles, Applied Energy, 226, 120–131, 2018.

[14] B. Gao, H. Chen, K. Sanada, and Y. Hu, Design and Control of Electric Vehicle Drivetrain with CVT System, SAE Technical Paper, 2017-01-1152, 2017.

TRACK-6

OTHERS

Data vs. Model Machine Learning Fairness Testing: An Empirical Study

RESHMA P K
Dept. of Artificial intelligence and machine learning
ASET, Palakkad,India
reshma.pk@ahalia.ac.in

VITHUL M
Dept. of Artificial intelligence and machine learning
ASET, Palakkad,India
vithul.m@ahalia.ac.in

SANKAR KRISHNA P
Dept. of Artificial intelligence and machine learning
ASET, Palakkad,India
sankar.p@ahalia.ac.in

SREYA S
Dept. of Artificial intelligence and machine learning
ASET, Palakkad,India
sreya.s@ahalia.ac.in

DARSHAN R M
Dept. of Artificial intelligence and machine learning
ASET, Palakkad,India
darshan.rm@ahalia.ac.in

***Abstract*—Machine learning systems are increasingly deployed in high-stakes domains such as healthcare, finance, and hiring, where fairness is critical. However, bias can arise from both data and model design, leading to discriminatory outcomes. This paper presents an empirical study comparing data-centric and model-centric approaches to fairness testing. The study evaluates how dataset imbalance, feature selection, and model architecture influence fairness metrics such as demographic parity, equal opportunity, and disparate impact. Experiments are conducted using multiple classifiers including Logistic Regression, Random Forest, and Neural Networks on benchmark datasets. Results indicate that data-level interventions such as re-sampling and re-weighting significantly improve fairness, while model-level techniques provide additional but limited gains. The study concludes that fairness is primarily a data problem, with model adjustments acting as secondary optimization. Ethical considerations and implications for responsible AI deployment are also discussed.**

Keywords—machine learning fairness, bias detection, data bias, model bias, ethical AI, fairness metrics, classification

I. INTRODUCTION

Machine learning models are widely used for automated decision-making, yet they often inherit biases from training data or amplify them through model design. These biases can lead to unfair outcomes, particularly for underrepresented groups. As machine learning systems are increasingly used in critical applications, ensuring fairness has become a key concern.

Fairness in machine learning has attracted significant research attention, with efforts focused on identifying, measuring, and mitigating bias. However, a fundamental question remains unresolved regarding whether fairness issues originate primarily from data or from the model itself. This distinction is important because it determines where corrective efforts should be concentrated.

This paper investigates the relative impact of data and model factors on fairness. The study evaluates how dataset imbalance and algorithm choice affect fairness metrics and overall model performance, with the objective of identifying the dominant source of bias..

II. BACKGROUND AND RELATED WORK

A. Fariness in Machine Learning

Fairness in machine learning refers to the absence of discrimination against individuals based on sensitive attributes such as gender, race, or age. Several formal definitions of fairness exist, including demographic parity, equal opportunity, and equalized odds. Each definition captures a different aspect of fairness and may lead to conflicting requirements in practical scenarios.

B. Sources of bias

Bias in machine learning systems arises from multiple sources. Data bias occurs when the training dataset is imbalanced or not representative of the population. Algorithmic bias originates from model assumptions and optimization strategies that may favor certain groups. Measurement bias results from incorrect labeling or poor feature selection. These sources often interact, making bias difficult to isolate.

III. PROPOSED FRAMEWORK

The proposed framework evaluates fairness from both data-centric and model-centric perspectives. The system consists of data preprocessing, model training, fairness evaluation, and comparative analysis stages. Controlled experiments are conducted to isolate the impact of each component.

The dataset used contains sensitive attributes such as gender or ethnicity. To simulate real-world bias, controlled imbalance is introduced. Multiple models including Logistic Regression, Random Forest, and Neural Networks are trained under identical conditions to ensure fair comparison.

A. Model Selection and Training

Three categories of models are evaluated. Traditional machine learning models include Logistic Regression and Random Forest. Neural network models are used to analyze complex patterns in data. All models are trained using standard optimization techniques with consistent training and testing splits.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Class imbalance is handled using re-sampling techniques such as oversampling and undersampling. Hyperparameters are tuned using cross-validation to ensure optimal performance.

B. Experimental Evaluation

Experiments are conducted to compare data-centric and model-centric fairness improvements. Fairness metrics including demographic parity difference, equal opportunity difference, and disparate impact ratio are computed for each model. Accuracy is also measured to evaluate performance trade-offs.

Results indicate that data balancing significantly improves fairness metrics across all models. Improvements range between 20 and 35 percent when data-level techniques are applied. Model-level techniques provide smaller improvements, typically between 5 and 15 percent.

C. Ethical and Practical Considerations

The deployment of fair machine learning systems requires careful consideration of ethical issues. Bias must be evaluated across different demographic groups to ensure equitable outcomes. Transparency in decision-making and explainability of models are essential for building trust.

The system is designed as a decision support tool rather than a replacement for human judgment. Human oversight remains necessary in critical applications to prevent unintended consequences. Proper data governance and privacy protection measures must also be implemented.

D. Conclusion

This study presents a comparative analysis of data-centric and model-centric approaches to fairness in machine learning. The results demonstrate that fairness is primarily influenced by data quality rather than model complexity. While model-level techniques contribute to fairness improvement, their impact is limited without proper data handling.

Future work will focus on real-world deployment scenarios and the development of adaptive systems capable of continuously monitoring and mitigating bias.

ACKNOWLEDGMENT

The authors acknowledge the support of their institution and peers for providing resources and guidance for this research.

REFERENCES

[1] S. Barocas, M. Hardt, and A. Narayanan, “Fairness in Machine Learning,” 2019.
[2] M. Hardt et al., “Equality of Opportunity in Supervised Learning,” NIPS, 2016.
[3] F. Kamiran and T. Calders, “Data preprocessing techniques for classification without discrimination,” 2012.
[4] A. Mehrabi et al., “A survey on bias and fairness in machine learning,” ACM, 2021.

Funded by: [Funding Agency Name]. Grant No: [Grant Number].

Regenerative Effects of α-Mangostin-Loaded ZnO Nanoparticles on Wound Healing: In Vitro and Zebrafish In Vivo Insights

Taniya Mary Martin,
Department of Anatomy, Zebrafish Facility Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai-600077, Tamil Nadu, India
taniyam.sdc@saveetha.com

Meenakshi Sundaram Kishore Kumar*
Department of Anatomy, Zebrafish Facility Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai-600077, Tamil Nadu, India
meenakshisundaram.sdc@saveetha.com

***Abstract*— Introduction:** α-Mangostin, a key bioactive compound derived from the peel of Garcinia mangostana, is known for its wound-healing potential. This study aims to evaluate the synergistic effects of α-Mangostin-derived zinc oxide nanoparticles (α-Mangostin–ZnO NPs) on wound healing, with emphasis on antimicrobial, anti-inflammatory, antioxidant, and molecular mechanisms.

Methods: α-Mangostin–ZnO nanoparticles were assessed using in vitro antimicrobial assays against wound-associated pathogens (Staphylococcus aureus, Pseudomonas aeruginosa, Escherichia coli, and Vibrio vulnificus). Anti-inflammatory activity was evaluated by protein denaturation inhibition using ibuprofen as a reference, while antioxidant potential was assessed using the DPPH free radical scavenging assay. In silico molecular docking was performed against key target proteins (7S53, 4NG2, 20MF, and 5TV2). In vivo wound-healing efficacy was examined using a zebrafish wound model, along with RT-PCR analysis of inflammatory gene expression.

Results: α-Mangostin–ZnO NPs exhibited enhanced antimicrobial activity compared to individual components. Significant inhibition of protein denaturation and strong DPPH radical scavenging activity indicated potent anti-inflammatory and antioxidant effects. Molecular docking revealed strong binding affinities with antimicrobial, oxidative stress, inflammatory, and wound-healing proteins. RT-PCR analysis showed marked downregulation of pro-inflammatory markers (mmp13, TNF-α, and IL-1β). Zebrafish studies demonstrated accelerated wound closure and improved tissue regeneration.

Conclusion: The findings demonstrate that α-Mangostin–ZnO NPs act as a multifunctional wound-healing agent by integrating antimicrobial, antioxidant, and anti-inflammatory actions, supported by molecular targeting. This nanoplatform shows strong potential for future wound-healing applications.

Keywords*— *α-Mangostin, Human, Health, Disease, Medicine

INTRODUCTION

Wound healing is a complex and dynamic process that involves a series of coordinated biological events, including haemostasis, inflammation, cell proliferation, and tissue remodelling[1]. These stages are essential for restoring the integrity of damaged tissues and ensuring effective recovery. Despite notable advancements in medical treatments, managing chronic wounds—those that do not progress through the normal stages of healing—continues to be a significant challenge, highlighting the need for innovative therapeutic strategies. Chronic wounds are often complicated by persistent infections, prolonged inflammation, and impaired tissue regeneration, which demand advanced solutions capable of addressing these multifaceted issues[2].

Recent developments in nanomedicine have highlighted the potential of α-Mangostin-ZnO NPs as novel agents for wound healing management. This combination harnesses the individual advantages of α-Mangostin and zinc oxide nanoparticles to tackle the diverse challenges associated with wound healing. The unique properties of α-Mangostin-ZnO NPs may enhance the healing process through their antibacterial, anti-inflammatory, and antioxidant effects, positioning them as promising candidates for the treatment of chronic wounds[3]. α-Mangostin-ZnO NPs have emerged as an innovative therapeutic approach in wound healing, integrating the unique biological activities of both α-Mangostin and Zinc Oxide Nanoparticles (ZnO NPs). ZnO NPs are renowned for their exceptional antibacterial properties, which are essential for preventing infections in chronic wounds. Their ability to reduce bacterial load is particularly important, as infections often complicate and hinder the healing process. In addition to their antimicrobial effects, ZnO NPs also exhibit anti-inflammatory properties that help regulate the inflammatory response, thus reducing excessive inflammation that can delay recovery. Furthermore, these nanoparticles possess antioxidant capabilities that combat oxidative stress, a factor detrimental to cellular health and wound healing. By promoting fibroblast proliferation and facilitating collagen synthesis, ZnO NPs play a crucial role in forming new connective tissue, ensuring tissue strength and integrity[4].

Complementing these advantages, α-Mangostin is a natural polyphenolic compound known for its wide range of biological activities[5]. It provides antioxidative, antimicrobial, and anti-inflammatory benefits and has traditionally been used in herbal medicine to support healing. α-Mangostin effectively scavenges free radicals, reducing oxidative stress and preventing cellular damage. Its potent antimicrobial properties extend to a variety of pathogens, including those resistant to antibiotics, making it a valuable tool in managing wound infections. Additionally, α-Mangostin helps regulate inflammatory pathways by inhibiting the production of pro-inflammatory cytokines and

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

enzymes, thereby enhancing its efficacy in controlling inflammation associated with wounds. Studies have shown that α-Mangostin-ZnO NPs can significantly enhance wound healing outcomes. For instance, these nanoparticles have been reported to expedite the healing of diabetic wounds, which are particularly challenging due to compromised immune responses and inadequate circulation. α-Mangostin-ZnO NPs have exhibited the ability to promote re-epithelialization, aiding the migration of new epithelial cells to cover wounds while effectively reducing bacterial load[6]. Their effectiveness has also been assessed in the context of burn wound healing, where α-Mangostin-ZnO NPs were found to decrease inflammation and facilitate faster tissue regeneration compared to conventional therapies. The potential of α-Mangostin-ZnO NPs is not only promising but also underscores the necessity of exploring their synergistic effects when combined with other therapeutic agents. This synergy reflects the enhanced therapeutic outcomes observed when α-Mangostin-ZnO NPs are administered alongside additional treatments, resulting in greater efficacy than when the components are used individually. Understanding these interactions is crucial for the development of more effective wound healing therapies that capitalize on the strengths of α-Mangostin-ZnO NPs. To further investigate these effects, the zebrafish model provides an excellent in vivo system for evaluating the impact of α-Mangostin-ZnO NPs on wound healing in a living organism. Zebrafish are particularly advantageous for biomedical research due to their genetic similarities to humans, rapid wound healing capabilities, and transparent embryos that enable real-time observation of healing processes[7]. This study aims to evaluate the efficacy of α-Mangostin-ZnO NPs in promoting wound closure, tissue regeneration, and reduction of inflammation, thereby providing insights into their potential mechanisms of action. The underlying hypothesis of this research posits that α-Mangostin-ZnO NPs, especially when combined with other therapeutic agents, will significantly enhance the wound healing process. This enhancement may occur through various mechanisms, such as the combined antimicrobial effects of α-Mangostin-ZnO NPs with additional antibacterial agents, concurrent anti-inflammatory actions with other anti-inflammatory drugs, or a synergistic promotion of collagen synthesis alongside other regenerative compounds. Exploring these interactions could lead to innovative wound management strategies that surpass the effectiveness of current treatments[5].

Research indicates that α-Mangostin-ZnO NPs possess robust antimicrobial activity against a wide range of microorganisms, including bacteria, fungi, and viruses. This broad-spectrum efficacy is especially valuable in combating pathogenic microorganisms, providing both preventive and therapeutic applications. Studies have demonstrated that α-Mangostin-ZnO NPs can inhibit the growth of antibiotic-resistant strains, presenting a promising alternative in an era marked by increasing antimicrobial resistance. Moreover, the anti-inflammatory properties of α-Mangostin-ZnO NPs further enhance their therapeutic profile. By modulating inflammatory pathways and inhibiting the production of pro-inflammatory cytokines and enzymes, these nanoparticles may help alleviate chronic inflammatory conditions. Evidence from various animal models supports the efficacy of α-Mangostin-ZnO NPs in reducing inflammation, indicating their potential usefulness in treating conditions such as inflammatory bowel disease and arthritis. In the realm of herbal medicine, α-Mangostin-ZnO NPs are increasingly utilized for their diverse therapeutic effects, often in combination with other bioactive compounds to improve efficacy. Their integration into modern formulations for skin health showcases their versatility. For example, topical preparations containing α-Mangostin-ZnO NPs have been developed for treating acne and eczema, capitalizing on their antioxidative and antimicrobial properties. Clinical trials have demonstrated significant reductions in acne lesions and improvements in skin texture, highlighting their potential in dermatological applications[8].

The food industry is also investigating the use of α-Mangostin-ZnO NPs as natural preservatives and additives. Their antimicrobial and antioxidative properties can help prolong the shelf life of perishable products, aligning with consumer preferences for clean-label solutions. Recent studies have shown that α-Mangostin-ZnO NPs can significantly slow spoilage in fresh-cut fruits, positioning them as a promising natural alternative in food preservation5. Additionally, the neuroprotective potential of α-Mangostin-ZnO NPs is garnering interest. Their ability to cross the blood-brain barrier and protect neuronal cells suggests they could be therapeutic options for neurodegenerative diseases such as Alzheimer's and Parkinson's. Research indicates that α-Mangostin-ZnO NPs can mitigate oxidative damage and inflammation in neuronal models, contributing to reduced amyloid-beta plaque formation and overall neuroprotection. The wide array of biological activities exhibited by α-Mangostin-ZnO NPs—including antioxidative, antimicrobial, and anti-inflammatory effects—makes them a highly versatile compound with numerous applications in medicine. The rising interest in natural products, coupled with the demand for effective, safe, and sustainable therapeutic agents, highlights the necessity for further research into the potential of α-Mangostin-ZnO NPs. As our understanding of their mechanisms improves, α-Mangostin-ZnO NPs are likely to play a significant role in the development of innovative treatments for a variety of health conditions, ranging from chronic inflammatory diseases to neurodegenerative disorders and skin health. The ongoing exploration of their applications in both traditional and modern medicine holds great promise for advancing natural therapeutic strategies that cater to the increasing demand for health products derived from nature. This study aims to rigorously examine the synergistic effects of α-Mangostin-ZnO NPs on wound healing through a combination of in vitro assays and zebrafish model studies, ultimately providing valuable insights into their efficacy and potential clinical applications[5,9].

MATERIALS AND METHODS

2.1 Preparation of α-Mangostin-ZnO NPs

α-Mangostin-ZnO NPs are synthesized by dissolving 24.55 g of zinc acetate dihydrate ($Zn(CH3CO2)2·2H2O$) in 100 ml of deionized water to create a 0.1 M solution. Concurrently, 4 g of sodium hydroxide (NaOH) is dissolved in 100 ml of deionized water to prepare another 0.1 M solution. α-Mangostin an active compound found in the peel of Garcinia mangostana is prepared at a concentration of 1 mg/ml and is utilized as a reducing agent. The zinc acetate solution is then added dropwise to the sodium hydroxide solution while stirring vigorously. The mixture is heated until a white precipitate forms, indicating the formation of the nanoparticles. After cooling, the precipitate is collected

through centrifugation and washed with ethanol and deionized water to remove impurities. The resulting α-Mangostin-ZnO NPs are then dried and characterized using various techniques. Fourier Transform Infrared Spectroscopy (FTIR) analysis is conducted by preparing a pellet of the nanoparticles for examination with an FTIR spectrometer. X-ray Diffraction (XRD) analysis is performed on a powdered sample of the nanoparticles. Lastly, UV analysis involves dispersing the nanoparticles in ethanol and measuring the absorbance spectrum using a UV-Vis spectrophotometer[5,9,10].

2.2 Invitro anti-inflammatory and antioxidant analysis

2.2.1 Albumin Inhibition assay

To evaluate the inhibition of albumin denaturation, a stock solution of human serum albumin (HSA) was prepared by dissolving it in phosphate buffer at pH 7.4 to achieve a concentration of 5 mg/ml. Concurrently, the synthesized α-Mangostin-ZnO NPs were dispersed in an appropriate solvent to reach the desired concentration. Test solutions of anti-inflammatory agents were then created by diluting them in the phosphate buffer. In a series of test tubes, equal volumes of the HSA solution, the test compounds, and the nanoparticle dispersion were mixed and incubated at a controlled temperature of 37°C for one hour. A separate set of HSA tubes underwent the same incubation conditions without the nanoparticles, serving as a positive control for denaturation. Following the incubation period, the samples were allowed to cool and then subjected to heat stress by heating at 70°C for 20 minutes. The absorbance of each sample was then measured at 660 nm using a UV-Vis spectrophotometer. A decrease in absorbance in the samples containing α-Mangostin-ZnO NPs, compared to the control, indicated inhibition of albumin denaturation, suggesting the potential anti-inflammatory properties of the nanoparticles[11].

2.2.2 Hydroxyl radical scavenging assay

A reaction mixture was prepared that included a hydroxyl radical source, specifically Fenton's reagent (a 1:1 mixture of 10 mM FeSO4 and 10 mM H2O2 solutions), along with varying concentrations of α-Mangostin-Derived Zinc Oxide Nanoparticles (α-Mangostin-ZnO NPs). The reaction mixture was incubated at room temperature. To initiate the reaction between the hydroxyl radicals and salicylic acid, salicylic acid was added, leading to the formation of 2,3-dihydroxybenzoic acid. After the incubation period, the reaction was halted by adding trichloroacetic acid (TCA), followed by the introduction of thiobarbituric acid (TBA) to produce a pink chromogen. The tubes were then heated, allowed to cool, and subsequently centrifuged. Finally, the absorbance of the supernatant was measured at a wavelength of 532 nm using a UV-Vis spectrophotometer to evaluate the progress of the reaction.

2.3. Minimum Inhibitory Concentration

Each pathogen—S. aureus, P. aeruginosa, E. coli, and V. vulnificus—was cultured separately in the appropriate growth media until reaching mid-log phase. The cultures were then diluted to obtain a standardized bacterial suspension of approximately 5 × 10^6 CFU/mL. A series of dilutions of the α-Mangostin-ZnO NPs stock solution were prepared in the respective growth media to cover a range of concentrations.

In a 96-well microplate, aliquots of the bacterial suspension were dispensed into each well, followed by the addition of different concentrations of α-Mangostin-ZnO NPs. Positive and negative controls were included, with amoxicillin serving as the positive control and an untreated sample as the negative control. The microplate was then incubated at the appropriate temperatures for each pathogen (37°C for S. aureus and E. coli, 42°C for V. vulnificus, and 30°C for P. aeruginosa) for 18-24 hours. After incubation, bacterial growth was assessed either visually or by measuring optical density at a suitable wavelength using a microplate reader. The Minimum Inhibitory Concentration (MIC) was defined as the lowest concentration of α-Mangostin-ZnO NPs at which no visible growth or a significant decrease in optical density was observed compared to the control[12].

2.4. Zone of Inhibition

In the well diffusion method, agar plates were prepared by pouring molten agar containing the appropriate growth media for each pathogen, such as nutrient agar for S. aureus and E. coli, Mueller-Hinton agar for P. aeruginosa, and tryptic soy agar supplemented with 2% NaCl for V. vulnificus. Once solidified, bacterial suspensions of 5 × 10^6 CFU/mL were evenly spread across the surface of the agar plates using a sterile swab. Wells were then created in the agar using a sterile cork borer, and various concentrations of the α-Mangostin-ZnO NPs solution were added to the wells. Amoxicillin was included as a positive control, along with an untreated control. The plates were incubated at room temperature, and after incubation, the zones of inhibition around the wells containing α-Mangostin-ZnO NPs were measured in millimeters (mm) using a calibrated ruler[13].

2.5 Ligand and Receptor interaction studies

The experimental procedure began with retrieving crystal structures of receptors relevant to S. aureus, E. coli, P. aeruginosa, and V. vulnificus from the Protein Data Bank (https://www.rcsb.org/). PyMOL (https://pymol.org/) was used to prepare these receptor structures by removing water molecules and non-essential ligands. The 3D structure of the α-Mangostin-ZnO NPs was obtained from relevant literature and converted into appropriate formats using Open Babel GUI Ver 2.0. AutoDock VER 1.5.7 was employed for molecular docking studies to predict binding interactions between α-Mangostin-ZnO NPs and the receptors. Docking parameters were set, and binding affinity was evaluated by analysing the results, including binding energy and interaction patterns. Finally, Discovery Studio Visualizer software was used to visualize the 3D and 2D binding interactions between α-Mangostin-ZnO NPs and the receptor proteins.

2.6 In vivo Studies

2.6.1. Zebrafish Toxicity

All zebrafish experiments were conducted following approval from the Institutional Animal Ethical Committee of Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS) (Approval No.: BRULAC/SDCH/SIMATS/IAEC/06-2023/15). The zebrafish were acclimatized in a controlled aquatic environment maintained at a temperature of 28°C ± 1°C, with a pH range of 7.0-7.5 and a 14:10 light-dark cycle. During the acclimatization period, the zebrafish were housed in 10 L glass tanks filled with dechlorinated water, equipped

with aeration and filtration systems. After a one-week acclimatization phase to ensure physiological adjustment, the zebrafish were randomly assigned to experimental groups, each consisting of 6 individuals to enable adequate statistical analysis. The experimental groups were exposed to varying concentrations of α-Mangostin-ZnO NPs, ranging from 40 µg/mL to 80 µg/mL. Throughout the one-week exposure

period, the survival rates of the zebrafish were monitored daily and calculated as the percentage of surviving zebrafish relative to the total number in each group at the start of the experiment.

2.6.2. Zebrafish wound closure experiments

In the in vivo wound healing experiment, zebrafish (n=6/group) were divided into four groups: a control group, wounded zebrafish, and two experimental groups treated with α-Mangostin-ZnO NPs at concentrations of 40 µg/mL and 80 µg/mL, respectively. The fish were anesthetized by immersion in a 0.2% Tricaine methane sulfonate solution before being wounded with a sterilized scalpel behind the gill region, near the lateral line. The α-Mangostin-ZnO NPs were administered to the wounded zebrafish every third day, and the healing process was monitored through daily photographs. Zebrafish were anesthetized at specific intervals (0, 7, and 14 days post-injury), and the wounded sites were photographed and analysed for closure using ImageJ software. The progress of wound healing was assessed by periodically measuring the wound area, with injured locations identified by observing differences in skin color between affected and unaffected regions. Healing progression continued to be tracked until the wounded area fully regenerated and returned to its original coloration, indicating the completion of the healing process[14].

2.6.4. Gene expression studies

Adult zebrafish were categorized into four groups: treatment 1 (α-Mangostin-Derived Zinc Oxide Nanoparticles at 40 µg/mL), treatment 2 (α-Mangostin-ZnO NPs at 80 µg/mL), model (wounded zebrafish), and unwounded control. Tissue samples from the wounded area were collected and stored at -80°C. The pooled wound tissue underwent homogenization using a Polytron PT 1200E (Kinematic, Switzerland) for RNA isolation with TRIzol reagent (Invitrogen, Carlsbad, CA, USA). The RNA concentration was determined using a NanoDrop One, followed by cDNA synthesis using 2.5 µg of the isolated RNA and a PrimeScript 1st strand cDNA synthesis kit (TaKaRa®, Tokyo, Japan) according to the manufacturer's instructions. The synthesized cDNA samples were diluted 30-fold with nuclease-free water and stored at -20°C until further analysis. The primers utilized in this study are detailed in Table 1, with zebrafish β-actin serving as the housekeeping gene. Quantitative real-time PCR (qRT-PCR) was performed using a Thermal Cycler Dice Real-Time System (TaKaRa®) following a previously published protocol with minor modifications. Each qRT-PCR reaction consisted of 3 µL of cDNA template, 1 µL (10 µM) of each forward and reverse primer, and 5 µL of TB Green Premix Ex Taq II (TaKaRa®) qPCR mix. Primer specificity was confirmed through a standard three-step thermal cycling profile with an annealing temperature of 56°C, followed by a single dissociation reading step. Relative expression fold changes were calculated using the 2-ΔΔCT method[15].

Table 1.

Primers used for gene expression

Gene	Forward primer (5´-3´)	Reverse primer (5´-3´)	Accession number
mmp13	GAGAAGGTTTGGGCTCTCTATG	TGAGTTGCTGTCTTCCTTGTAG	AF506756
tnf-α	AGAAGGAGAGTTGCCTTTACCGCT	AACACCCTCCATACACCCGACTTT	AF025305
il-1β	TCAAACCCCAATCCACAGAG	TCACTTCACGCTCTTGGATG	AY340959.
β actin	AATCTTGCGGTATCCACGAGACCA	TCTCCTTCTGCATCCTGTCAGCAA	AF025305

2.6.5 Statistical Analysis

Statistical analyses were conducted using GraphPad Prism software (Version 8, San Diego, CA, USA). To evaluate the significance between the control and treated groups, one-way analysis of variance (ANOVA) was employed, followed by Tukey's post hoc test. A significance threshold of $p < 0.05$ was established, and all data are presented as mean ± standard deviation.

3. RESULTS

The characterization results indicate that the α-Mangostin-ZnO NPs possess distinct structural and chemical properties. FTIR analysis identifies characteristic peaks at wave numbers 1644.96, 1353.42, 1204.72, 1025.82, 888.15, 761.99 and 541.38, indicating the presence of functional groups associated with both α-Mangostin and zinc oxide (Fig 1). These peaks correspond to various chemical bonds and molecular vibrations, offering insight into the composition and bonding within the α-Mangostin-ZnO NPs. XRD analysis reveals that the nanoparticles exhibit a crystalline phase of approximately 30.1% and an amorphous phase of approximately 69.9% (Fig 2). This crystallinity indicates the presence of ordered atomic arrangements within the nanoparticles, which can influence their physical and chemical properties. Furthermore, UV spectroscopy displays a peak at 268 nm, suggesting the occurrence of electronic transitions within the α-Mangostin-ZnO NPs (Fig 3).

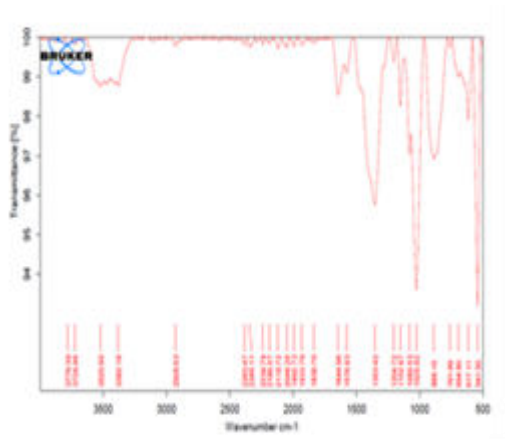


Fig.1: FTIR spectrum represent specific vibrational modes corresponding to functional groups present in the α-Mangostin-ZnO NPs

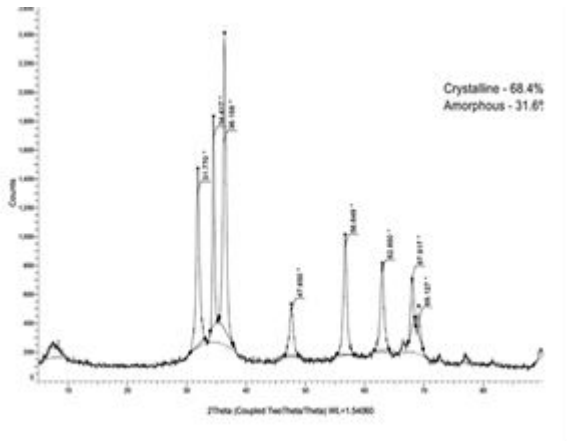


Fig. 2: XRD analysis of α-Mangostin-ZnO NPs illustrating the presence of a crystalline phase and amorphous phase

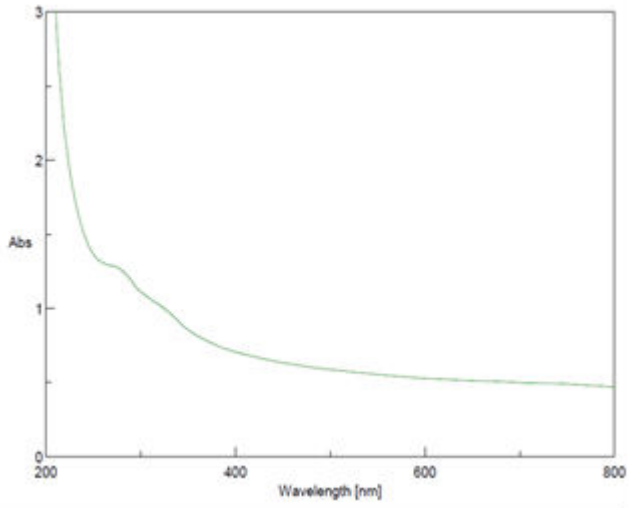
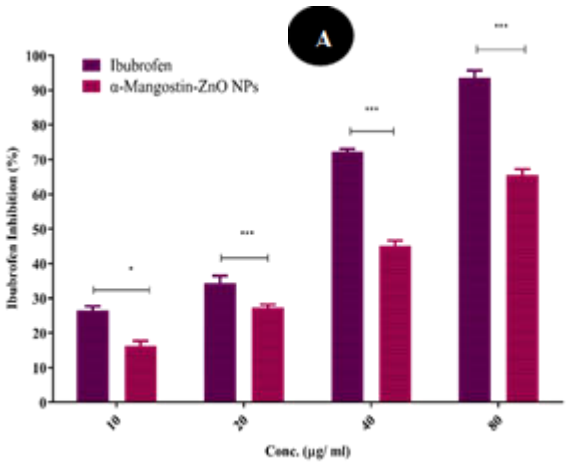


Fig. 3: UV-Vis spectrum of synthesized α-Mangostin-ZnO NPs

3.2. Inhibition of albumin denaturation and Free radical scavenging activity

The results of the study on α-Mangostin-ZnO NPs indicate promising properties in combating protein denaturation and oxidative stress. In the albumin denaturation experiment, α-Mangostin-ZnO NPs at the higher concentration of 80 µg/mL exhibited 63% activity, which is slightly lower than the positive control, Ibuprofen, showing 72% activity. This suggests that α-Mangostin-ZnO NPs effectively prevent protein denaturation in a manner similar to diclofenac. Similarly, in the hydroxyl radical scavenging assay, α-Mangostin-ZnO NPs at 80 µg/mL demonstrated 69% activity, comparable to the positive control, ascorbic acid, which exhibited 75% activity. These results indicate that α-Mangostin-ZnO NPs possess significant antioxidant capabilities, effectively neutralizing harmful hydroxyl radicals. Overall, these findings suggest that α-Mangostin-ZnO NPs hold promise as therapeutic agents for addressing conditions related to protein misfolding and oxidative damage (Fig 4).



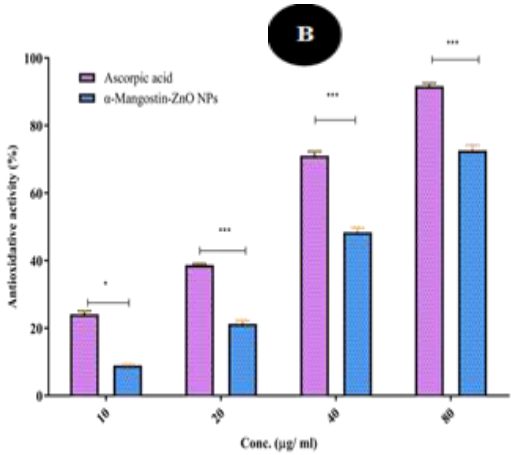


Fig. 4: (A) Albumin denaturation Inhibition and (B) Hydroxyl radical scavenging activity of α-Mangostin-ZnO NPs at different concentrations (10 µg/mL, 20 µg/mL, 40 µg/mL, and 80 µg/mL). Diclofenac and Trolox was used as a positive control. The *** represented the level of significance ($p < 0.05$) when the results were compared to the control. Data were presented as mean ± SD of three independent experiments.

3.3. Inhibition of Biofilm formation

The MIC experiment with α-Mangostin-ZnO NPs demonstrated their effectiveness against a variety of pathogenic bacteria, including S. aureus, P. aeruginosa, E. coli, and V. vulnificus. The MIC values were consistently recorded at 40 µg/mL for all tested bacteria, indicating that concentrations at or above this threshold significantly ($p < 0.05$) inhibited bacterial growth (Fig 5). This underscores the broad-spectrum antimicrobial activity of α-Mangostin-ZnO NPs across both Gram-positive and Gram-negative bacteria. Furthermore, the zone of inhibition assay confirmed the antimicrobial potential of α-Mangostin-ZnO NPs. At a concentration of 80 µg/mL, the NPs produced inhibition zones measuring 7 mm, 6 mm, 4 mm, and 8 mm against S. aureus, P. aeruginosa, E. coli, and V. vulnificus, respectively. Increasing the concentration to 80 µg/mL resulted in larger zones of inhibition (8 mm, 10 mm, 5 mm, and 7 mm), demonstrating a dose-dependent response. Notably, the zones of inhibition exceeded those of the positive control, amoxicillin (50 µg/mL), indicating that at 80 µg/mL, the antimicrobial activity of α-Mangostin-ZnO NPs significantly outperformed that of amoxicillin against all tested pathogens (Fig 6).

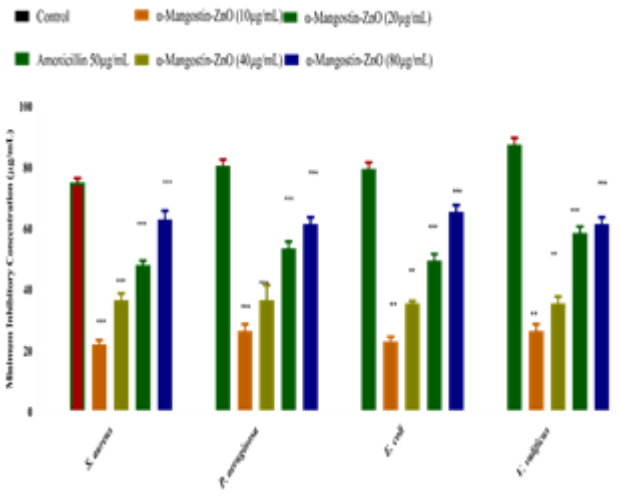


Fig. 5: MIC of α-Mangostin-ZnO NPs at different concentration against dental pathogens of S. aureus, P. aeruginosa, E. coli, and V. vulnificus. Amoxicillin was used as positive control. The *** represented the level of significance ($p < 0.05$) when the results were compared to the control. Data were presented as mean ± SD of three independent experiments.

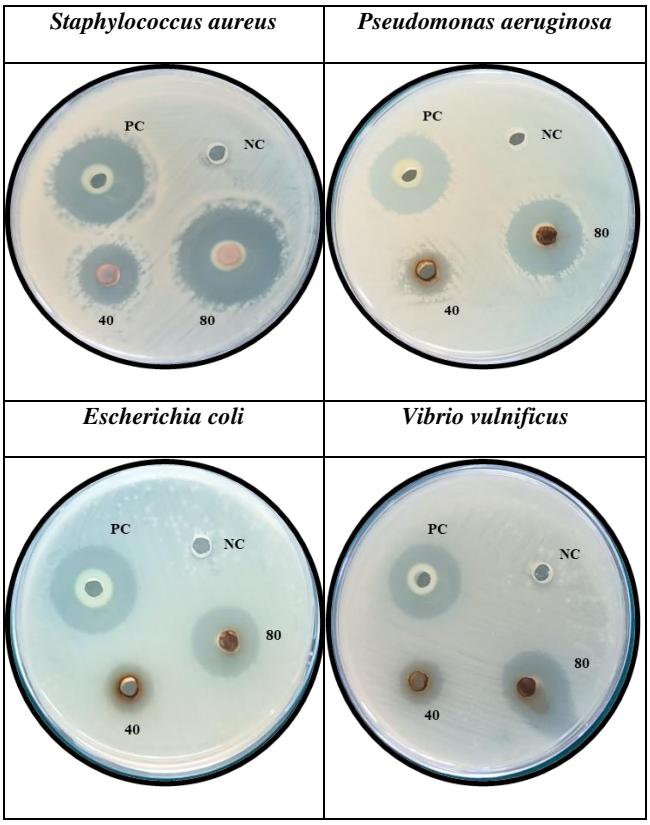


Fig.6: Zone of inhibition of different dental pathogen by α-Mangostin-ZnO NPs (A). A – Control, B – Amoxicillin (50 µg/mL), at 40 µg/mL, and D - α-Mangostin-ZnO NPs at 80 µg/mL. (B) Zone of inhibition measured at mm

3.4. Binding interaction with Pathogenic receptor

Mangostin exhibits varying binding affinities with different bacterial receptors. In the S. aureus Sortase A-substrate complex (PDB: 7S53), it binds with an affinity of -5.2 kcal/mol, interacting with amino acids such as GLU, LEU, ALA, VAL, LYS, and MET, which likely stabilizes the ligand and inhibits Sortase A activity. In P. aeruginosa LasR (PDB: 4NG2), mangostin has a binding affinity of -4.6 kcal/mol, with interactions involving PHE, TYR, GLU, and GLY, suggesting potential π-π stacking that disrupts LasR's signalling pathway (Table 2). For E. coli's outer membrane porin (PDB: 2OMF), mangostin shows a strong affinity of -5.5 kcal/mol, interacting with a diverse set of amino acids, including PRO, ALA, GLY, HIS, LEU, and LYS, indicating it may affect membrane permeability or nutrient uptake. Lastly, in the V. vulnificus nuclear protein (PDB: 5TV2), mangostin binds with an affinity of -4.4 kcal/mol, interacting with LYS, ARG, and MET, which could suggest a role in modulating nuclear functions (Fig 7).

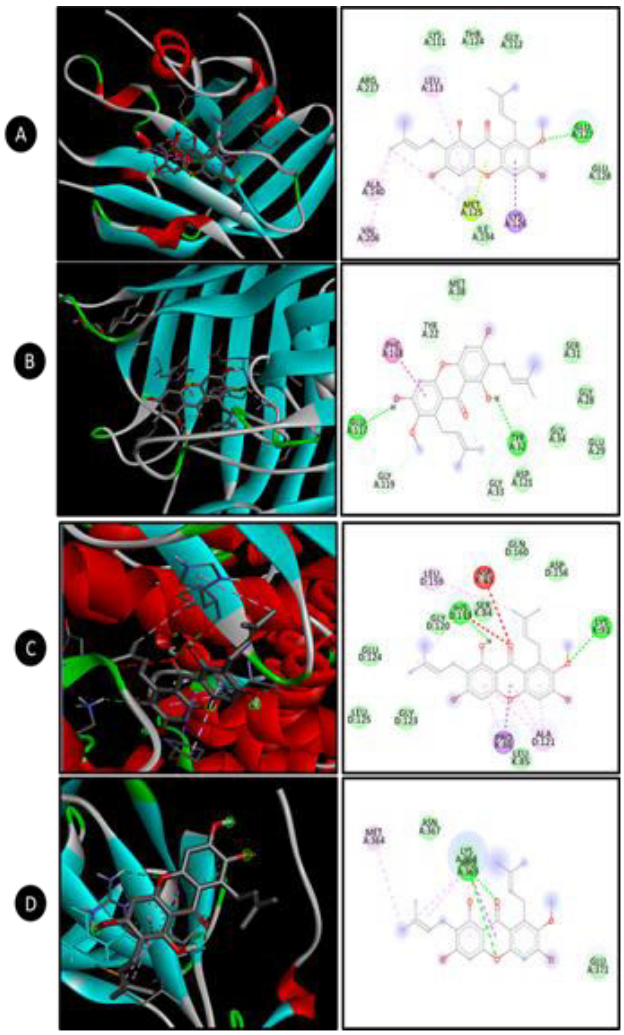


Fig. 7: 3D and 2D interaction images of Mangostin and wound-infecting bacterial pathogens. (A) 7S53 (antimicrobial protein), (B) 4NG2 (oxidative stress regulator), (C) 20MF (inflammatory mediator), (D) 5TV2 (wound healing protein). The interaction is observed with amino acids such as GLU (Glutamic acid), LEU (Leucine), ALA (Alanine), VAL (Valine), LYS (Lysine), MET (methionine), PHE (Phenylalanine), TYR (Tyrosin), GLY (glycine), PRO (proline), HIS (Histidine), ARG (Arginine).

Table 2: Amino acid and Binding affinity value of the vanillic acid interaction with different dental pathogen receptors as GLU (Glutamic acid), LEU (Leucine), ALA (Alanine), VAL (Valine), LYS (Lysine), MET (methionine), PHE (Phenylalanine), TYR (Tyrosin), GLY (glycine), PRO (proline), HIS (Histidine), ARG (Arginine).

	Receptor (PDB ID)	**Ligand**	**Binding affinity (kcal/mol)**	**Amino acid interaction**
1	*S. aureus*: Sortase A-substrate Complex (SrtA) (PDB: 7S53)	Mangostin	-5.2	GLU, LEU, ALA, VAL, LYS, MET
2	*P. aeruginosa:* LasR (PDB: 4NG2)	Mangostin	-4.6	PHE, TYR, GLU, GLY
3	*E. coli:* Outer membrane porin (integral membrane protein porin) (PDB: 2OMF)	Mangostin	-5.5	PRO, ALA, GLY, HIS, LEU, LYS
4	*V. vulnificus:* Nuclear Protein (PDB: 5TV2)	Mangostin	-4.4	LYS, ARG, MET

3.5 Survival rate analysis in zebrafish

The study examined the toxicity of α-Mangostin-ZnO NPs on zebrafish, comparing four groups: a control group (unwounded fish), wounded zebrafish, and zebrafish treated with varying concentrations of α-Mangostin-ZnO NPs. The results indicated a 69% decrease in survival rate among the wounded zebrafish, highlighting the potential adverse effects of injury. Notably, zebrafish exposed to α-Mangostin-ZnO NPs at concentrations of 40 µg/mL and 80 µg/mL showed no significant toxicity compared to the control group. This suggests that within the tested concentrations, α-Mangostin-ZnO NPs do not pose considerable toxicity to zebrafish (Fig 8).

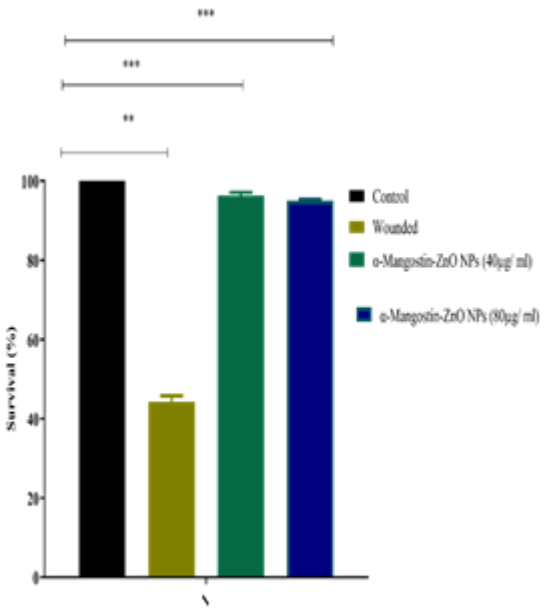


Fig.8: The survival rate analysis of zebrafish with different groups of control, wounded zebrafish, and exposure to α-Mangostin-ZnO NPs (40 µg/mL and 80 µg/mL). The independent three experimental data were expressed as mean ± SD (n=6/group). The significant difference at $p < 0.05$ was expressed as *

3.6. Wound healing activity

The wound healing activity of α-Mangostin-ZnO NPs was investigated on wounded zebrafish, focusing on concentrations of 40 µg/mL and 80 µg/mL over periods of 7 and 14 days post-wounding (dpw). Notably, treatment with α-Mangostin-ZnO NPs at 40 µg/mL resulted in a 54% closure of wounds at 7 dpw, while at 80 µg/mL, the closure significantly increased to 62% ($p < 0.05$). Remarkably, by 14 dpw, the closure rates further improved to 70% and 93% for the 40 µg/mL and 80 µg/mL concentrations, respectively, compared to the control group. These findings underscore the potential efficacy of α-Mangostin-ZnO NPs in promoting accelerated wound closure in zebrafish models, suggesting their promise for further exploration in wound healing applications (Fig 9).

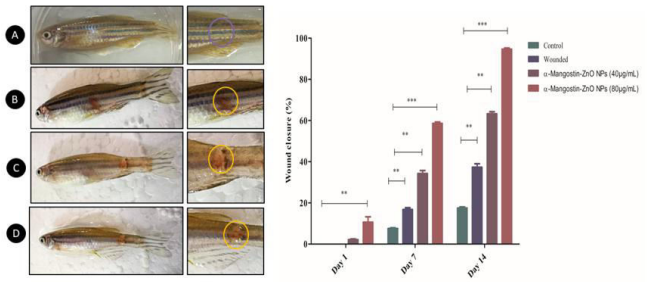


Fig. 9. (A – D) Representative photograph of wounded zebrafish treated with α-Mangostin-ZnO NPs at different concentrations. (E) The graphical presentation of the wound closure was measured using Image J software. The independent three experimental data were expressed as mean ± SD (n=6/group). The significant difference at $p < 0.05$ was expressed as **

3.7. Wound morphological analysis

Histological analysis of wounded zebrafish tissue revealed distinct differences across various groups: (A) Control (unwounded zebrafish), (B) Wounded Zebrafish, (C) α-Mangostin-ZnO NPs treatment at 40 µg/mL, and (D) α-Mangostin-ZnO NPs treatment at 80 µg/mL. The tissue layers examined included skeletal muscle, dermal layer, granulation tissue, and epithelial layer with immune cell infiltration. Microscopic examination of zebrafish treated with α-Mangostin-ZnO NPs, compared to the untreated control group, revealed notable differences in tissue composition. The treated groups exhibited a densely layered epidermis, significant proliferation of granulation tissue, and clear formation of new blood vessels (neoangiogenesis). Remarkably, the treated zebrafish displayed rapid development of a thin layer of new epidermis just hours after nanoparticle application. Additionally, the presence of abundant dispersed fibroblasts within the pronounced granulation tissue, along with an influx of inflammatory cells, indicated an active healing process facilitated by the nanoparticles. This was further supported by the emergence of newly formed vasculature, signifying an accelerated healing response. The enhanced wound healing induced by α-Mangostin-ZnO NPs was evident in the treated group's significantly quicker wound contraction and the formation of fresh scales within an impressive timeframe of 4 to 6 days post-wound creation, contrasting sharply with the slower healing observed in the untreated control group (Fig 10).

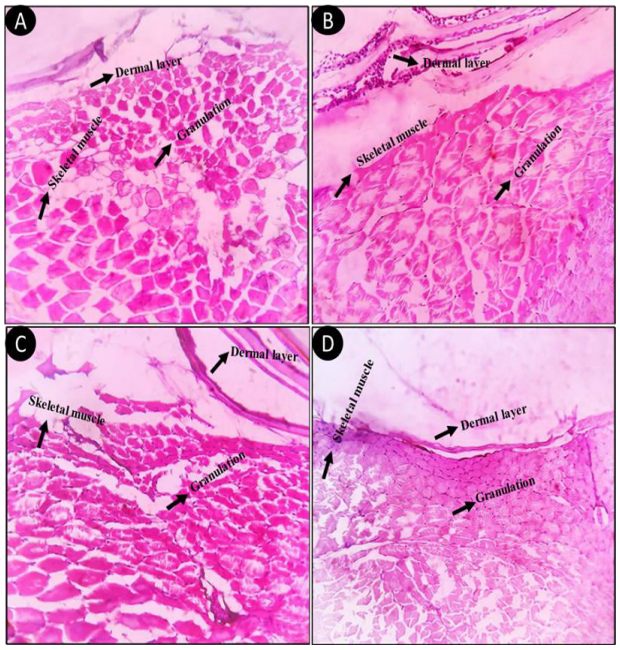


Fig. 10. Histological analysis of wounded zebrafish tissue. (A) Control (unwounded zebrafish), (B) Wounded Zebrafish, (C) α-Mangostin-ZnO NPs treatment at 40 µg/mL, and (D) α-Mangostin-ZnO NPs treatment at 80 µg/mL. The tissue

layers include skeletal muscle, dermal layer, granulation tissue, and epithelial layer with immune cell infiltration.

3.8 Expression level of wound healing related genes

To understand the molecular mechanism of α-Mangostin-ZnO NPs treatment in wound healing, the gene expression levels of matrix metalloproteinase 13 (mmp13) and pro-inflammatory cytokines such as tnf-α and il-1β were evaluated. The higher concentration of α-Mangostin-ZnO NPs at 80 µg/mL was selected for further gene expression studies due to its observed superior activity and lower toxicity. The results showed that the wound-administered group demonstrated a significant upregulation in the gene expression of mmp13 (3.5-fold), tnf-α (5.7-fold), and il-1β (4.6-fold). In contrast, the expression levels of mmp13 (1.7-fold), tnf-α (3.4-fold), and il-1β (2.3-fold) were significantly downregulated in the α-Mangostin-ZnO NPs treated group (80 µg/mL) compared to the control group (Fig 11).

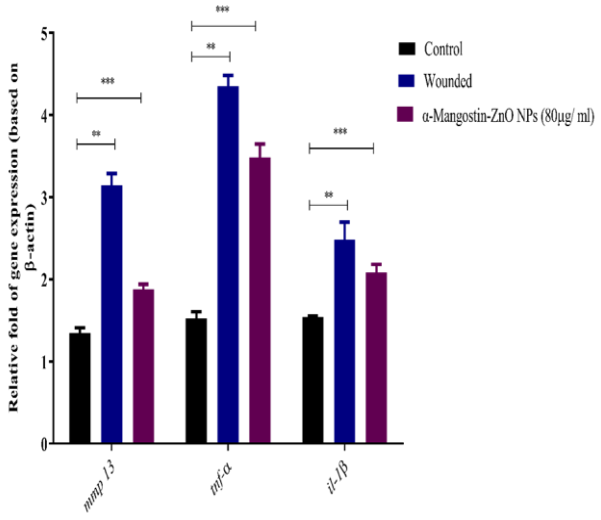


Fig.11. Effect of α-Mangostin-ZnO NPs treatment group on the mRNA expression level of mmp13, tnf-α, il-1β. Data were expressed as mean + SD of three independent experiments. * $p < 0.05$ as compared to the control.

DISCUSSION

The emergence of α-Mangostin-ZnO NPs as a novel therapeutic strategy for wound healing highlights the innovative combination of ZnO NPs with α-Mangostin, utilizing their distinct biological activities to tackle the complex challenges of wound management. ZnO NPs are well-known for their potent antibacterial properties, crucial for preventing infections, which are a significant barrier to healing in chronic wounds[16]. Their effectiveness in reducing bacterial load is enhanced by anti-inflammatory effects that help regulate the inflammatory response, thereby minimizing delays in healing caused by excessive inflammation. Additionally, their antioxidant capabilities are vital in countering oxidative stress, which can negatively impact cellular health and hinder recovery. When paired with α-Mangostin, a natural polyphenolic compound recognized for its wide range of biological activities, the wound healing potential is further amplified. α-Mangostin not only provides antioxidative, antimicrobial, and anti-inflammatory benefits but also effectively scavenges free radicals, reducing oxidative damage. This synergy between ZnO NPs and α-Mangostin presents a promising path for advancing wound healing therapies, addressing both infection control and tissue regeneration[17].

Characterization of α-Mangostin-ZnO NPs reveals critical structural and chemical properties essential for potential applications in wound healing. In this study, the promising antioxidant and anti-inflammatory therapeutic potential of α-Mangostin-ZnO NPs in wound healing can be attributed to their ability to stabilize proteins and scavenge harmful oxidative agents. By preventing protein denaturation, these NPs help maintain cellular integrity in the wound environment, reducing inflammation in a manner similar to established anti-inflammatories like diclofenac. Their significant antioxidant activity also alleviates oxidative stress—a key barrier to healing—by neutralizing reactive oxygen species (ROS) that can damage tissue at the wound site. These functionalities not only accelerate the healing process but also enhance tissue repair, corroborated by previous studies indicating that Aerva persica-mediated synthesis of ZnO NPs demonstrated a dual therapeutic effect of antioxidant and anti-inflammatory properties[18].

The antimicrobial activity of α-Mangostin-ZnO NPs is crucial for effective wound healing, as it directly combats infection, one of the primary complications that can impede the healing process. Infections can delay healing by causing inflammation, increasing necrotic tissue burden, and disrupting new tissue formation. Effective antimicrobial agents speed up the healing process by maintaining a cleaner wound environment while minimizing the potential for developing antibiotic resistance, a significant concern with traditional antimicrobial treatments. Recent studies have shown that electrospun membranes loaded with α-Mangostin-ZnO NPs exhibited antibacterial activity against E. coli and S. aureus. Consistent with previous findings, our study demonstrated that the minimum inhibitory concentration (MIC) and zone of inhibition assays highlighted the broad-spectrum antimicrobial activity of α-Mangostin-ZnO NPs, suggesting their significant utility in wound healing applications. This antimicrobial efficacy is critical in preventing and controlling infections at wound sites, particularly against common pathogens such as S. aureus, P. aeruginosa, E. coli, and V. vulnificus. The ability of α-Mangostin-ZnO NPs to inhibit the growth of both Gram-positive and Gram-negative bacteria makes them excellent candidates for topical wound dressings and coatings that require robust antimicrobial properties[19].

In the context of wound infections, the Sortase A-substrate complex (SrtA) in S. aureus is significant because it facilitates the attachment of surface proteins, including adhesins and toxins, to the bacterial cell wall. This attachment is essential for S. aureus to adhere to and invade host tissues, initiating infection. Adhesins enable bacteria to adhere to damaged tissues and host cells at the wound site, promoting colonization and biofilm formation, which can protect bacteria from host defenses and antibiotics. Additionally, toxins anchored to the cell wall by SrtA can damage host cells and tissues, contributing to the pathogenesis of S. aureus infections[20].

One of the key virulence factors of P. aeruginosa in wound infections is the LasR protein, a transcriptional

regulator controlling the expression of numerous virulence genes. P. aeruginosa is adept at forming biofilms, complex communities of bacteria encased in a matrix of extracellular polymeric substances. LasR regulates the expression of genes involved in biofilm formation, promoting attachment to wound surfaces and protecting bacteria from host immune defenses and antibiotics. In E. coli, the outer membrane porin OmpC contributes to its virulence in wound infections by allowing the passage of small molecules, including nutrients, toxins, and antibiotics, into the bacterial cell. This can enhance the pathogenicity of E. coli in wound infections, facilitating the uptake of toxic substances produced by the host or promoting antibiotic resistance[21].

In V. vulnificus wound infections, the protein EpsD is involved in the biosynthesis of extracellular polysaccharides (EPS), a crucial component of the bacterial biofilm matrix. This protein enables biofilm formation on both biotic and abiotic surfaces, including wounds. Biofilms are complex communities of bacteria encased in a self-produced matrix of EPS, proteins, and DNA, which provides structural support and protects bacteria from environmental stresses, including host immune responses and antibiotic treatments. EpsD is essential for EPS synthesis and secretion, contributing to the stability of V. vulnificus biofilms on wound surfaces[22].

In this study, α-Mangostin, which demonstrated significant binding affinity to essential bacterial proteins like Sortase A in S. aureus, LasR in P. aeruginosa, OmpC in E. coli, and EpsD in V. vulnificus, offers promising prospects for clinical applications in preventing wound infections. These compounds could serve as lead molecules for novel antimicrobial therapeutics, complementing existing treatments or being formulated into topical preparations and wound dressings for direct application. Through their ability to disrupt key virulence factors and inhibit biofilm formation, α-Mangostin-ZnO NPs hold potential for enhancing wound healing and reducing infection risk, particularly in high-risk populations or clinical settings where infection control is critical.

Wound healing is essential for restoring the structural and functional integrity of tissues, promoting overall patient health and well-being. Proper wound closure reduces the risk of infection, prevents further tissue damage, and facilitates the restoration of normal physiological function, allowing patients to resume their daily activities more quickly. Previous studies have explored the wound healing activity of nano emulsions with Elaeis guineensis leaf extract in zebrafish models, observing effective wound closure after treatment. Similarly, our study's results regarding the toxicity analysis and wound healing activity of α-Mangostin-ZnO NPs in zebrafish models present significant implications for potential applications in wound management. The absence of discernible toxicity in zebrafish, even at concentrations as high as 80 μg/mL, underscores the favourable safety profile of α-Mangostin-ZnO NPs. This finding is particularly promising, suggesting that these nanoparticles may be well-tolerated in biological systems, a crucial consideration for any therapeutic intervention[23].

Furthermore, the remarkable wound healing activity observed with α-Mangostin-ZnO NPs highlights their potential as effective wound healing agents. Concentrations of 40 μg/mL and 80 μg/mL demonstrated accelerated wound closure compared to the control group, with the higher concentration showing particularly notable results, achieving a closure rate of 93% by 14 days post-wounding. Additionally, tissue repair during the wound healing process was confirmed through histological examination, revealing the development of skeletal muscle, dermal layer, granulation tissue, and epithelial layer with immune cell infiltration. This accelerated wound closure suggests that α-Mangostin-ZnO NPs promote tissue regeneration and repair, potentially enhancing the overall wound healing process. These findings open exciting possibilities for developing α-Mangostin-ZnO NPs-based wound dressings or topical formulations for clinical use.

In the context of wound healing, mmp13 and pro-inflammatory cytokines tnf-α and il-1β play significant roles. Mmp13 facilitates tissue remodelling by degrading extracellular matrix components, while pro-inflammatory cytokines like tnf-α and il-1β initiate and amplify the inflammatory response crucial for wound healing. Mmp13 activity promotes the breakdown of damaged tissue and the migration of inflammatory cells to the wound site, aiding in tissue repair. Tnf-α and il-1β induce the expression of adhesion molecules and chemokines, recruiting leukocytes and amplifying the inflammatory cascade. Moreover, il-1β stimulates fibroblast proliferation and collagen synthesis, promoting tissue repair. Together, these molecules orchestrate various phases of wound healing, from inflammation to tissue remodelling[24].

According to earlier reports, downregulation of mmp13, tnf-α, and il-1β was observed during the wound healing process in zebrafish models. In line with this, our study's upregulation of these genes in the wound-administered group indicates an active inflammatory phase necessary for initiating the healing process. However, the downregulation of these genes in the α-Mangostin-ZnO NPs treated group suggests modulation of the inflammatory response, potentially leading to a more controlled and efficient healing process. These findings highlight the therapeutic potential of α-Mangostin-ZnO NPs in wound management by regulating inflammatory signalling pathways and promoting tissue repair.

CONCLUSION

In conclusion, the promising findings regarding α-Mangostin-ZnO NPs highlight their potential as effective agents in wound healing through their antimicrobial, anti-inflammatory, and antioxidant properties. The observed mechanisms suggest that these nanoparticles can modulate inflammatory responses while promoting tissue regeneration, paving the way for innovative therapeutic approaches. Future research should focus on optimizing formulations for clinical applications, exploring their synergistic effects with other therapeutic agents, and conducting in vivo studies to further assess safety and efficacy. Additionally, investigating the mechanisms underlying their interactions with bacterial virulence factors will be crucial for developing targeted treatments. Overall, α-Mangostin-ZnO NPs hold significant promise for enhancing wound care strategies in both clinical and community settings.

ACKNOWLEDGEMENT

I would like to extend my sincere appreciation to Saveetha Dental and Medical College, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai, India for their invaluable support and resources that made this research possible. Their commitment to excellence in

education and research has significantly contributed to the advancement of knowledge in our field.

CRediT authorship contribution statement

Meenakshi Sundaram Kishore Kumar: Conceptualization, Supervision, Methodology, Formal analysis, Writing – original draft, Writing – review & editing.Taniya Mary Martin: Investigation, Methodology, Data curation, Writing – review & editing. Selvi Rajavel: Investigation, Methodology, Data curation, Visualization.

REFERENCES

[1] Mamun AA, Shao C, Geng P, Wang S, Xiao J. Recent advances in molecular mechanisms of skin wound healing and its treatments. Frontiers in immunology. 2024 May 21;15:1395479.

https://doi.org/10.3389/fimmu.2024.1395479

[2] Garza-Cerna JA, Mecott-Rivera GÁ, Coronado EP, Castro-Govea Y, Pérez-Trujillo JJ, de Oca-Luna RM, Salas-Treviño D, González-Mateo GT, Kopytina V, Berroa JL. Therapeutic Advances in Chronic Wound Management: A Review of Debridement Strategies and Emerging Approaches. Wound Healing-Mechanisms and Pharmacological Interventions. 2025 Sep 25.

DOI: 10.5772/intechopen.1011088

[3] Mansur S, Zainol Abidin MN. Development of α-mangostin-loaded silica nanoparticles with enhanced protein-bound uremic toxin adsorption and biocompatibility. Journal of Porous Materials. 2025 Dec;32(6):2231-43.

https://doi.org/10.1007/s10934-025-01843-5

[4] Szlachcikowska D, Mazurek K, Magiera M, Jama G, Tabęcka-Łonczyńska A. Current Insights and Future Directions in Scar Management and Skin Regeneration. International Journal of Molecular Sciences. 2025 Oct 31;26(21):10636.

doi: 10.3390/ijms262110636

[5] Majdalawieh AF, Khatib BK, Terro TM. α-Mangostin Is a Xanthone Derivative from Mangosteen with Potent Immunomodulatory and Anti-Inflammatory Properties. Biomolecules. 2025 May 7;15(5):681.

doi: 10.3390/biom15050681

[6] Sureendar B, Chellappa LR. In Vitro Evaluation of Antifungal Activity of a Dental Varnish Formulated with Ginger and Rosemary Mediated Titanium Oxide Nanoparticles. Journal of Pioneering Medical Sciences. 2025 Apr 27;14:278-82.

DOI : https://doi.org/10.47310/jpms202514S0135

[7] Senthil R. Bevacizumab-conjugated curcumin nanoparticles promote cytotoxicity and apoptosis in human malignant oral keratinocytes in vitro. Journal of Oral and Maxillofacial Surgery. 2025 May 27.

https://doi.org/10.1016/j.joms.2025.05.014

[8] Pan-In P, Wongsomboon A, Kokpol C, Chaichanawongsaroj N, Wanichwecharungruang S. Depositing α-mangostin nanoparticles to sebaceous gland area for acne treatment. Journal of pharmacological sciences. 2015 Dec 1;129(4):226-32.

https://doi.org/10.1016/j.jphs.2015.11.005

[9] Varshasree S, Sinduja P, Manjesh D, Sharma S, Kalra IK, Karunakaran JV. Effectiveness of Piper longum Extract as Natural Irrigant on Antimicrobial Properties and Smear Layer Removal: A Scanning Electron Microscope Study. The Journal of Contemporary Dental Practice. 2025 Nov 25;26(11):1073-8.

DOI: 10.5005/jp-journals-10024-3875

[10] Anbarasu M, Martin TM, Priya P, Sivamurugan V, Kumar MS, Shaik MR, Kari ZA, Guru A. Assessing the impact of Ag-ZnO nanoparticle on the induction of oxidative stress, hematological, and molecular changes in zebrafish (Danio rerio) and McCoy fibroblast cell lines. Aquaculture International. 2024 Aug;32(4):5373-92.

https://doi.org/10.1007/s10499-024-01611-3

[11] Nguyen PT, Nguyen MT, Bolhuis A. Inhibition of biofilm formation by alpha-mangostin loaded nanoparticles against Staphylococcus aureus. Saudi Journal of Biological Sciences. 2021 Mar 1;28(3):1615-21.

https://doi.org/10.1016/j.sjbs.2020.11.061

[12] Gopika GG, Nagesh S, Mani SP, Pitchaipillai S. Assessment of antimicrobial properties of silver-doped hydroxyapatite nanoparticle-modified orthodontic composite: an in-vitro study. Journal of Stomatology.;78(1).

DOI: https://doi.org/10.5114/jos.2025.155774

[13] Sivaperumal P, Ganapathy D, Kamala K. Evaluating the efficacy of doripenem against Staphylococcus aureus in vancomycin-resistant strains. Microbial Pathogenesis. 2025 May 1;202:107449.

https://doi.org/10.1016/j.micpath.2025.107449

[14] Ramachandran T, Mohanraj KG, Martin TM. Enhanced wound healing with β-chitosan-zinc oxide nanoparticles: insights from zebrafish models. Cureus. 2024 Sep 21;16(9).

DOI: 10.7759/cureus.69861

[15] Dandagi P, Martin TM, Babu Y. In silico and glioblastoma cell line evaluation of thioflavin-derived zinc nanoparticles targeting beclin protein. Cureus. 2024 Sep 13;16(9).

DOI: 10.7759/cureus.69319

[16] Chan YB, Aminuzzaman M, Chuah XT, Li K, Balu P, Wong LS, Guha SK, Tey LH. Review in green synthesis mechanisms, application, and future prospects for Garcinia mangostana L.(mangosteen)-derived nanoparticles. Nanotechnology Reviews. 2025 Apr 28;14(1):20250157.

https://doi.org/10.1515/ntrev-2025-0157

[17] Arun Dharsaun SR, Francis AP, Gayathri R, Veeraraghavan VP, Sankaran K. Synthesis of Hemidesmus indicus-mediated biogenic titanium oxide nanoparticles for potential applications in dental adhesive. Nano Life. 2025 Jun 30;15(03):2450021.

https://doi.org/10.1142/S1793984424500211

[18] Fatima K, Asif M, Farooq U, Gilani SJ, Bin Jumah MN, Ahmed MM. Antioxidant and anti-inflammatory applications of Aerva persica aqueous-root extract-mediated synthesis of ZnO nanoparticles. ACS omega. 2024 Mar 29;9(14):15882-92.

https://doi.org/10.1021/acsomega.3c08143

[19] Górecka H, Guźniczak M, Buzalewicz I, Ulatowska-Jarża A, Korzekwa K, Kaczorowska A. Alpha-Mangostin: A Review of Current Research on Its Potential as a Novel Antimicrobial and Anti-Biofilm Agent. International Journal of Molecular Sciences. 2025 May 30;26(11):5281.

https://doi.org/10.3390/ijms26115281

[20] Kruger RG, Otvos B, Frankel BA, Bentley M, Dostal P, McCafferty DG. Analysis of the substrate specificity of the Staphylococcus aureus sortase transpeptidase SrtA. Biochemistry. 2004 Feb 17;43(6):1541-51.

https://doi.org/10.1021/bi035920j

[21] Elnegery AA, Mowafy WK, Zahra TA, Abou El-Khier NT. Study of quorum-sensing LasR and RhlR genes and their dependent virulence factors in Pseudomonas aeruginosa isolates from infected burn wounds. Access Microbiology. 2021 Mar 5:000211.

https://doi.org/10.1099/acmi.0.000211

[22] Limoli DH, Jones CJ, Wozniak DJ. Bacterial extracellular polysaccharides in biofilm formation and function. Microbial Biofilms. 2015 Oct 7:223-47.

https://doi.org/10.1128/9781555817466.ch11

[23] Bajhal S, Seenivasan B, Arockiaraj J, Sundaramurthy A. Antimicrobial Potentials of Haliclona fibulata Derived Compounds for Combatting Drug‐Resistant Bacterial Infections in Zebrafish Model. Chemistry & Biodiversity. 2025 May;22(5):e202402717.

https://doi.org/10.1002/cbdv.202402717

[24] Chidambaram D, Subashini V, Nanthanalaxmi M, Saranya I, Selvamurugan N. Regulation of matrix metalloproteinase-13 in cancer: Signaling pathways and non-coding RNAs in tumor progression and therapeutic targeting. World Journal of Clinical Oncology. 2025 Jun 24;16(6):105996.

doi: 10.5306/wjco.v16.i6.105996

Determining the Dominant Social Media Platform Using Hexagraph Topology and Matrix Representation

S. Santhiya
Assistant Professor
Department of Science and Humanities
Easa College of Engineering and Technology
Tamil Nadu, India
santhiyamrs26@gmail.com

***Abstract*—The application of hexagraph topology combined with matrix representation provides a systematic framework for analyzing social media platforms. This approach enables researchers to study platform usage patterns, relationships, and user interaction behaviors in a structured manner. By modeling social media networks using hexagraph structures and representing their connections through matrices, complex interactions between platforms can be analyzed more effectively. Applying this methodology across different social media platforms allows researchers to compare levels of connectivity, engagement, and interaction patterns. Such analysis helps in identifying platforms that are more interconnected, exhibit higher engagement rates, or demonstrate unique usage behaviors. The insights derived from this approach are valuable not only for academic research but also for practical applications in social media marketing, digital strategy development and platform management.**

***Index Terms*—Social Media, Hexa Graph Topology, Matrix Representation**

I. Introduction and Preliminaries

Nano graph topology refers to a network structure where each node has a very small number of connections relative to the size of the network. Unlike traditional graph topologies that may have nodes with numerous connections (like social networks), nano graph topology typically involves nodes with only a handful of connections, often two or three, sometimes referred to as low-degree networks.

Using graph theory and matrix representations allows us to analyze complex networks like social media platforms effectively. By constructing and analyzing the incidence matrix and computing centrality metrics, we can determine which platform is the leading one based on the interactions and connectivity observed in the collected data from 2024.

Working on this paper that aims to identify the most prominent social network using hexa graph topology and an incidence matrix representation.

Definition 1.1 (14). Let G be a undirected graph with six disjoint graph decomposition $J_h(s)$ where $h = 1, 2, 3, 4, 5, 6$ of G. Consider $S_i, i = 1, 2, 3, 4, 5, 6$ be any six subgraphs of G, then we define a lower hexa subgraph and upper hexa subgraph as $H_-(S_i) = \bigcup_{s\in G}\{s : J_1(s) \subseteq S_1 \bigvee J_2(s) \subseteq S_2 \bigvee J_3(s) \subseteq S_3 \bigvee J_4(s) \subseteq S_4 \bigvee J_5(s) \subseteq S_5 \bigvee J_6(s) \subseteq S_6\}$, $H^-(S_i) = \bigcup_{s\in G}\{s : J_1(s) \cap S_1 \neq \phi \bigwedge J_2(s) \cap S_2 \neq \phi \bigwedge J_3(s) \cap S_3 \neq \phi \bigwedge J_4(s) \cap S_4 \neq \phi \bigwedge J_5(s) \cap S_5 \neq \phi \bigwedge J_6(s) \cap S_6 \neq \phi\}$.The boundary hexa subgraph is defined as $B_N(S) = H^-(S_i) - H_-(S_i)$.

Definition 1.2 (14). Let G be the non-empty undirected graph and $J_h(s), h = 1, 2, 3, 4, 5, 6$ be any six disjoint graph decomposition of G.Then for any six subgraphs $S_i, i = 1, 2, 3, 4, 5, 6$ of G and $\zeta_H(S_i) = \{G, \phi, H_-(S_i),$
$H^-(S_i), B_h(S_i)\}$. Let $S_i \subseteq G, \zeta_H(S_i)$ satisfies the following axioms

1) $G, V_o \in \zeta_H(S_i)$, where G = full graph with edge set, V_o = Null graph
2) Any union of elements of $\zeta_H(S_i)$ is in $\zeta_H(S_i)$
3) The finite intersection of the elements of $\zeta_H(S_i)$ is in $\zeta_H(S_i)$.

Then a pair $\zeta_H(S_i)$ is known as Hexa graph topology on G with regard to a subgraph S_i of G.We call the pair $(G, \zeta_H(S_i))$ as the Hexa graph topological space on G.The elements of Hexa graph topological spaces are regard as Hexa open subgraph of G and the complement of the Hexa open subgraph of G is called an Hexa closed subgraph of G.

Definition 1.3 (14). Let G be a non-empty undirected graph and for a six subgraph S_i of G. If $\zeta_H(S_i)$ is the Hexa Graph topology on G concerning to S_i, then $B = \{G, \mathrm{H}_-(S_i), \mathrm{B}_h(S_i)\}$ is the basis for $\zeta_H(S_i)$.

II. Social Media Platform

Social media has revolutionized online communication, enabling individuals to create and share content at an unprecedented rate. This transformation reshapes how people interact and disseminate information due to its ease of use, speed, and extensive reach.

Types of Social Media Platforms include:

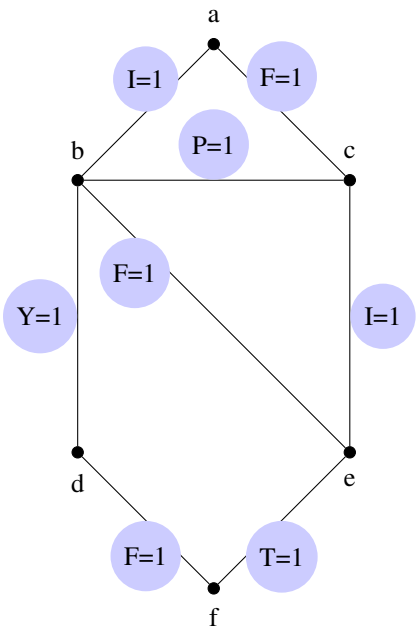
- Social Networking: Facebook, Twitter, LinkedIn, Instagram
- Media Sharing: Instagram, TikTok, Snapchat
- Microblogging: Twitter
- Discussion Forums: Reddit
- Video Sharing: YouTube, Vimeo, TikTok

III. Representation of Hexa Graph Topology with Incidence Matrix

Let G be a weighted graph of social network with vertex of persons are $\{a, b, c, d, e, f\}$ and edges of social networks are F- stands for facebook ,I-stands for Instagram ,T-stands for twitter,Y-stands for youtube, PI- stands for P-interest.where the vertices represent persons and the edges represent different types of social network connections (Facebook, Instagram, Twitter, YouTube, Pinterest).
The following table shows that the sample of 6 persons and their respective social network connections.

persons	linked with
$\{a\}$	$\{F, I\}$
{b}	$\{F, I, Y\}$
{c}	$\{F, I, P\}$
{d}	$\{F, Y\}$
{e}	$\{F, I, T\}$
{f}	$\{F, T\}$

Figure 2: Social network analysis weighted graph



Consider the above social network analysis weighted graph G with there are 5 persons (vertices) and 6 different social networks (Facebook, Instagram, Twitter, YouTube, Twitter,P-Interest)

Then its incidence matrix M is

$$\text{M= } [bij] = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

(assuming 1 indicates a connection and 0 indicates no connection). Let's denote the 6 persons as $\{a, b, c, d, e, f\}$. And denote the 5 social networks as $\{F, I, Y, T, P\}$. The incidence matrix M will have 6 rows (one for each person) and 5 columns (one for each social network).

- The first row corresponds to person (a). The 1's in the first and second columns indicate that person (a) is connected to Facebook (F) and Instagram (I) respectively. The 0's in the remaining columns, indicate that person (a) is not connected to Youtube(Y),Twitter(T) and P-Interest (P).
- The second row corresponds to person (b). The 1's in the first, second,third,fifth columns indicate that person (b) is connected to Facebook (F) ,Instagram (I) , youtube (Y) and P-Interest (P) respectively. The 0's in the remaining columns, indicate that person (b) is not connected to Twitter(T).
- The third row corresponds to person (c). The 1's in the first,second and fifth columns indicate that person (c) is connected to Facebook (F), Instagram (I) and P-Interest(P) respectively. The 0's in the remaining columns, indicate that person (c) is not connected to Youtube(Y),Twitter(T).
- The fourth row corresponds to person (d). The 1's in the first and third columns indicate that person (d) is connected to Facebook (F) and youtube(Y) respectively. The 0's in the remaining columns, indicate that person (d) is not connected to Instagram(I),Twitter(T) and P-Interest (P).
- The fifth row corresponds to person (e). The 1's in the first,second and fourth columns indicate that person (e) is connected to Facebook (F),Instagram (I) and Twitter(T) respectively. The 0's in the remaining columns, indicate that person (e) is not connected to Youtube(Y) and P-Interest (P).
- The sixth row corresponds to person (f). The 1's in the first and fourth columns indicate that person (f) is connected to Facebook (F) and Twitter(T) respectively. The 0's in the remaining columns, indicate that person (f) is not connected to Instagram(I), Youtube(Y) and P-Interest (P).

N_i	$J_h(n), h = 1,2,3,4,5,6$	$H_-(N_i)$	$H^-(N_i)$
$\{F, I\}$	$\{\{F, I, Y\}, \{T, P\}\}$	ϕ	$\{F, I, Y\}$
$\{F, I, Y, P\}$	$\{\{F, Y\}, \{I, T\}, \{P\}\}$	$\{F, Y, P\}$	G
$\{F, I, P\}$	$\{\{F, T\}, \{I, Y\}, \{P\}\}$	$\{P\}$	G
$\{F, Y\}$	$\{\{F, I, T\}, \{Y, P\}\}$	ϕ	G
$\{F, I, T\}$	$\{\{F, I\}, \{Y, T\}, \{P\}\}$	$\{F, I\}$	$\{F, I, Y, T\}$
$\{F, T\}$	$\{\{F, Y, T\}, \{I, P\}\}$	ϕ	$\{F, Y, T\}$

Here $N_i, i = 1, 2, 3, 4, 5, 6$ be six persons with their linked social networks and $J_h(n), h = 1, 2, 3, 4, 5, 6$ be any six disjoint graph decomposition of G. Hence the Lower Hexa subgraph and Upper Hexa subgraph and Boundary hexa subgraph is $H_-(N_i) = \{F, Y, I, P\}$, $H^-(N_i) = \{F, Y\}$, $B_h(N_i) = \{I, P\}$ then
$\zeta_H(N_i) = \{G, \emptyset, \{F, Y\}, \{I, P\}, \{F, Y, I, P\}\}$ is a Hexa graph topology on G.

Identifying the leading social media platform through Hexa graph topology with

representation of Matrix

Case 1: Consider the above example, G be a weighted graph of social network with vertex of persons are $\{a, b, c, d, e, f\}$ and edges of social networks $\{F, I, Y, T, P\}$ and the Hexa graph topology on G is
$\zeta_H(N_i) = \{G, \emptyset, \{F, Y\}, \{I, P\}, \{F, Y, I, P\}\}$ and the basis is $\mathrm{B} = \{G, \{I, P\}, \{F, Y, I, P\}\}$...(1)

Case 1.1: When we remove Instagram from the incidence matrix then we will get the new incidence matrix is M= $[bij]$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

N_i	$J_h(n), h = 1,2,3,4,5,6$	$H_-(N_i)$	$H^-(N_i)$
$\{F\}$	$\{\{F, Y\}, \{T, P\}\}$	ϕ	$\{F, Y\}$
$\{F, Y, P\}$	$\{\{F, Y\}, \{T\}, \{P\}\}$	$\{F, Y, P\}$	$\{F, Y, P\}$
$\{F, P\}$	$\{\{F, T\}, \{Y\}, \{P\}\}$	$\{P\}$	$\{F, T, P\}$
$\{F, Y\}$	$\{\{F, T\}, \{Y, P\}\}$	ϕ	G
$\{F, T\}$	$\{\{F\}, \{Y, T\}, \{P\}\}$	$\{F\}$	$\{F, Y, T\}$
$\{F, T\}$	$\{\{F, Y, T\}, \{P\}\}$	ϕ	$\{F, Y, T\}$

Here $N_i, i = 1,2,3,4,5,6$ be six persons with their linked social network of G and $J_h(n), h = 1,2,3,4,5,6$ be any six disjoint graph decomposition of G. Hence the Lower Hexa subgraph and Upper Hexa subgraph and Boundary hexa subgraph is $H_-(N_i) = \{F, Y, P\}$, $H^-(N_i) = \{F\}$, $B_h(N_i) = \{Y, P\}$ then
$\zeta_H(N_i) = \{G, \emptyset, \{F\}, \{Y, P\}, \{F, Y, P\}\}$ is a Hexa graph topology on G and the basis is
$\mathrm{B} = \{G, \{Y, P\}, \{F, Y, P\}\}$.....(2). Here $(1) \neq (2)$

Case 1.2: When we remove Youtube from the incidence matrix then we will get the new incidence matrix is

$$\text{M= } [bij] = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

N_i	$J_h(n), h = 1,2,3,4,5,6$	$H_-(N_i)$	$H^-(N_i)$
$\{F, I\}$	$\{\{F, T\}, \{I, P\}\}$	ϕ	G
$\{F, I, P\}$	$\{\{F\}, \{I, T\}, \{P\}\}$	$\{F, P\}$	G
$\{F, I, P\}$	$\{\{F, T\}, \{I\}, \{P\}\}$	$\{I, P\}$	G
$\{F\}$	$\{\{F, I, T\}, \{P\}\}$	ϕ	$\{F, I, T\}$
$\{F, I, T\}$	$\{\{F, I\}, \{T\}, \{P\}\}$	$\{F, I, T\}$	$\{F, I, T\}$
$\{F, T\}$	$\{\{F, I\}, \{T, P\}\}$	ϕ	G

Hence the Lower Hexa subgraph and Upper Hexa subgraph and Boundary hexa subgraph is $H_-(N_i) = \{F, I, P, T\}$, $H^-(N_i) = \{F, I, T\}$, $B_h(N_i) = \{P\}$ then
$\zeta_H(N_i) = \{G, \emptyset, \{P\}, \{F, I, T\}, \{F, I, P, T\}\}$ is a Hexa graph topology on G and the basis
$\mathrm{B} = \{G, \{F, I, P, T\}, \{P\}\}$.....(3). Here $(1) \neq (3)$

Case 1.3: When we remove Twitter from the incidence matrix then we will get the new incidence matrix is

$$\text{M= } [bij] = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

N_i	$J_h(n), h = 1,2,3,4,5,6$	$H_-(N_i)$	$H^-(N_i)$
$\{F, I\}$	$\{\{F, I, Y\}, \{P\}\}$	ϕ	$\{F, I, Y\}$
$\{F, I, Y, P\}$	$\{\{F, Y\}, \{I\}, \{P\}\}$	$\{F, Y, I, P\}$	G
$\{F, I, P\}$	$\{\{F\}, \{I, Y\}, \{P\}\}$	$\{F, P\}$	G
$\{F, Y\}$	$\{\{F, I\}, \{Y, P\}\}$	ϕ	G
$\{F, I\}$	$\{\{F, P\}, \{I, Y\}\}$	ϕ	G
$\{F\}$	$\{\{F, I, P\}, \{Y\}\}$	ϕ	$\{F, I, P\}$

Hence the Lower Hexa subgraph and Upper Hexa subgraph and Boundary hexa subgraph is $H_-(N_i) = \{F, Y, I, P\}$, $H^-(N_i) = \{F, I\}$, $B_h(N_i) = \{Y, P\}$ then
$\zeta_H(N_i) = \{G, \emptyset, \{F, I\}, \{Y, P\}, \{F, Y, I, P\}\}$ is a Hexa graph topology on G and the basis
$\mathrm{B} = \{G, \{Y, P\}, \{F, Y, I, P\}\}$.....(4). Here $(1) \neq (4)$

Case 1.4: When we remove P-Interest from the incidence matrix then we will get the new incidence matrix is

$$\text{M= } [bij] = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

N_i	$J_h(n), h = 1,2,3,4,5,6$	$H_-(N_i)$	$H^-(N_i)$
$\{F, I\}$	$\{\{F, I, Y\}, \{T\}\}$	ϕ	$\{F, I, Y\}$
$\{F, I, Y\}$	$\{\{F, Y\}, \{I, T\}\}$	$\{F, Y\}$	G
$\{F, I\}$	$\{\{F, T\}, \{I, Y\}\}$	ϕ	G
$\{F, Y\}$	$\{\{F, I, T\}, \{Y\}\}$	$\{Y\}$	G
$\{F, I, T\}$	$\{\{F, I\}, \{Y, T\}\}$	$\{F, I\}$	G
$\{F, T\}$	$\{\{F, Y, T\}, \{I\}\}$	ϕ	$\{F, Y, T\}$

Hence the Lower Hexa subgraph and Upper Hexa subgraph and Boundary hexa subgraph is $H_-(N_i) = \{F, Y, I\}$, $H^-(N_i) = \{F, Y\}$, $B_h(N_i) = \{I\}$ then
$\zeta_H(N_i) = \{G, \emptyset, \{I\}, \{F, Y\}, \{F, Y, I\}\}$ is a Hexa graph topology on G and the basis
$\mathrm{B} = \{G, \{I\}, \{F, Y, I\}\}$.....(5).Here $(1) \neq (5)$

Case 1.5: When we remove Facebook we will get the Lower Hexa subgraph and Upper Hexa subgraph and Boundary hexa subgraph is $H_-(N_i) = \{F, Y, I, P\}$, $H^-(N_i) = \{F, Y\}$, $B_h(N_i) = \{I, P\}$ then $\zeta_H(N_i) = \{G, \emptyset, \{F, Y\}, \{I, P\}, \{F, Y, I, P\}\}$ is a Hexa graph topology on G and the basis $\mathrm{B} = \{G, \{I, P\}, \{F, Y, I, P\}\}$...(6). Here $(1) = (6)$, CORE = { Facebook }

Conclusion

Using hexa graph topology in matrix form allows us to model and analyze complex networks like social media platforms. By constructing and analyzing the Incidence matrix can identify Facebook as a leading social media platform based

on its connectivity and influence within the network of social media platforms.

1) **Global Reach**: Facebook has a massive user base that spans across continents, making it a platform with extensive global reach.
2) **Diverse Features**: Over the years, Facebook has expanded its features beyond just social networking, incorporating messaging (Facebook Messenger), photo and video sharing, events, groups, and more.
3) **Integration with Other Platforms**: Facebook has acquired other popular platforms like Instagram and WhatsApp, which has further strengthened its user base and offerings.
4) **Advertising Platform**: It's also a significant advertising platform, allowing businesses to reach specific demographics effectively.
5) **Established User Habits**: Many users have established habits and connections on Facebook, making it difficult for other platforms to completely displace its usage.

facebook emerge as a leading platform due to its extensive user connectivity, influential user base and well-defined community structures. However, it's important to note that social media trends can change rapidly, and new platforms and technologies can emerge that may alter the landscape in the future.

DECLARATIONS

Consent to participate – Not Applicable
Consent for publication – Not Applicable
Funding – Not Applicable
Data Availability Statement – No associated data.

CONFLICT OF INTEREST

Author declares that there is no conflict of interest.

REFERENCES

[1] Asmaa S, Qaddoori, Identification Functions in hexa Topological Spaces, Tikrit Journal of Pure Science Vol. 26 (6) 202, ISSN: 1813 - 1662,doi.org/10.25130/tjps.v26i6.198

[2] AL arjan JF, The Prevalence of Obesity. Overweight and Underweight Among Students of Al-Balqa Applied University in Jordan, Educational Sciences Studies, 38(6), 2011;2019-2036. doi: 10.2147/JMDH.S313172

[3] Bondy, J.A., and Murty, U.S.R. (2008), "Graph Theory", Springer, Berli.

[4] Bhuvaneswari K, Rasya Banu H, On Nano Forms of Weakly open sets, International Journal Math.Stat. Inven., 1 (1), 31-37, (2013)

[5] Chandramani, I., and Sakthivadivu, M. (2022), "Factorization of Complete Graph and its Application", Advances in Mathematical and Statistical Science,ISBN 978-93-91768-62-1.

[6] Chandrasekar S, Banupriya V, Sureshkumar J. Nano b-open sets in nano tri star Topological space, International Journal of Scientific Engineering and Applied Science(IJSEAS),3,61-67,(2017).

[7] Chandrasekar S, Banupriya V, Sureshkumar J, A Review On Nano Quad Topological Spaces, International Journal of Computer Science and Network, 6, 230-236, (2017)

[8] Chandra RV, Pushpalatha V,Introduction to hexa Topological spaces (6-tuple topology), International Journal of Management, Technology And Engineering ISSN NO: 2249-7455, Volume X, Issue I, JANUARY (2020), 16-18.

[9] Ekram, A., Salh, Taha, H., Jasim (2021), "On Certain Types of Set in Micro Topological Spaces with an Application in Thalassemia Sick", Tikrit Journal of Pure Science, Vol. 26 No. 2.

[10] Gary Chartrand, Linda Lesniak, and Ping Zhang (2015), "Graphs and Digraphs", Taylor and Francis Group,doi.org/10.1201/b19731.

[11] John Levi Martin and King-To Yeung. (2006), "Persistence of close personal ties over a 12-year period" , Social Networks, 28, 331â362, doi:10.1016/j.socnet.2005.07.008

[12] Khan M, Khan G, p-Continuity and p-Homeomorphism in Penta Topological Spaces, European International Journal of Science and Technology,7,(5), 2018, pp.1-8.

[13] Kirchuffs Atengble (2014), "Social Media and Ghanaâs 2012 Election Petition -A Discussion" , Advances in Journalism and Communication, 2, 121-126.

[14] Levine, N.(1970), "Generalized closed sets in topology", Rend. Circ Mat. Palermo, 19(2), 89-96.

[15] Marwa Fathi Ahmed, Taha H. Jasim (2022), "On Regular Semi Supra Open set and Regular Semi Supra Continuity", Tikrit Journal of Pure Science, Vol. 27 No. 1.

[16] Muralidhar Reddy, Y., Crissy Methew, and Suresh (2020), "Social Media: Internet Trends In India And Growth Of Social Media In The Recent Times", International Journal of Business Administration and Management Research, 8, 20-24

[17] Mukundan DV, Introduction to Quad topological spaces (4 - tuple topology), International Journal of Scientific and Engineering Research, Volume 4, (2013), Issue 7, 2483-2485.

[18] Qadouri AS, Mohammed NJ, Abdel Qader ZT, Use Factor Analysis in Determining Most Important Factors Affecting Childhood Obesity, Turkish Journal of Computer and Mathematics Education,12 (12) 2021;4621-4630 doi:10.17762/turncoat.v12i12.8501

[19] Resnicow K, Davis R, Rollnick S. Motivational interviewing for pediatric obesity: conceptual issues and evidence review. J Am Diet Assoc. 2006;106:2024â33. doi: 10.1016/j.jada.2006.09.015.

[20] Saelens BE, Lozano P, Scholz K. A randomized clinical trial comparing delivery of behavioral pediatric obesity treatment using standard and enhanced motivational approaches. J Pediatr Psychol. 2013;38:954â64. doi: 10.1093/jpepsy/jst054.

[21] Silva MN, Markland D, CarraÃ§a EV, Vieira PN, Coutinho SR, Minderico CS, et al. Exercise autonomous motivation predicts 3-yr weight loss in women. Med Sci Sports Exerc. 2011;43:728â37. doi: 10.1249/MSS.0b013e3181f3818f.

[22] Salh E, Jasim T, On Certain Types of Set in Micro Topological Spaces with an Application in Thalassemia Sick, Tikrit Journal of Pure Science,26(2), 2021;103-107.doi.org/10.25130/tjps.v26i2.132

[23] Sitaram Asur, and Bernardo A. Huberman (2010), "Predicting the Future With Social Media" Social Computing Lab: HP Labs, Palo Alto, California. 1-8.

[24] Seema MD, Aaron S, Review of Childhood Obesity: From Epidemiology, Etiology, and Comorbidities to Clinical Assessment and Treatment, Mayo Clin,92(2), Feb;92(2):251-265. doi: 10.1016/j.mayocp.2016.09.017.

[25] Sushma Rawath, S., Satheeshkumar, R., and Venkatesh Kumar. (2019), "A Study on the Impact of Social Media on Youth" Journal of Management, 6, 89-96

[26] Thivagar L, Richard C, On Nano Forms Of Weakly Open Sets, International Journal Math. Stat. Inven.,2013;31-37.

[27] Thivagar L, Richard C, Computing technique for recruitement process via nanotopology, Sohag J. Math. 3(1), 2016;37-45. doi: 10.18576/sjm/030106.

[28] Teixeira PJ, Going SB, Houtkooper LB, Cussler EC, Metcalfe LL, Blew RM, et al. Exercise motivation, eating, and body image variables as predictors of weight control. Med Sci Sports Exerc. 2006;38:179â88. doi: 10.1249/01.mss.0000180906.10445.8d.

[29] Taha, H., Jasim, and Aiad I., Awad. (2020). Some Topological Concepts Via Graph Theory. Tikrit Journal of Pure Science, 25(4), 117â122. DOI:org/10.25130/tjps.v25i4.280

[30] Wilfley DE, Hayes JF, Balantekin KN, van Buren DJ, Epstein LH. Behavioral interventions for obesity in children and adults: evidence base, novel approaches, and translation into practice. Am psychol. 2018;73:981â93. doi: 10.1037/amp0000293.

[31] Wong EM, Cheng MM. Effects of motivational interviewing to promote weight loss in obese children. J Clin Nurs. 2013;22:2519â30. doi: 10.1111/jocn.12098.

[32] Yaseen RB, Shihab A, Alobaidi M,On Nano Penta Topological Spaces,AIP Publishing(ICCEPS-2021),doi.org/10.1063/5.0115365

[33] Yaseen RB, Shihab A, and Alobaidi M, Characteristics of Penta- open sets in Penta topological spaces, Int. J. Nonlinear Anal. Appl.,12,(2), 2021, pp. 2463-2475, doi:10.22075/IJNAA.2021.5388

Analytical Computation of Atangana–Baleanu Differintegrals of Classical Special Functions

1st Sunitha K G
Department of Science & Humanities
Ahalia School of Engineering and Technology
Palakkad, India
sunitha.kg@ahalia.ac.in

2nd Sangeetha P S
Department of Science & Humanities
Ahalia School of Engineering and Technology
Palakkad, India
sangeetha.ps@ahalia.ac.in

3rd Preeja V
Department of Science & Humanities
Ahalia School of Engineering and Technology
Palakkad, India
preeja.v@ahalia.ac.in

4th Prabha R
Department of Science & Humanities
Ahalia School of Engineering and Technology
Palakkad, India
prabha.r@ahalia.ac.in

***Abstract*—In this paper, we investigate the Atangana–Baleanu (AB) differintegrals of several special functions, including trigonometric, exponential, and hyperbolic functions. The AB differintegral represents a relatively recent class of fractional operators characterized by non-local and non-singular kernels, and it satisfies important properties such as the semigroup property.**

The fractional derivatives are obtained by employing the Riemann–Liouville fractional integral formulation, together with the known expression for the fractional integral of a power function. The proposed approach enables the computation of AB differintegrals for a broad class of functions belonging to $L^1[a, b]$, for all orders α satisfying $0 < \alpha < 1$.

Furthermore, the normalization function $B(\alpha)$ can be suitably chosen to ensure that the operator satisfies desired analytical properties. The results presented in this work contribute to the growing body of research on fractional calculus with non-singular kernels and provide useful insights into the behavior of special functions under AB fractional operators.

***Index Terms*—Atangana Baleanu fractional derivatives, Mittag Leffler kernel, Normalisation function**

I. Introduction

Fractional calculus, which generalizes the concepts of differentiation and integration to non-integer orders, has emerged as a powerful mathematical tool for modeling complex systems with memory and hereditary properties [3]–[5]. In contrast to classical integer-order calculus, fractional-order operators provide a more accurate framework for describing various real-world phenomena, particularly those exhibiting anomalous diffusion, viscoelasticity, and non-local behavior [19], [20]. As a result, fractional calculus has found widespread applications in physics, engineering, biology, control theory, and finance.

Over the years, several definitions of fractional derivatives and integrals have been proposed, among which the Riemann–Liouville and Caputo operators are the most extensively studied [3], [4]. Despite their success, these classical operators are associated with singular kernels, which may not adequately capture certain physical processes. This limitation has motivated the development of new fractional operators with non-singular and non-local kernels.

In this context, Atangana and Baleanu introduced a new class of fractional operators, now widely known as the Atangana–Baleanu (AB) fractional derivatives and integrals [1]. These operators are defined using a Mittag–Leffler function kernel, which is both non-singular and non-local, thereby providing a more realistic description of systems with fading memory effects [6]. Due to these advantageous properties, AB fractional operators have attracted considerable attention in recent years and have been successfully applied in diverse areas such as heat transfer, fluid flow, signal processing, and epidemiological modeling [11]–[13].

An important feature of the AB fractional operators is that they satisfy several desirable mathematical properties, including linearity and, under suitable conditions, the semigroup property [15], [18]. Furthermore, the presence of the normalization function $B(\alpha)$ allows flexibility in the formulation of the operator, enabling it to be adapted to specific modeling requirements.

Special functions, such as trigonometric, exponential, and hyperbolic functions, play a fundamental role in mathematical analysis and its applications. The study of their fractional-order derivatives and integrals provides deeper insights into their behavior and facilitates their use in solving fractional differential equations [4], [5]. In particular, evaluating fractional operators of special functions is essential for constructing analytical and approximate solutions to various applied problems.

In this paper, we focus on the computation of the Atangana–Baleanu differintegrals of certain special functions. The analysis is carried out by utilizing the relationship between AB fractional operators and the Riemann–Liouville fractional integral, along with the well-known formula for the fractional integral of a power function [3]. This approach enables us to derive explicit expressions for the AB differintegrals of trigonometric,

exponential, and hyperbolic functions.

Moreover, the proposed formulation is applicable to a wide class of functions belonging to the space $L^1[a,b]$, for all fractional orders α satisfying $0 < \alpha < 1$. The role of the normalization function $B(\alpha)$ is also discussed, highlighting how it can be chosen to ensure specific analytical properties of the operator.

The results obtained in this study contribute to the theoretical development of fractional calculus with non-singular kernels and provide useful tools for further research involving fractional differential equations and mathematical modeling [7], [9]. It is expected that the findings of this work will be beneficial for researchers working in both theoretical and applied aspects of fractional calculus.

II. Preliminaries

In this section, we recall some basic definitions and results from fractional calculus that will be used throughout the paper.

The Riemann–Liouville fractional integral of order $\alpha > 0$ is defined by

$$I^{\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau)\, d\tau. \tag{1}$$

The Atangana–Baleanu fractional integral is defined as

$${}^{AB}I^{\alpha} f(t) = \frac{1-\alpha}{B(\alpha)} f(t) + \frac{\alpha}{B(\alpha)\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau)\, d\tau. \tag{2}$$

In this paper, we study the existence of AB integrals for the trignometric, exponential and hyperbolic functions which are useful in finding the Laplace transforms of the iterated differintegrals of certain functions.The normalisation function B(α) can be chosen according to our wish to satisfy certain properties. So, Throughout this paper, in the derivation of AB differintegrals of these functions, we take $a = 0$, $\alpha = \frac{1}{2}$ and $B\left(\frac{1}{2}\right) = 1$ and the RL intergal of the power function of order $\frac{1}{2}$. Also the final results of AB integrals of special functions are obtained as a sum of Mittag-Leffler function in two parameters or as a linear combination of the Mittag-Leffler function in two parameters.

III. Properties of Atangana–Baleanu Operators

In this section, we present some important properties of the Atangana–Baleanu (AB) fractional operators which are useful in the subsequent analysis.

A. Linearity

The AB fractional derivative is a linear operator. For any functions $f, g \in L^1[a,b]$ and constants $c_1, c_2 \in \mathbb{R}$, we have

$${}^{AB}D^{\alpha}\left(c_1 f(t) + c_2 g(t)\right) = c_1\, {}^{AB}D^{\alpha} f(t) + c_2\, {}^{AB}D^{\alpha} g(t). \tag{3}$$

B. Semigroup Property

Under suitable conditions, the AB fractional operators satisfy a semigroup-type property given by

$${}^{AB}I^{\alpha}\left({}^{AB}I^{\beta} f(t)\right) = {}^{AB}I^{\alpha+\beta} f(t), \tag{4}$$

for $\alpha, \beta > 0$.

C. Boundedness

If $f \in L^1[a,b]$, then the AB fractional integral is bounded, i.e.,

$$\|{}^{AB}I^{\alpha} f\| \le C\|f\|, \tag{5}$$

for some constant $C > 0$.

D. Reduction to Classical Case

As $\alpha \to 1$, the AB fractional derivative reduces to the classical first-order derivative, i.e.,

$$\lim_{\alpha \to 1} {}^{AB}D^{\alpha} f(t) = \frac{d}{dt} f(t). \tag{6}$$

E. Action on Constant Function

For a constant function $f(t) = C$, the AB fractional derivative satisfies

$${}^{AB}D^{\alpha} C = 0. \tag{7}$$

IV. AB Integrals of Sine and Cosine functions

A. AB Integral of sine function

$$\begin{aligned}
{}^{AB}I_{0+}^{\frac{1}{2}} \sin t &= \frac{\frac{1}{2}}{B(\frac{1}{2})} \sin t + \frac{\frac{1}{2}}{B(\frac{1}{2})} \frac{1}{\Gamma(\frac{1}{2})} \int_0^t (t-x)^{-\frac{1}{2}} \sin x dx \\
&= \frac{1}{2} \sin t + \frac{1}{2} \frac{1}{\Gamma(\frac{1}{2})} \int_0^t (t-x)^{-\frac{1}{2}}\, x E_{2,2}\left(-x^2\right) dx \\
&= \frac{1}{2} \sin t + \frac{1}{2} \frac{1}{\Gamma(\frac{1}{2})} \int_0^t (t-x)^{-\frac{1}{2}}\, x \sum_{k=0}^{\infty} \frac{\left(-x^2\right)^k}{\Gamma(2k+2)} dx \\
&= \frac{1}{2} \sin t + \frac{1}{2} \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(2k+2)} \frac{1}{\Gamma(\frac{1}{2})} \int_0^t (t-x)^{-\frac{1}{2}}\, x^{2k+1} dx \\
&= \frac{1}{2} \left\{ t E_{2,2}\left(-t^2\right) + t^{\frac{3}{2}} E_{2,\frac{5}{2}}\left(-t^2\right) \right\}
\end{aligned}$$

B. AB Integral of cosine function

$$\begin{aligned}
{}^{AB}I_{0+}^{\frac{1}{2}} \cos t &= \frac{\frac{1}{2}}{B(\frac{1}{2})} \cos t + \frac{\frac{1}{2}}{B(\frac{1}{2})} \frac{1}{\Gamma(\frac{1}{2})} \int_0^t (t-x)^{-\frac{1}{2}} \cos x dx \\
&= \frac{1}{2} \cos t + \frac{1}{2} \frac{1}{\Gamma(\frac{1}{2})} \int_0^t (t-x)^{-\frac{1}{2}}\, E_{2,1}\left(-x^2\right) dx \\
&= \frac{1}{2} \cos t + \frac{1}{2} \frac{1}{\Gamma(\frac{1}{2})} \int_0^t (t-x)^{-\frac{1}{2}} \sum_{k=0}^{\infty} \frac{\left(-x^2\right)^k}{\Gamma(2k+1)} dx \\
&= \frac{1}{2} \cos t + \frac{1}{2} \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(2k+1)} \frac{1}{\Gamma(\frac{1}{2})} \int_0^t (t-x)^{-\frac{1}{2}}\, x^{2k} dx \\
&= \frac{1}{2} \left\{ E_{2,1}\left(-t^2\right) + t^{\frac{1}{2}} E_{2,\frac{3}{2}}\left(-t^2\right) \right\}
\end{aligned}$$

V. AB Integral of Exponential function

$$
\begin{aligned}
{}^{AB}I_{0+}^{\frac{1}{2}}e^{t} &= \frac{\frac{1}{2}}{B(\frac{1}{2})}e^{t} + \frac{\frac{1}{2}}{B(\frac{1}{2})}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\, e^{x}dx \\
&= \frac{1}{2}e^{t} + \frac{1}{2}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\, E_{1,1}\left(x\right)dx \\
&= \frac{1}{2}e^{t} + \frac{1}{2}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\sum_{k=0}^{\infty}\frac{x^{k}}{\Gamma(k+1)}dx \\
&= \frac{1}{2}e^{t} + \frac{1}{2}\sum_{k=0}^{\infty}\frac{(-1)^{k}}{\Gamma(k+1)}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\, x^{k}dx \\
&= \frac{1}{2}\left\{E_{2,1}\left(-t^{2}\right) + t^{\frac{1}{2}}E_{2,\frac{3}{2}}\left(-t^{2}\right)\right\}
\end{aligned}
$$

VI. AB Integrals of Hyperbolic Sine and Cosine functions

A. AB Integral of Hyperbolic sine function

$$
\begin{aligned}
{}^{AB}I_{0+}^{\frac{1}{2}}\sinh t &= \frac{\frac{1}{2}}{B(\frac{1}{2})}\sinh t \\
&\quad + \frac{\frac{1}{2}}{B(\frac{1}{2})}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\, sinhxdx \\
&= \frac{1}{2}\sinh t + \frac{1}{2}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\, xE_{2,2}\left(x^{2}\right)dx \\
&= \frac{1}{2}\sinh t \\
&\quad + \frac{1}{2}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\, x\sum_{k=0}^{\infty}\frac{\left(x^{2}\right)^{k}}{\Gamma(2k+2)}dx \\
&= \frac{1}{2}\sinh t \\
&\quad + \frac{1}{2}\sum_{k=0}^{\infty}\frac{1}{\Gamma(2k+2)}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\, x^{2k+1}dx \\
&= \frac{1}{2}\left\{tE_{2,2}\left(t^{2}\right) + t^{\frac{3}{2}}E_{2,\frac{5}{2}}\left(t^{2}\right)\right\}
\end{aligned}
$$

B. AB Integral of Hyperbolic cosine function

$$
\begin{aligned}
{}^{AB}I_{0+}^{\frac{1}{2}}\cosh t &= \frac{\frac{1}{2}}{B(\frac{1}{2})}\cosh t \\
&\quad + \frac{\frac{1}{2}}{B(\frac{1}{2})}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\cosh x dx \\
&= \frac{1}{2}\cosh t + \frac{1}{2}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\, E_{2,1}\left(x^{2}\right)dx \\
&= \frac{1}{2}\cosh t + \frac{1}{2}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\sum_{k=0}^{\infty}\frac{\left(x^{2}\right)^{k}}{\Gamma(2k+1)}dx \\
&= \frac{1}{2}\cosh t \\
&\quad + \frac{1}{2}\sum_{k=0}^{\infty}\frac{1}{\Gamma(2k+1)}\frac{1}{\Gamma(\frac{1}{2})}\int_0^t (t-x)^{-\frac{1}{2}}\, x^{2k}dx \\
&= \frac{1}{2}\left\{E_{2,1}\left(t^{2}\right) + t^{\frac{1}{2}}E_{2,\frac{3}{2}}\left(t^{2}\right)\right\}
\end{aligned}
$$

VII. Summary of Results

TABLE I
ATANGANA–BALEANU FRACTIONAL INTEGRALS OF SPECIAL FUNCTIONS FOR $\alpha = \frac{1}{2}$ AND $B(\frac{1}{2}) = 1$

Function $f(t)$	${}^{AB}I^{1/2}f(t)$
$\sin t$	$\frac{1}{2}\left[t\,E_{2,2}\left(-\frac{t^2}{2}\right) + t^{3/2}E_{2,\frac{5}{2}}\left(-\frac{t^2}{2}\right)\right]$
$\cos t$	$\frac{1}{2}\left[E_{2,1}\left(-\frac{t^2}{2}\right) + t^{1/2}E_{2,\frac{3}{2}}\left(-\frac{t^2}{2}\right)\right]$
e^t	$\frac{1}{2}\left[E_{2,1}\left(-\frac{t^2}{2}\right) + t^{1/2}E_{2,\frac{3}{2}}\left(-\frac{t^2}{2}\right)\right]$
$\sinh t$	$\frac{1}{2}\left[t\,E_{2,2}\left(\frac{t^2}{2}\right) + t^{3/2}E_{2,\frac{5}{2}}\left(\frac{t^2}{2}\right)\right]$
$\cosh t$	$\frac{1}{2}\left[E_{2,1}\left(\frac{t^2}{2}\right) + t^{1/2}E_{2,\frac{3}{2}}\left(\frac{t^2}{2}\right)\right]$

VIII. CONCLUSION

In this paper, we have introduced the Atangana–Baleanu (AB) differintegrals of several special functions, including trigonometric, hyperbolic, and exponential functions. These computations can be effectively used in evaluating iterated AB differintegrals and in solving fractional differential equations [1], [3]. The AB integrals of these functions are particularly useful in modeling complex systems with memory and non-local effects, where classical integer-order models may not be sufficient [12], [13], [19], [20].

Furthermore, the iterated AB integrals possess an important property known as the semigroup property, which is not generally satisfied by classical fractional operators. This property plays a significant role in describing real-world dynamical systems and enhances the applicability of AB fractional models in various scientific fields [6], [15].

REFERENCES

[1] A. Atangana and D. Baleanu, New fractional derivatives with nonlocal and nonsingular kernel: Theory and application to heat transfer model, *Thermal Science*, 20(2) (2016), 763–769.

[2] A. Atangana and D. Baleanu, Caputo–Fabrizio derivative applied to groundwater flow within confined aquifer, *European Physical Journal Plus*, 132 (2017), 1–12.

[3] I. Podlubny, *Fractional Differential Equations*, Academic Press, 1999.

[4] A.A. Kilbas, H.M. Srivastava, and J.J. Trujillo, *Theory and Applications of Fractional Differential Equations*, Elsevier, 2006.

[5] S.G. Samko, A.A. Kilbas, and O.I. Marichev, *Fractional Integrals and Derivatives: Theory and Applications*, Gordon and Breach, 1993.

[6] A. Fernandez, A complex analysis approach to Atangana–Baleanu fractional calculus, *Mathematical Methods in the Applied Sciences*, 2019. :contentReference[oaicite:1]index=1

[7] M.I. Syam and M. Al-Refai, Fractional differential equations with Atangana–Baleanu fractional derivative: Analysis and applications, *Chaos, Solitons & Fractals*, 2019. :contentReference[oaicite:2]index=2

[8] S.T. Sutar et al., Existence and data dependence results for fractional differential equations involving AB derivative, *Mathematics*, 2021. :contentReference[oaicite:3]index=3

[9] N. Pandey, N. Pandey, and R.K. Shukla, A comprehensive review of the Atangana–Baleanu fractional operator: Theory and applications, *Journal of Computational Analysis and Applications*, 2024. :contentReference[oaicite:4]index=4

[10] R.T. Alqahtani, Atangana–Baleanu derivative applied to groundwater model, *Journal of Nonlinear Science and Applications*, 2016. :contentReference[oaicite:5]index=5

[11] K.M. Owolabi, Modelling and simulation of a dynamical system with AB fractional derivative, *European Physical Journal Plus*, 133 (2018), 15.

[12] N. Sene, Stokes' first problem for heated flat plate with AB fractional derivative, *Chaos, Solitons & Fractals*, 117 (2018), 68–75.

[13] A. Yusuf, M. Inc, A.I. Aliyu, and D. Baleanu, Efficiency of fractional derivative with Mittag-Leffler kernel, *Chaos, Solitons & Fractals*, 116 (2018), 220–226.

[14] T. Abdeljawad and A. Fernandez, On discrete Atangana–Baleanu fractional operators, *Applied Mathematics*, 2019. :contentReference[oaicite:6]index=6

[15] A.O. Akdemir et al., Fractional integral inequalities via AB operators, *Mathematics*, 2021. :contentReference[oaicite:7]index=7

[16] O. Martínez-Fuentes and S.M. Delfín-Prieto, Stability of nonlinear systems with Mittag-Leffler kernel, *Applied Mathematics*, 2020. :contentReference[oaicite:8]index=8

[17] A.S. Alshehry et al., Comparative analysis of fractional advection–dispersion equations with AB derivative, *Symmetry*, 2023. :contentReference[oaicite:9]index=9

[18] V. Pandiyammal and U. Karthik Raja, Existence solutions for AB fractional neutral differential equations, *Journal of Mathematical and Computational Science*, 2021. :contentReference[oaicite:10]index=10

[19] K. Diethelm et al., Trends and open problems in fractional calculus, *Nonlinear Dynamics*, 2020.

[20] A.A. Tateishi, H.V. Ribeiro, and E.K. Lenzi, The role of fractional derivatives in anomalous diffusion, *Frontiers in Physics*, 2017.

Graph Analytics for Large-Scale Networks

Binoy Balan K.
Department of Science and Humanities
Ahalia School of Engineering and Technology
Palakkad,Kerala.
hodsh.aset@ahalia.ac.in

G.Murugananth
Department of Electrical and Electronics Engineering
Ahalia School of Engineering and Technology
Palakkad,Kerala.
hodsh.aset@ahalia.ac.in

Theertha K.
Department of Science and Humanities
Ahalia School of Engineering and Technology
Palakkad,Kerala.
theertha_k@ahalia.ac.in

Rajitha M.
Department of Science and Humanities
Ahalia School of Engineering and Technology
Palakkad,Kerala.
hodsh.aset@ahalia.ac.in

Jishnu P.
Department of Science and Humanities
Ahalia School of Engineering and Technology
Palakkad,Kerala.
hodsh.aset@ahalia.ac.in

***Abstract*—The rapid proliferation of digital technologies has led to the generation of massive datasets, with unstructured data forming the vast majority. Traditional data analysis techniques, which rely on structured data streams, are often inadequate for modeling these complex systems. This paper provides a comprehensive overview of graph analytics as applied to large-scale networks. By representing complex systems—such as social relationships, scientific data, and sensor grids—as mathematical graphs, structural and behavioral insights can be extracted. This article reviews fundamental network models, key centrality measures, the significance of cliques and clustering coefficients, and prominent algorithmic approaches for community detection, including relevant mathematical formulations and algorithmic complexities.**

Keywords—Graph Analytics, Unstructured Data, Centrality Measures, Community Detection, Clustering Coefficients

I. INTRODUCTION

Data analytics fundamentally encompasses the life cycle of data: creating data, analyzing data, and drawing conclusions. In modern data storage architectures, there is a clear distinction between a Data Lake, which easily accommodates long-historical data, and a Data Stream, which processes real-time data but makes parallel analysis difficult.

Currently, unstructured data comprises approximately 80% of all data, with total volume estimates reaching 3,000,000 petabytes. This massive scale of data creation is driven by diverse sources. For example, nearly one million bits of content are shared on Facebook every minute, and over 100,000 tweets are posted per minute. Data is continuously generated by social media platforms, scientific instruments collecting diverse metrics, mobile devices tracking geographic objects constantly, and extensive sensor technology networks

II. SOCIETY AS A GRAPH AND NETWORK MODELS

When analyzing large-scale systems, society can be effectively modeled as a graph where people are represented as nodes and their relationships (such as acquaintanceships, friendships, or co-authorships) are represented as edges.

A. The Small-World Phenomenon.

A prominent characteristic of human networks is the "Six Degrees of Separation" principle, which posits that all people are six or fewer social connections away from each other. The concept was originally introduced in a 1929 short story by Frigyes Karinthy. Common manifestations include:

The Kevin Bacon Game: Invented by Albright College students in 1994, this exercise aims to connect any actor to Kevin Bacon by linking actors who have appeared in the same movie.

The Erdős Number: A metric calculating the number of links required to connect scholars to the prolific mathematician Paul Erdős, who authored over 1,500 papers with 507 co-authors

B. Graph Centrality Measures

Centrality measures are vital for understanding networks, identifying the most important nodes such as influencers in a social network or super-spreaders of a disease.

- **Degree Centrality:** Defined as the number of links incident upon a node. In a directed graph, this is sub-divided into in-degree and out-degree. Real-world networks frequently form "Scale-Free Networks" where the degree distribution follows a power law.
- **Closeness Centrality:** Defined by Bavelas (1950) as the reciprocal of farness, this is calculated as the sum of the length of the shortest paths between the node and all other nodes.
- **Betweenness Centrality:** The betweenness centrality for a vertex is the number of shortest paths between vertex pairs that pass through it.
- **Eigenvector Centrality:** This measure posits that endorsements from important nodes count more. A node with high eigenvector centrality is linked to by other important nodes, differing from simple in-degree centrality.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

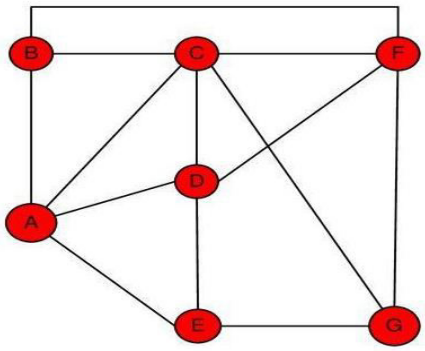


Figure 1.

Centrality measures provide a rigorous framework for identifying structurally influential nodes within complex interaction networks. In the present example, shown in Figure 1, an undirected network consisting of seven nodes (A–G) is examined to quantify node importance using four classical centrality metrics: degree, closeness, betweenness, and eigenvector centrality. The network topology reflects heterogeneous connectivity patterns in which node C maintains the highest number of direct links relative to other vertices. Such structural asymmetry is characteristic of many empirical interaction systems, including collaboration networks, communication infrastructures, and online discussion environments, where certain actors function as hubs facilitating information flow across the network. These measures collectively enable a multidimensional interpretation of influence that extends beyond simple connection counts to include accessibility, brokerage capacity, and recursive prestige effects.

Degree centrality was first computed as the normalized number of direct connections incident on each node. With a total of seven nodes in the network, normalization was performed by dividing node degree by $n - 1 = 6$. The results indicate that node C possesses the highest degree centrality (0.833), followed by nodes A, D, and F (each 0.667), while nodes B, E, and G exhibit lower values (0.500). These findings suggest that node C occupies a dominant local position within the network structure, enabling it to directly interact with a larger proportion of nodes compared with others. Consequently, node C can be interpreted as the principal local hub supporting rapid dissemination of information within the interaction topology.

For an undirected graph:

$$C_D(v) = \frac{\deg(v)}{n-1}$$

Since $n - 1 = 6$:

Node	Degree	Normalized Degree Centrality
A	4	0.667
B	3	0.500
C	5	0.833
D	4	0.667
E	3	0.500
F	4	0.667
G	3	0.500

Closeness centrality was subsequently evaluated to determine the efficiency with which each node can access all other vertices through shortest communication paths. The computed results show that node C again achieves the highest closeness centrality (0.857), followed by nodes A, D, and F (each 0.750), whereas nodes B, E, and G display comparatively lower accessibility (0.667). These results indicate that node C maintains the shortest average geodesic distance to all other nodes, thereby enabling faster interaction reachability and improved transmission efficiency across the network. From a functional perspective, such nodes often serve as optimal coordinators for distributing information or resources in collaborative environments.

Closeness centrality is computed as:

$$C_C(v) = \frac{n-1}{\sum d(v,u)}$$

Shortest-path distance sums:

Node	Distance Sum	Closeness Centrality
A	8	0.750
B	9	0.667
C	7	0.857
D	8	0.750
E	9	0.667
F	8	0.750
G	9	0.667

Betweenness centrality was then calculated to evaluate the extent to which nodes mediate communication between other node pairs. The analysis reveals that node C attains the highest normalized betweenness centrality value (0.433), substantially exceeding those of nodes D and F (each 0.167), node A (0.133), and nodes B, E, and G (each 0.033). This result demonstrates that node C frequently lies along the shortest paths connecting other vertices and therefore functions as a critical intermediary controlling information flow between otherwise weakly connected substructures. Such brokerage roles are particularly significant in collaborative and communication networks because they enhance structural cohesion while simultaneously increasing the strategic importance of intermediary actors

Betweenness centrality measures how often a node lies on shortest paths between node pairs:

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Normalized results:

Node	Betweenness Centrality
A	0.133
B	0.033
C	0.433
D	0.167
E	0.033
F	0.167
G	0.033

Finally, eigenvector centrality was computed to assess node influence based not only on connection quantity but also on the structural importance of neighboring nodes. The resulting scores indicate that node C again ranks highest (0.521), followed by nodes A, D, and F(each 0.394), while nodes B, E, and Gshow comparatively lower influence levels (0.284). Unlike degree centrality, this measure captures recursive prestige effects, demonstrating that node $C$is connected to other structurally significant vertices and therefore exerts amplified influence across the network. Taken together, the consistent dominance of node C across all four centrality metrics confirms its role as the most structurally influential node in the examined topology, functioning simultaneously as a hub, mediator, and prestige-bearing actor within the interaction system.

Eigenvector centrality satisfies:

$$Ax = \lambda x$$

where A is the adjacency matrix and x is the dominant eigenvector. Normalized principal eigenvector scores are approximately:

Node	Eigenvector Centrality
A	0.394
B	0.284
C	0.521
D	0.394
E	0.284
F	0.394
G	0.284

III. CLIQUES AND CLUSTERING COEFFICIENTS

A. Cliques

A clique is a complete subgraph where edges exist between all pairs of nodes. Finding a maximal clique is generally an NP-Complete problem, though it is polynomial for interval graphs.Cliques constitute a fundamental structural concept in network analysis, representing *complete subgraphs* in which every pair of vertices is directly connected. Formally, a clique of size k is a subset of k nodes such that all possible $\frac{k(k-1)}{2}$edges among them are present. The identification of cliques enables the detection of tightly cohesive substructures that often correspond to functional communities in social, biological, and collaboration networks. In practical applications, locating the **maximum clique** (largest fully connected subgraph) is computationally intractable in general graphs because the problem is NP-complete; however, efficient polynomial-time solutions exist for restricted graph classes such as interval graphs. For example, in a research collaboration network, if authors A, B, C, and Dhave each co-authored papers with one another, they form a 4-node clique, indicating a strongly integrated scholarly subgroup. Such cohesive clusters frequently correspond to specialized research teams sharing resources and knowledge intensively, thereby accelerating innovation within localized substructures

Beyond identifying cohesive communities, clique detection also supports the characterization of influence propagation and redundancy in communication pathways. Nodes belonging to multiple overlapping cliques typically act as bridges between otherwise distinct collaborative groups, enhancing both robustness and information diffusion efficiency across the network. For instance, consider an online academic discussion forum in which participants C, D, and Finteract mutually, forming a 3-node clique, while node Csimultaneously participates in another clique with nodes $A$and B. In this scenario, node C plays a structurally strategic role by linking two tightly connected subgroups, thereby facilitating interdisciplinary exchange. Algorithms such as the Bron–Kerbosch method are widely employed for maximal clique enumeration in empirical studies, enabling researchers to reveal hidden community structures within complex relational systems (Bron & Kerbosch, 1973).

B. Clustering Coefficients

Introduced by Duncan J. Watts and Steven Strogatz in 1998, the clustering coefficient quantifies how close neighbors are to forming a clique. The local clustering coefficient (Ci) for a vertex with n neighbors is calculated as the number of links between those n neighbors divided by [n(n-1)/2]. The clustering coefficient of the whole network is the average of all local coefficients.

Closely related to clique structure is the **clustering coefficient**, introduced by Watts and Strogatz (1998) to quantify the degree to which neighboring nodes of a given vertex tend to form tightly interconnected groups. The local clustering coefficient of a node $i$with $n$neighbors is defined as the ratio between the number of actual links among those neighbors and the maximum possible number of such links, expressed as $\frac{n(n-1)}{2}$. This metric therefore measures the extent to which a node's neighborhood approximates a clique-like structure. For example, if node A has three neighbors (B, C, and D) and only two of the three possible links among them exist (e.g., B–$C$and C–D, but not B–D), then the clustering coefficient equals $2/3$. A high local clustering coefficient indicates strong local cohesion and redundancy in communication pathways, which can enhance network resilience and trust formation in social interaction environments. At the global level, the clustering coefficient of an entire network is computed as the average of all local clustering coefficients, providing an aggregate indicator of structural cohesiveness. Networks exhibiting high average clustering coefficients often display *small-world characteristics*, combining dense local neighborhoods with relatively short global path lengths. For instance, in collaboration networks among researchers within a specialized conference community, participants frequently co-author papers with mutual collaborators, producing high clustering values that reflect strong thematic alignment and repeated interaction patterns. Conversely, technological infrastructure networks such as power grids typically exhibit lower clustering coefficients due to their hierarchical rather than community-based connectivity patterns. Consequently, clustering coefficient analysis provides a critical quantitative lens for distinguishing between decentralized community-driven systems and centrally coordinated infrastructures, thereby enriching structural interpretation in complex network studies.

IV. Community Detection In Large Networks

Real-world networks exhibit distinct community structures. The objective is to divide the network into non-empty groups.

1. **Spectral Bisection:** This uses the eigenvectors of the graph Laplacian matrix (L = D - A, where A is the adjacency matrix and D is the diagonal matrix of vertex degrees). Bisection relies on the eigenvector corresponding to the lowest non-zero eigenvalue. The time complexity is $O(n^3)$.
2. **Kernighan-Lin (KL) Algorithm:** A greedy algorithm that swaps pairs of vertices between two randomly divided groups to maximize a benefit function (ΔQ), which calculates the difference between internal and external edges.
3. **Hierarchical Clustering:** This approach builds a dendrogram based on a calculated similarity measure. The overall time complexity is $O(n^2 \log n)$.
4. **Girvan and Newman (GN) Algorithm:** A divisive algorithm relying on edge betweenness. By iteratively removing the edge with the highest betweenness, the graph fractures into components. The time complexity ranges from $O(m^2n)$ to $O(n^3)$.

Conclusion

Graph analytics offers critical mathematical frameworks to navigate unstructured datasets. By abstracting entities into nodes and edges, data scientists can utilize centralities, clustering coefficients, and algorithms like Girvan-Newman or Spectral Bisection to unravel the structures embedded within modern networks. Eigenvector centrality has proven to be an indispensable analytical tool in physics and chemistry. Its ability to capture global influence patterns, model energy propagation, and identify structurally or functionally critical components makes it highly valuable for studying complex systems. As spectral methods and computational capabilities advance, eigenvector centrality is expected to play an even more prominent role in understanding and designing physical, chemical, biological, and quantum systems.

References

[1] F. Karinthy, "Chain-Links," *Everything is Different*, 1929.

[2] C. Fass, B. Turtle, and M. Ginelly, "The Kevin Bacon Game," Albright College, 1994.

[3] J. Grossman, "The Erdős Number Project," Oakland University.

[4] P. Erdős and A. Rényi, "On Random Graphs I," *Publicationes Mathematicae*, vol. 6, pp. 290-297, 1959.

[5] A. Bavelas, "Communication Patterns in Task-Oriented Groups," *Journal of the Acoustical Society of America*, vol. 22, no. 6, pp. 725-730, 1950.

[6] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440-442, 1998.

[7] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *PNAS*, vol. 99, no. 12, pp. 7821-7826, 2002.

[8] Bron, C., & Kerbosch, J. (1973). Algorithm 457: Finding all cliques of an undirected graph. *Communications of the ACM, 16*(9), 575–577. https://doi.org/10.1145/362342.362367

[9] Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature, 393*(6684), 440–442.

.

Comparative Structural Analysis of ERCO Ercoupe Wing Rib with Different Cut-Out Geometries

Dr. Ajith V S,
Associate Professor
Dept.of Aeronautical Engineering,
Jawaharlal College of Engineering and Technology, Lakkidi, Palakkad.
vs.ajith.aero@gmail.com

Dr. Rajeev N,
Dean Academics & Professor,
Ahaliya School of Engineering and Technology,
Palakkad

FATHIMA FIDHA K
UG student, Dept.of Aeronautical Engineering, Jawaharlal College of Engineering and Technology, Lakkidi, Palakkad

GEETHIKA K G
UG student, Dept.of Aeronautical Engineering, Jawaharlal College of Engineering and Technology, Lakkidi, Palakkad

ARDRA P
UG student, Dept.of Aeronautical Engineering, Jawaharlal College of Engineering and Technology, Lakkidi, Palakkad.

HARITHA R S
UG student, Dept.of Aeronautical Engineering, Jawaharlal College of Engineering and Technology, Lakkidi, Palakkad

***Abstract*—The structural efficiency of aircraft components is important to enhance performance, weight reduction, and safety. The wing rib is one of these components that help in the maintenance of the airfoil shape, as well as passing the aerodynamic loads to the primary structural members. The paper under discussion is based on the structural investigation of a wing rib of the ERCO Ercoupe aircraft to examine how various cut-out shapes impact its performance. The wing rib design was based on a plane that was situated in the hangar and designed using CAD software. It was introduced with three cut-out patterns of circular, elliptical and rectangular to minimize the structural weight. Under the same loading and boundary conditions, Finite Element Analysis (FEA) was conducted to analyze the distribution of stress, deformation and strain features of each configuration. The findings revealed that the use of cut-outs has a tremendous effect on structural behaviour of the wing rib. The rib with no cut-outs exhibited little deformation and stress, but was relatively heavier. The elliptical cut-out was one of the altered designs that had enhanced structural performance characterized by reduced stress concentration and controlled deformation. Conversely, the circular cut-out caused more deformation and stress, which implies that it was not as stiff when compared to the rectangular cut-out which had a localized stress in the corners. All in all, the research finds that cut-out geometry is an important aspect in the achievement of weight reduction and structural integrity. The cut-out shape is an elliptical shape which offers a good balance between strength and weight and is therefore the best design out of the cases under consideration. The results of this work help to create light and efficient aircraft constructions and form a basis of further optimization investigations.**

Keywords— Wing Rib, Cut-Out Geometry, Finite Element Analysis, Stress Distribution, Structural Optimization, Deformation Analysis, Aircraft Structures, Weight Reduction.

I. Introduction

Designed by Fred E. Weick, the ERCO Ercoupe is a two-seat American light aircraft that first flew in 1937. It was developed with a strong focus on safety, featuring spin-resistant characteristics and simplified, coordinated controls that make it easier to operate than conventional aircraft. It was also one of the first light aircraft to use tricycle landing gear, allowing it to handle on the ground like a car. After World War II, it became popular as an affordable and easy-to-fly personal aircraft. The ERCO Ercoupe 415-C features a low-wing monoplane design, providing better aerodynamic efficiency and stability. Its wings produce smooth and predictable lift, and its unique control system links ailerons with the rudder to enable coordinated turns without separate rudder input, reducing the risk of spins and stalls. The wing structure is lightweight yet strong, typically made of metal with fabric covering in early models, ensuring a balance between durability and performance[1][2].

Wing ribs shown in fig:1, are essential structural components that define the airfoil shape and extend from the leading edge to the trailing edge, connecting to the spars. They help maintain the aerodynamic profile, transfer loads from the skin to the spars, prevent buckling, and ensure structural rigidity. In the Ercoupe, ribs are placed at regular intervals and connected to a main spar located around 25% of the chord, forming a strong internal framework. To improve efficiency, ribs are designed with lightening holes (cut-outs), which reduce weight while maintaining strength and also provide space for internal systems like cables and fuel lines. However, these cut-outs can cause stress concentration and local deformation, so their geometry must be carefully designed. Overall, wing ribs play a crucial role in achieving lightweight construction, proper load distribution, and safe flight performance[3][4].

Fig. 1 ERCO Ercoupe 415-C wing rib

The advancement of digital technologies has introduced the concept of digital twins in aerospace engineering, enabling real-time monitoring, simulation, and optimization of structural components. A digital twin is a virtual representation of a physical system that integrates design data, operational conditions, and performance metrics to predict structural behavior and support decision-making. In the context of aircraft structures, digital twins facilitate continuous assessment of components such as wing ribs by combining computational models with real-world data.

The integration of digital twin frameworks with structural analysis tools, such as finite element methods[5], allows for improved prediction of stress distribution, deformation, and potential failure regions. This approach enhances the capability to

optimize structural designs for weight reduction while maintaining safety and reliability. Furthermore, digital twin-based methodologies enable iterative design improvements and lifecycle management of aerospace components, making them highly relevant for modern aircraft design and maintenance[6][7].

II. LITERATURE SURVEY

The literature survey on wing rib design and analysis highlights the importance of achieving an optimal balance between structural strength and weight reduction in aircraft components[8][3][9]. Various studies emphasize that wing ribs are essential for maintaining the aerofoil shape, distributing loads, and ensuring overall structural stability[10][11]. Introducing cutouts such as circular, rectangular, and elliptical shapes is a common technique used to reduce the weight of the wing rib. However, these cutouts also lead to stress concentration around their edges, making those regions critical for structural analysis and design considerations[12][13].

Several research works indicate that the shape and size of the cutouts significantly influence the stress distribution and deformation characteristics of the wing rib. Circular cutouts often result in higher stress concentration, whereas rectangular and elliptical cutouts may provide better performance depending on their geometry and placement. It is also observed that wing ribs without cutouts exhibit lower stress and deformation, but they are heavier and less efficient. Therefore, the use of optimized cutout designs is necessary to achieve both strength and weight reduction[14].

Advanced analysis methods such as the Finite Element Method (FEM) are widely used to evaluate the structural behaviour of wing ribs under various loading conditions[15]. Studies involving composite materials reveal that specific ply orientations, such as 0/45/90/90/45/0, offer improved strength, stiffness, and reduced deflection compared to conventional materials. Additionally, optimization techniques, including topology optimization and computational simulations, help in identifying efficient rib configurations that enhance load distribution and structural reliability[16][17][18].

Furthermore, research highlights the role of aerodynamic performance and material selection in wing rib design. Optimized configurations, especially elliptical designs, contribute to improved lift-to-drag ratio, fuel efficiency, and reduced emissions. Composite materials are increasingly preferred due to their high strength-to-weight ratio, cost-effectiveness, and resistance to extreme conditions[19]. Overall, the literature concludes that a well-optimized wing rib design, considering cutout geometry, material properties, and loading conditions, is essential for achieving safe, efficient, and high-performance aircraft structures[20][21][22].

III. METHODOLOGY

The methodology adopted for this study focuses on the structural analysis and optimization of an aircraft wing rib using a systematic and step-by-step approach. It begins with collecting geometric data from an actual aircraft, followed by aerofoil identification and reconstruction. The wing rib is then modeled using Autodesk Fusion 360, where different cutout configurations such as circular, elliptical, and rectangular shapes are introduced. Material properties are assigned, and Finite Element Modelling (FEM) is performed by applying suitable boundary conditions and loads. The structural behaviour is then analysed in terms of stress distribution, deformation, and weight reduction, followed by comparison of results and optimization of the most efficient cutout design.

Further, the wing rib geometry is developed based on the extracted aerofoil profile, extending from the leading edge to the trailing edge and incorporating structural elements like flanges and web regions. A baseline model without cutouts is first created for comparison. CAD modelling involves sketching the aerofoil profile, defining thickness, and generating a solid model. Separate models are created for each cutout configuration to ensure accurate analysis.

To achieve weight reduction while maintaining structural integrity, different cutout shapes are carefully introduced into the rib web (shown from figure 2 to 5). These cutouts are positioned in similar regions across all models to ensure a fair comparison. Proper design considerations are taken to avoid excessive weakening of the structure, ensuring that the wing rib remains strong, efficient, and suitable for aircraft applications.

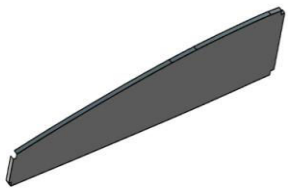


Fig: 2, wing rib of ERCO ERCOUPE without cutout shapes

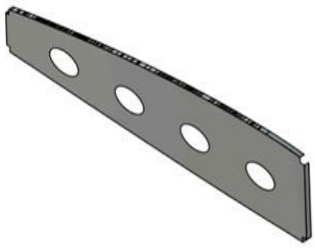


Fig: 3, wing rib of ERCO ERCOUPE with elliptical cutout shapes

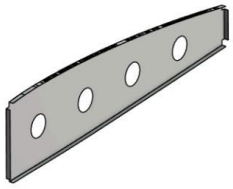


Fig:4, wing rib of ERCO ERCOUPE with circular cutout shapes

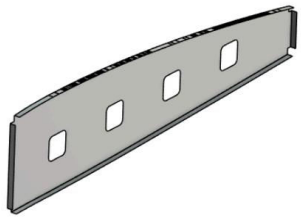


Fig:5, Wing rib of ERCO Ercoupe with rectangular cutout shapes

IV. MESH ANALYSIS OF DIFFERENT CUT-OUTS

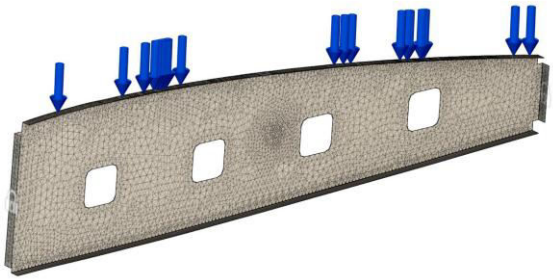


Fig: 6, Mesh analysis of wing rib with rectangular cut-outs

In the fig: 6, the plate contains rectangular-shaped holes instead of circular ones. Mesh refinement is more critical at the sharp corners of the square cut-outs, where stress concentration is higher compared to circular holes. The loading condition remains similar with downward forces along the top edge. This model is ideal for comparing stress concentration between circular and square discontinuities.

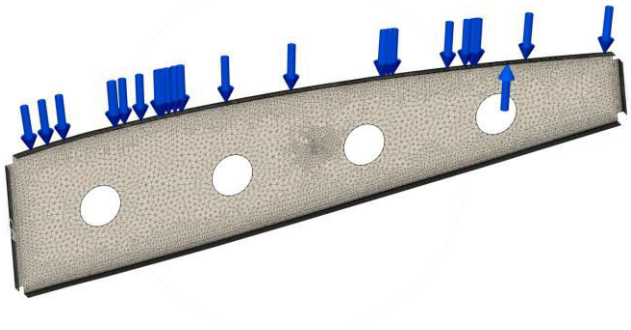


Fig:7, Mesh analysis of wing rib with circular cut-outs

The structure is a rectangular plate with multiple circular holes as shown in fig 7. A fine triangular mesh is used, with higher refinement near the holes to capture stress concentration effects. Downward distributed loads are applied along the top edge, while one location shows an upward reaction force, indicating a support or constraint. This model helps study how circular discontinuities influence stress distribution under loading.

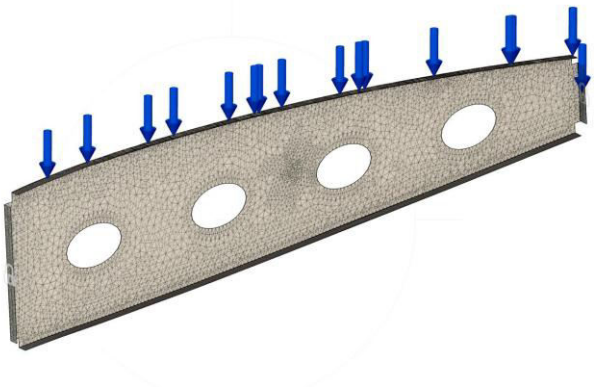


Fig:8, Mesh analysis of wing rib with elliptical cut-outs

The geometry is slightly curved along the top edge, introducing geometric nonlinearity compared to circular cut-outs as shown in fig:8. Similar triangular meshing strategy is used, with denser mesh near the holes. Multiple downward forces are applied at discrete points along the top boundary. The curvature affects load distribution and bending behavior, making this model useful for analyzing real-world structural panels like aircraft skin sections.

V. ANALYTICAL COMPUTATION

The aerodynamic load acting on the wing was converted into an equivalent uniform pressure. This pressure was applied uniformly on the rib surface to simulate aerodynamic loading.

Total aircraft weight(W) = 572 kg =5609.4 N
Wing span(b) = 9.14 m
Chord length(C) = 5 ft = 1.524 m
Number of wings = 2
Number of ribs(N) = 8×2 = 16
Calculation:
W=L (straight level flight)
Load carried by wing(1w) = W/2
Load carried by 1 rib (Lrib) = W/2N
= 5609.4/ (2×16) = 175.29 N
Area supported by 1 rib (Arib) = Spacing ×Chord
= b/(2N) ×C
= 9.14/(2×16) × 1.524
Arib = 0.4352925 m2
Average pressure on rib = F/A= Lrib /Arib
=175.29/0.4352925
= 40.26 MPa

The study focuses on improving aircraft structural efficiency by optimizing wing rib design. The ERCO Ercoupe, a light two-seat aircraft known for its safety and simplified control system, uses wing ribs to maintain airfoil shape and transfer aerodynamic loads[23]. In this study, different cutout shapes—circular, elliptical, and rectangular—are introduced into the wing rib to reduce weight while maintaining structural integrity. The models are designed and analyzed using Autodesk Fusion 360, applying finite element analysis to evaluate stress distribution, strain, and deformation under loading conditions. The results show that while cutouts significantly reduce weight, they also create stress concentration near their edges; however, elliptical cutouts provide a more uniform stress distribution compared to other shapes. Overall, the study concludes that optimized cutout designs can enhance performance, reduce weight, and maintain safety in aircraft structures.

VI. ANALYSIS OF WING RIB STRUCTURE

This chapter presents the finite element analysis (FEA) of the aircraft wing rib to evaluate its structural behaviour under aerodynamic loading. The analysis focuses on determining the

stress distribution, deformation characteristics, and overall structural performance of the rib for different cut-out configurations. Fusion 360 is used as the simulation tool to perform static structural analysis.

This project analyzes the structural integrity of wing ribs for the ERCO Ercoupe aircraft by evaluating the impact of different lightening hole geometries circular, elliptical, and rectangular using Finite Element Analysis (FEA) in ANSYS. Utilizing Aluminium 7075 T6 material and applying an aerodynamic load of 40.2 MPa to the upper flange, the study determined that cut-out configuration significantly affects both displacement and stress concentration. The results revealed that while the original rib without cut-outs was the most stable with a displacement of 0.001 mm, the elliptical cut-out design proved to be the most efficient modification, showing the lowest stress (19.111 MPa) and most uniform load distribution among the lightened structures. Conversely, circular cut-outs experienced the highest stress (75.315 MPa) and displacement (26.892 mm), while rectangular shapes suffered from stress concentrations due to their sharp corners. Ultimately, the analysis confirms that optimizing wing rib geometry through elliptical configurations effectively reduces structural weight while maintaining the necessary strength and safety for aircraft performance.

Results of baseline model (without cut-out)

The rib without cut-outs represents the original structural configuration.

- Stress: 4.143 MPa
- Deformation: 0.001 mm
- Strain: 9.809×10^{-5}

The baseline model shown in the figure: 8 exhibits the lowest stress and deformation due to the absence of material removal. The stress distribution is uniform across the rib, indicating strong structural integrity.

However, this configuration is not efficient from a design perspective, as it results in higher weight. In aircraft structures, minimizing weight is equally important as maintaining strength.

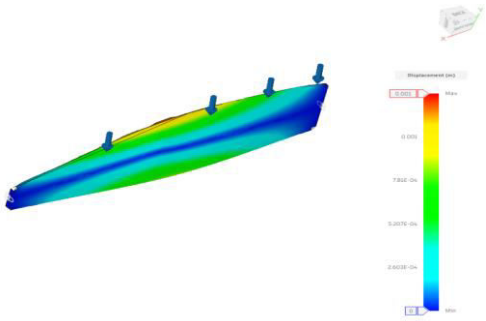


Fig:8, Displacement of wing rib of ERCO Ercoupe without cut

The strain analysis in the figure 9, below shows very low and mostly uniform strain across the airfoil section, indicating good stiffness. Slight increases occur near the load application points due to localized effects. The maximum strain remains within safe limits, confirming that the structure can withstand the applied loads without significant deformation.

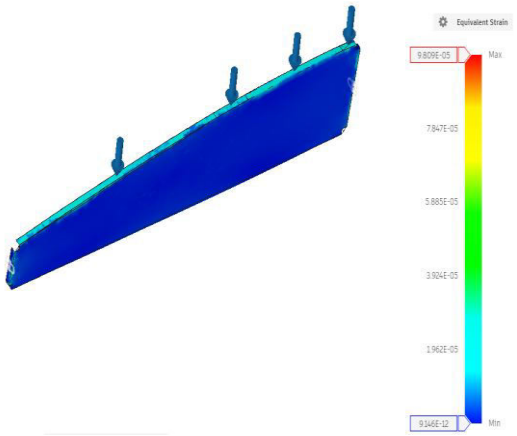


Fig:9, strain analysis of wing rib of ERCO Ercoupe without cut

The stress analysis in the figure 10, below shows a mostly uniform distribution of Von Mises stress across the airfoil section, indicating effective load transfer. Higher stress concentrations are observed near the load application points along the upper surface. The maximum stress remains within safe limits, confirming that the structure can withstand the applied loads without risk of failure.

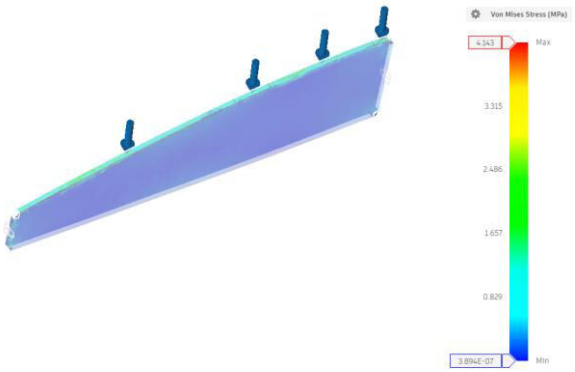


Fig:10, Stress on wing rib of ERCO Ercoupe without cut.

Results of elliptical cut-out

- Stress: 19.111 MPa
- Deformation: 2.905 mm
- Strain: 4.585×10^{-4}

The elliptical cut-out configuration in figure 11, shows the most balanced performance among all modified designs.The stress distribution is smoother due to the elongated shape, which aligns better with the direction of load transfer. This reduces stress concentration compared to circular and rectangular cut-outs. Deformation is moderate and significantly lower than the circular case, indicating better stiffness retention. This configuration achieves effective weight reduction while maintaining structural stability.

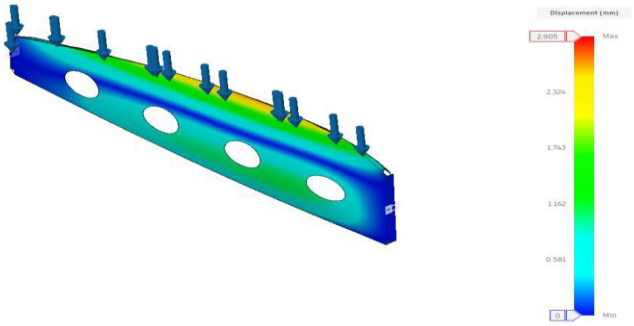


Fig: 11, Displacement of wing rib of ERCO Ercoupe with elliptical cutout

The strain analysis in the figure12, shows that the rib with material removal (lightening holes) experiences slightly higher strain compared to the solid model. The strain distribution remains mostly uniform (blue), indicating adequate stiffness, while localized increases are observed around the hole edges and near the load application points due to stress concentration effects. The maximum strain is higher but still within acceptable limits, demonstrating that the structure maintains its integrity despite weight reduction. This indicates an efficient design with a good balance between reduced weight and structural performance.

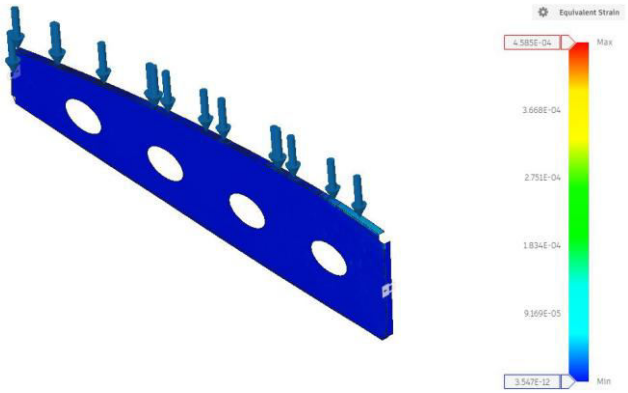


Fig:12, strain of wing of ERCO Ercoupe with elliptical cutout

The elliptical cut-out configuration in the figure 13 provides a highly efficient stress flow, with a peak Von Mises stress of 19.111 MPa concentrated primarily at the load points as shown in figure 13. This elongated geometry maintains superior stiffness and significantly lower deformation compared to circular designs by aligning material with the primary load path. Ultimately, it achieves an optimal balance of mass reduction and structural stability, making it ideal for weight-sensitive aeronautical applications.

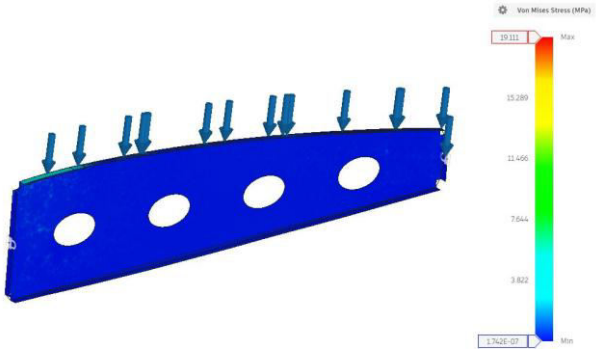


Fig 13, Stress of wing rib of ERCO Ercoupe with elliptical cutout

Results of circular cut-out

- Stress: 75.315 MPa
- Deformation: 26.892 mm
- Strain: 0.002

The introduction of circular cut-out significantly reduces the weight of the rib but leads to a considerable increase in stress and deformation shown in figure 14.

The stress is concentrated around the edges of the circular holes due to geometric discontinuity. Although circular shapes provide smoother stress flow compared to sharp geometries, the large increase in deformation indicates reduced stiffness.

This configuration is structurally acceptable only if the stress remains within allowable limits, but the high deformation may affect aerodynamic performance.

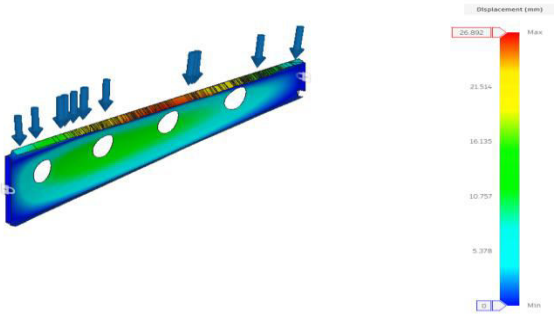


Fig 14 Displacement of wing rib of ERCO Ercoupe with circular cutout

The circular cut-out configuration in shows higher localized deformation, with a peak equivalent strain of 0.002 concentrated around the hole boundaries and load points as shown in figure 15. While it provides effective weight reduction, the geometry results in lower stiffness compared to the elliptical design, leading to more pronounced structural strain. This design represents a standard weight-saving approach but is less optimized for maintaining structural rigidity under high-stress aeronautical conditions.

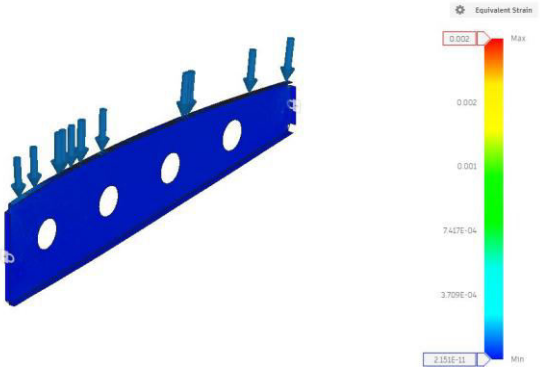


Fig:15 Strain of wing rib of ERCO Ercoupe with circular cutout

The circular cut-out configuration in experiences a significantly higher maximum Von Mises stress of 75.315 MPa, primarily concentrated at the load application points and hole edges as shown in figure 16. While this design provides substantial weight reduction, the circular geometry disrupts the internal load path more abruptly than the elliptical version, leading to higher stress intensities. Overall, this configuration remains stable but is less efficient at distributing loads, resulting in lower structural safety margins compared to the optimized modified designs.

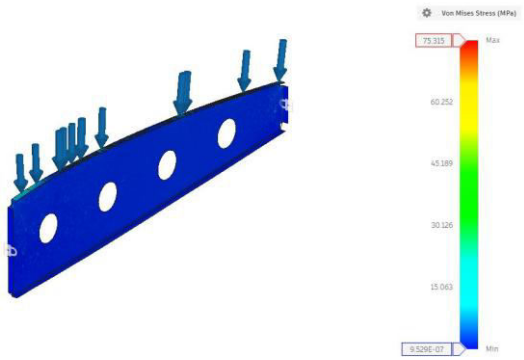


Fig:16, Stresses in wing rib of ERCO Ercoupe with rectangular cutout

Results of rectangular cut-out

- Stress: 28.94 MPa
- Deformation: 3.496 mm
- Strain: 4.974×10^{-4}

Rectangular cut-outs result in higher stress concentration, particularly at the corners, even when fillets are applied.

The presence of geometric discontinuities leads to localized stress intensification. Although deformation is moderate(fig 16), the stress levels are higher than elliptical configurations. This makes the rectangular cut-out less favourable for structural applications.

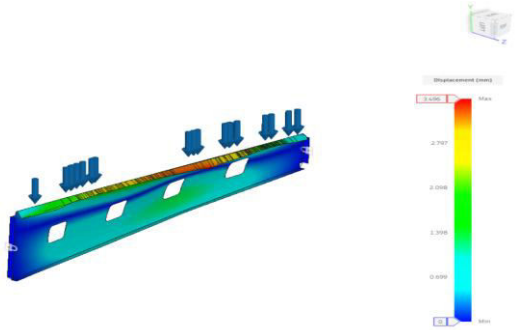


Fig: 17, Deformation of wing rib of ERCO Ercoupe with rectangular cutout

The rectangular cut-out configuration in the figure 18, shows higher localized deformation at the corners, where sharp geometry creates significant strain concentrations compared to the previous elliptical design. While it effectively reduces weight, the equivalent strain peak of 4.974E-04 indicates reduced stiffness and a greater risk of structural fatigue under the applied downward loads. Overall, this design provides a functional weight-to-strength ratio but lacks the optimized load-path distribution found in more rounded geometries.

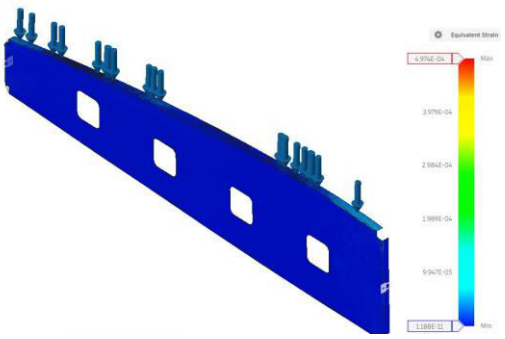


Fig:18 Strain of wing rib of ERCO Ercoupe with rectangular cutout

The rectangular cut-out configuration in the figure 19, exhibits a maximum Von Mises stress of 28.942 MPa, which is notably higher than the elliptical case due to significant stress concentrations at the sharp corners. These geometric discontinuities disrupt the load path, leading to higher localized stress peaks despite effective overall weight reduction. Consequently, while the design is functional, the increased stress at the corners suggests a higher susceptibility to fatigue and reduced structural efficiency compared to rounded alternatives.

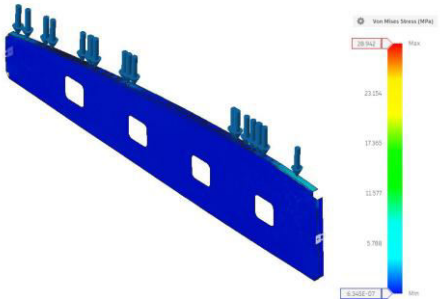


Fig:19, Stress of wing rib of ERCO Ercoupe with rectangular cutout

Table: 1; Comparison of result analysis

Types of geometries	*Displacement*	*Stress*	*Strain*
Wing rib without cutout shapes	0.001mm	4.143MPa	9.809E-05
Wing rib with elliptical cutout shapes	2.905mm	19.111MPa	4.585E-04
Wing rib with circular cutout shapes	26.892mm	75.315MPa	0.002
Wing rib with rectangular cutout shapes	3.496mm	28.94MPa	4.974E-04

VII. RESULT AND DISCUSSION

This chapter presents a detailed analysis of the results obtained from the finite element simulation of the aircraft wing rib. The study evaluates the influence of different cut-out geometries on structural performance under identical loading and boundary conditions. The results are interpreted in terms of stress distribution, deformation characteristics, and structural efficiency to identify the most suitable configuration for weight reduction while maintaining structural safety.

The baseline wing rib without cut-outs exhibited minimal deformation of 0.001 mm and a low stress value of 4.143 MPa, indicating high structural strength and stiffness. However, this configuration results in higher weight, making it less efficient for practical aircraft applications where weight reduction is a critical design requirement. The introduction of cut-outs significantly

alters the structural behaviour of the rib. Among the configurations analysed, the elliptical cut-out demonstrated superior performance, with a deformation of 2.905 mm and a stress of 19.111 MPa. The smooth curvature of the elliptical geometry allows gradual stress flow, thereby reducing stress concentration and maintaining structural integrity. This results in an efficient balance between weight reduction and strength.

In contrast, the circular cut-out configuration resulted in a deformation of 26.892 mm and a maximum stress of 75.315 MPa, indicating a considerable reduction in stiffness. Although circular holes are commonly used due to their simplicity and uniform geometry, the larger deformation and higher stress values suggest that this configuration is less efficient compared to the elliptical design. Similarly, the rectangular cut-out configuration showed a deformation of 3.496 mm and a stress of 28.94 MPa, with stress concentration occurring at the corners due to geometric discontinuities, even when fillets are provided.

All models were analysed under identical conditions, including the same material properties, boundary conditions, applied pressure, and rib geometry, ensuring that the variations in results are solely due to cut-out geometry. Based on the comparative analysis, it is evident that cut-out geometry plays a crucial role in determining structural performance. Among all configurations, the elliptical cut-out provides the most optimal balance between strength, deformation, and weight reduction, making it the most suitable design for the aircraft wing rib.

REFERENCES

[1] Ajith V S, Mohamed Sameer and K Kavin Kumar., "Design and structural analysis of wing-fuselage bracket of a UAV," *Int. J. Creat. Res. Thoughts*, vol. 5, no. 4, pp. 896–903, 2017, doi: 10.6084/m9.figshare.17163473.v1.

[2] T. Raju, K. Kishore Kumar, and M. Jawahar, "Design of an aircraft wing structure for static & fatigue life prediction," *Int. J. Recent Technol. Eng.*, vol. 8, no. 2 Special Issue 11, pp. 2614–2618, 2019, doi: 10.35940/ijrte.B1316.0982S1119.

[3] R. Sharma and G. Garg, "Design and Analysis of Wing Rib of Aircraft Review," *Int. J. Innov. Res. Sci. Eng. Technol.*, vol. 03, no. 12, pp. 17999–18005, 2014, doi: 10.15680/ijirset.2014.0312033.

[4] A. V. S and J. K. G, "Hybrid deep learning for object detection in drone imagery: a new metaheuristic based model," *Multimed. Tools Appl.*, vol. 83, pp. 8551–8589, 2023, doi: https://doi.org/10.1007/s11042-023-15785-0.

[5] R. Balaji, E. R. Talasila, L. Oblisamy, V. S. Ajith, and A. Basithrahman, "Design and Structural Analysis of Nano-satellite," in *Recent Advances in Manufacturing, Automation, Design and Energy Technologies: Proceedings from ICoFT 2020*, Springer, 2022, pp. 515–521.

[6] V. S. Ajith, P. R. Ragesh, and K. Shine, "Digital Twin–CPS Synergy in Manufacturing: A Pathway to Intelligent and Sustainable Industry 5.0," in *Manufacturing Process Optimization for Sustainable Development Using Digital Twin Applications*, 1st ed., CRC Press, 2026, ch. 10, pp. 159–168. doi: 10.1201/9781003504825-10.

[7] R. P. R. Ajith V. S. Shine K., Khaled Shahin, Vishnu Vijay Kumar, "Carbon Footprint Reduction Strategies Using Digital Twin Integration Strategy," in *Manufacturing Process Optimization for Sustainable Development Using Digital Twin Applications*, 1st ed., CRC Press, 2026, ch. 5, pp. 86–99. doi: 10.1201/9781003504825-5.

[8] S. Bairavi, "Design And Stress Analysis of Aircraft Wing Rib With Various Cut Outs," vol. 358, no. 4, pp. 511–514, 2016.

[9] J. Muthuraman, M. Shankar, and M. Vivekanandhan, "Design and Analysis of Wing Rib Using Finite Element Method," *Int. Conf. Emerg. Trends Eng. Technol. Res.*, pp. 55–62, 2019.

[10] Ajith V S, M. S. Shihab, A. S. Nath, and M. Shamil Mashhood, "Design and analysis of batwing for lift enhancement," 2024, [Online]. Available: www.ijrar.org

[11] A. Bharath, E. Keshav Kumar, V. M. Sreehari, and B. Ravikumar, "Design and analysis of an aircraft wing structure," *Test Eng. Manag.*, vol. 82, no. June, pp. 6952–6958, 2020, [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083427484&partnerID=40&md5=3691b948e1779536662452880c7eb91d

[12] m. Kondal rao savidiboina, kota venkateswarlu, "design and analysis of a typical wing rib for passenger aircraft," vol. 10, no. 11, pp. 560–564, 2019.

[13] V. Kalavagunta and S. Hussain, "Wing Rib Stress Analysis of DLR-F6 aircraft," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 455, no. 1, 2018, doi: 10.1088/1757-899X/455/1/012033.

[14] R. J. S. Rohit Kumar, Syed Mohsin, "Review of Wing Rib Design and Analysis of Aircraft," *J. Emerg. Technol. Innov. Res.*, vol. 07, no. 06, pp. 1711–1719, 2014, doi: 10.15680/ijirset.2014.0312033.

[15] Kannan and M. Veeranjaneyulu, "Structural modeling and analysis of composite wing rib using finite element METHOD," *Int. J. Eng. Sci. Manag. Res.*, no. 2, pp. 20–41, 2015, [Online]. Available: http://www.ijesmr.xn--com-1ea

[16] M. S. T. K. Kavin kumar K, Ajith V.S, Selvakumari G, "Mechanical Characterisation of Glass/Epoxy,Carbon/Epoxy, and Hybrid Composite," *Int. J. Creat. Res. Thoughts*, vol. 5, no. 4, pp. 990–1003, 2017, doi: https://doi.org/10.6084/m9.figshare.17056979.

[17] A. VS and K. Jolly, "Unmanned aerial systems in search and rescue applications with their path planning: a review," in *Journal of Physics: Conference Series*, IEEE, Nov. 2021, pp. 1–12. doi: 10.1088/1742-6596/2115/1/012020.

[18] V. S. Ajith, K. Vidhya, M. S. TK, K. Karunakaran, and N. Rajeev, "Enhancing UAV Endurance and Operational Efficiency in Search and Rescue Missions through Hybrid Fuel Cell and Battery

Systems.," *Aerosp. Traffic Saf.*, 2026.

[19] D. E. P. E. Naveen, M. Priya, N. Maheswaran, Ajith V S, "Acetylation chemical treatment for improve the surface of natural fiber," *Int. Res. J. Nat. Appl. Sci.*, vol. 3, no. 12, pp. 121–127, 2016, doi: https://doi.org/10.6084/m9.figshare.17162750.

[20] Y. Shwetha, G. S. D. Madhav, and D. Govardhan, "Buckling analysis of carbon epoxy trailing edge wing RIB," *Int. J. Mech. Prod. Eng. Res. Dev.*, vol. 8, no. 2, pp. 659–670, 2018, doi: 10.24247/ijmperdapr201877.

[21] Ajith vs, D. R. Paramasivam, and K. Vidhya., "Study of Optimal Design of Spar Beam for the Wing of an Aircraft," *Int. J. Eng. Dev. Res.*, vol. 5, no. 3, pp. 179–193, 2017, doi: 10.6084/m9.figshare.17162744.v1.

[22] S. Noman, M. Siddique, A. Z. Immad, B. R. Kumar, and D. Ravi, "Design Analysis of Air Craft Wing RIB," pp. 5051–5058, 2017, doi: 10.15680/IJIRSET.2017.0603306.

[23] K. Asghari, M. Boroji, and N. Salehi, "Structural Design and Analysis of Aircraft Wing Box : Load Distribution , Strength , and Buckling Assessment for NACA 23015 Airfoil Configuration," vol. 2, 2022.