e-ISBN: 978-81-973215-9-7
Doi: 10.17492/jpi.pcet.022501

Frontiers of Innovation

Exploring Emerging Technologies and Computing

Editors

Dr. Archana Chaugule, Dr. Archana Kollu and Dr. Vaishali Latke

Publisher

Journal Press India

Frontiers of Innovation: Exploring Emerging Technologies and Computing

Edited by

Dr. Archana Chaugule

Head of the Department, Department of Computer Engineering, PCET's Pimpri Chinchwad College of Engineering and Research, Pune, Maharashtra, India

Dr. Archana Kollu

Assistant Professor, R&D Coordinator, Department of Computer Engineering, PCET's Pimpri Chinchwad College of Engineering and Research, Pune, Maharashtra, India

Dr. Vaishali Latke

Assistant Professor, Department of Computer Engineering, PCET's Pimpri Chinchwad College of Engineering and Research, Pune, Maharashtra, India

Publication and Editorial Services facilitated by

Journal Press India
Delhi, India

Book Title: Frontiers of Innovation: Exploring Emerging Technologies and Computing

Editors: Dr. Archana Chaugule, Dr. Archana Kollu and Dr. Vaishali Latke

Online e-published by: JOURNAL PRESS INDIA
Publisher's address: A-4/17, 1st Floor,
Sector-15, Rohini,
Delhi - 110 089, India
Mobile: 8826623730; Tel: +91-11-42631517
Website: www.journalpressindia.com
E-mail: info@journalpressindia.com

First Edition, February 2025

e-ISBN: 978-81-973215-9-7

About the Book

“Frontiers of Innovation: Exploring Emerging Technologies and Computing” is a comprehensive compilation of cutting-edge research and insights presented at the International Conference on Emerging Technologies and Computing (ICETC 2025). This book serves as a vital resource for understanding the transformative role of advanced technologies such as artificial intelligence, machine learning, information security, quantum computing, and the Internet of Things in reshaping industries and society. Each chapter is carefully curated to reflect the dynamic and interdisciplinary discussions at the conference, showcasing pioneering ideas, innovative solutions, and practical applications that address real-world challenges in sectors like healthcare, agriculture, finance, and manufacturing.

In the context of India’s remarkable technological evolution, fueled by initiatives like Digital India, the National AI Strategy, and the Smart Cities Mission, this book highlights the nation’s strides in adopting and advancing emerging technologies. It explores critical themes such as cyber security, ethical implications, digital infrastructure development, and sustainable innovation, aligning with the conference’s objective to foster collaboration and drive technological progress.

Contributors to this book include researchers and academics whose work underscores the importance of addressing pressing issues while unlocking opportunities in the rapidly evolving tech landscape. By providing insights into groundbreaking developments and fostering a deeper understanding of emerging technologies, this book equips readers with the knowledge to navigate and contribute to the future of computing.

“Frontiers of Innovation” is not just an academic resource but also a guide for innovators, policymakers, and students keen on exploring the frontiers of technology. It captures the essence of the conference by promoting knowledge-sharing, inspiring collaboration, and catalyzing India’s leadership in the global technological arena. This publication is an indispensable tool for anyone seeking to understand and participate in the ongoing revolution in computing and innovation.

About the Editors

Dr. Archana Chaugule, awarded a Ph.D. in 2017, is a recognized researcher and educator with significant contributions to intellectual property and higher education. She has developed a popular lecture series on Discrete Mathematics on Udemy. She received the Best Researcher Award from DK International Research Foundation in 2018 and the Institute of Scholars in 2019 and was honored by PCCOER for her role in coordinating a record-breaking event where 154 copyrights were filed in one day in January 2018. Dr. Chaugule serves as a Technical Committee member for ICRSA 2021, an editorial board member for IGI Global's book on data-driven UI design, and as a Board of Studies member at Pillai College. She has conducted a Facebook live session on copyright filing and achieved "Elite" in NPTEL's Machine Learning course. Her academic portfolio includes 52 research publications, 8 copyrights, 1 granted patent, and 6 filed patents.

Dr. Archana Kollu is an Assistant Professor and R&D Coordinator in the Department of Computer Engineering, PCET's Pimpri Chinchwad College of Engineering and Research. She was awarded a Ph.D. in Computer Science and Engineering in 2022 by Koneru Lakshmaiah Education Foundation (Deemed to be University). With over 20 years of academic and research experience, Dr. Kollu specializes in Artificial Intelligence, focusing on Evolutionary Computing, Computational Intelligence, Energy Informatics, and Machine Learning. She has an extensive portfolio of publications in international journals, conferences, and symposiums. Dr. Kollu also contributes to the academic community as a reviewer for reputed journals and conferences, demonstrating her commitment to advancing research and innovation in her field.

Dr. Vaishali Latke is an Assistant Professor and Academic coordinator in the Department of Computer Engineering, PCET's Pimpri Chinchwad College of Engineering and Research. With over 15 years of experience in engineering education, she has made significant contributions to both teaching and research. In 2024, Dr. Latke completed her PhD in Computer Engineering from D.Y. Patil Deemed to be University, Nerul, Mumbai. Her research interests are well-regarded, with several of her articles being published in prestigious SCI and Scopus indexed journals. She has also actively participated in numerous national and international conferences, further establishing her expertise in the field. Her academic journey and contributions highlight her commitment to advancing the field of computer engineering.

Contents

Preface

In an era defined by rapid technological advancements, the exploration of emerging technologies and computing has never been more critical. This edited volume compiles a diverse range of technical papers that delve into the forefront of innovation, addressing the significant challenges and opportunities presented by the digital revolution. The contributions within this book highlight the transformative power of technologies such as artificial intelligence, machine learning, the Internet of Things (IoT), and Robotics, showcasing their applications across various sectors including healthcare, finance, agriculture, and manufacturing.

The purpose of this book is to provide a comprehensive overview of the current landscape of emerging technologies and computing, offering insights from leading experts, researchers, and practitioners in the field. Each chapter presents rigorous research and case studies that illuminate the practical implications of these technologies, as well as the ethical, social, and economic considerations that accompany their implementation. We believe that fostering a deeper understanding of these topics is essential for academics, industry professionals, and students alike as they navigate an increasingly complex technological environment.

We are grateful to all the authors who contributed their expertise and research to this volume. Their commitment to advancing knowledge in this dynamic field is reflected in the quality and depth of the papers included here. We also extend our appreciation to the reviewers for their invaluable feedback, which has enhanced the overall rigor and clarity of the content.

As you embark on this journey through the pages of this book, we hope you find inspiration and insight that will not only inform your understanding of emerging technologies but also encourage you to engage with these innovations in your academic and professional endeavors. Together, let us explore the endless possibilities that lie ahead in the realm of technology and computing.

Editors

Dr. Archana Chaugule
Dr. Archana Kollu
Dr. Vaishali Latke

Acknowledgments

We extend our deepest gratitude to everyone who contributed to the success of the International Conference on Emerging Technologies and Computing (ICETC 2025) and the publication of the book, *Frontiers of Innovation: Exploring Emerging Technologies and Computing*. This achievement is the result of the collective efforts and unwavering support of numerous individuals and teams.

First, we express our sincere thanks to the PCET's management for their visionary leadership, encouragement, and constant support. Their commitment to fostering innovation and academic excellence laid the foundation for this conference and the creation of this book.

We are profoundly grateful to our respected principal Prof. Dr. H.U. Tiwari for his guidance and support, which have been instrumental at every step of this journey. His belief in the importance of this initiative has inspired everyone involved.

A special note of appreciation is extended to our keynote speaker for his insightful and thought-provoking address, which enriched the conference and inspired meaningful discussions among participants.

We sincerely thank the authors for their invaluable contributions, which form the essence of *Frontiers of Innovation*. Your dedication to advancing research and sharing knowledge has greatly enhanced the academic and practical relevance of this publication.

Our heartfelt appreciation goes to the reviewers for their meticulous evaluation and constructive feedback. Your efforts have been essential in maintaining the high standards of the research presented at the conference and in the book.

We are also deeply grateful to the advisory committee for their invaluable guidance and strategic input, which have been instrumental in shaping the direction and success of both the conference and the publication.

Finally, we recognize the organizing team for their tireless efforts behind the scenes. Your dedication and expertise ensured the seamless execution of the event and the successful production of this book.

To everyone who contributed, thank you for making ICETC 2025 and *Frontiers of Innovation* a resounding success.

Editors

Dr. Archana Chaugule
Dr. Archana Kollu
Dr. Vaishali Latke

CHAPTER 1

A System to Detect and Monitor Real-time Soil Fertilization using IOT

Rudragouda Patil and Kanchan Pradhan***

ABSTRACT

India's economy is based on agribusiness. In India, agribusiness is essential to completing 50% of the current work. Agriculture in India has a greater level of commitment than some other sectors. In any event, farmers developed harvests using traditional methods, which results in lower yield profitability. Additionally, a primary factor contributing to lower crop profitability is soil erosion and disintegration. This will result in decreased fruitfulness. The development of a plant depends on nutrients like potassium (K), nitrogen (N), and phosphorus (P). IoT devices might make this possible. Farmers may obtain the required information as soon as the screen yields. The Internet of Things (IoT) links the entire universe through sensors and other installed devices. A pH sensor is going to be used to compute the soil values after subtracting the nutrient value from the various soil samples that were obtained from different fields. At that time, depending on the information or data, it decomposed, and finally, it started to see signals of increasing crop output. Farmers may efficiently monitor their land and soil by utilizing the proposed system. Instead of going to the lab for soil testing, the farmer may obtain comprehensive information on the soil from their own homes.

Keywords: NPK; pH Sensors; Internet of Things; Recommendations.

1.0 Introduction

Farming is needed for human life, and IOT is a better stage of savvy agribusiness. With the expansion in the populace, there is a requirement for nourishment creation. The use of IoT in agriculture is steadily growing. For savvy agriculture alongside supplements, innovation assumes a key job for the manageable cultivating and diminished natural impacts and financial misfortunes. Various soil samples were gathered from different fields, and a pH sensor will be used to compute the soil values after deducting the nutrient value (Padole *et al.*, 2017). Yield creation is, for the most part, dependent on soil properties, and harvest development relies upon supplements accessible in the soil.

**Corresponding author; Assistant Professor, Department of Computer Engineering, Marathwada Mitramandal's College of Engineering, Pune, Maharashtra, India (E-mail: patilrudrag@gmail.com)*
***Assistant Professor, Department of Computer Engineering, MIT Academy of Engineering, Pune, Maharashtra, India (E-mail: mailme.kanchan27@gmail.com)*

Appropriate soil testing and observation will cause the farmers to get great harvests dependent on the supplements present in the soil (Agraj Aher *et al.*, 2018). The fundamental or significant supplements accessible in the soil are nitrogen (N), potassium (K), and phosphorous (P). Root retains the required measure of supplements and water from the soil. With each collected harvest, the richness level of the yield diminishes (Mohanraj *et al.*, 2017). To expand crop richness level, estimation of ripeness level is essential for a farmer to develop great harvests. With the estimation of soil supplement focus; this worth will be useful to build crop fruitfulness. The soil NPK sensor is designed to measure soil's nitrogen, phosphorus, and potassium levels, helping evaluate soil fertility and overall health. This sensor allows for a detailed assessment of soil conditions. It can remain in the soil for extended periods, thanks to its durable probe, which is resistant to rust, electrolytic effects, and corrosion from salts and alkalis, making it suitable for various soil types, with alkaline, acidic, substrate, seedling beds, and coconut bran soils.

Because IoT solutions guarantee excellent yields, financial success, and environmental sustainability, they are essential in helping farmers narrow the supply-demand gap. IoT technology is used in precision agriculture to maximize resource utilization, which improves agricultural yields and lowers expenses. IoT in agriculture comprises specialized machinery, wireless communication, software, and IT services. A survey by "BI Intelligence" reveals that by 2020, 75 million IoT devices were expected to be used in agriculture, with an annual growth rate of 20%. The global market for smart agriculture is projected to grow from just over $5 billion in 2016 to $15.3 billion by 2025. IoT-based smart farming allows farmers to reduce waste and enhance productivity across several areas, such as fertilizer usage, farm vehicle trips, and resource efficiency (e.g., water and energy). These smart farming systems use sensors to automate irrigation and monitor fields by measuring factors like light, humidity, temperature, soil moisture, and crop health. Using data from remote field monitoring, farmers can decide whether to take automatic or manual action. Example: Sensors can trigger irrigation when soil moisture levels fall, making intelligent farming highly efficient.

2.0 Literature Review

A smart agriculture stick promotes the innovative IoT-based smart agriculture program. It gives farmers access to real-time data to monitor the state of the environment, allowing them to practice smart farming, boost their final yield, and guarantee that their product satisfies quality standards (Dhanaraju *et al.*, 2022).

The plan addresses the negative effects of excessive fertilizer use on the environment by introducing an Internet of Things-based Plant Nutrient Control System. The system provides real-time data on soil nutrient concentrations using NPK sensors and is saved in a cloud service such as Google Firebase. In addition to monitoring soil temperature and moisture and notifying farmers when necessary, the system analyzes the data to identify the minerals required for healthy plant growth. Kodular API is used to develop a user-friendly interface (Anjaneyulu *et al.*, 2024). To cut down on fertilizer use and resource loss in agriculture, researchers have created a water-fertilizer management system that uses inexpensive wireless sensor networks. Water fertilizer application is guided by the system using data on soil conductivity thresholds. In comparison to conventional systems, the tested method saved 0.76–0.87 tons of fertilizer during cotton growth, a 10.89% reduction in fertilizer usage. This creative method can lessen soil pollution and farmers' workload (Du *et al.*, 2021).

In 2022, the author (Blesslin Sheeba *et al.*, 2022) conducted a study examining the nutrient composition of mulberry gardens in Tamil Nadu, categorizing the soil based on fertility and pH levels. The research employed the extreme learning machine (ELM), an efficient classification technique, to detect micronutrients in the soil. The findings revealed that Tamil Nadu soils typically have normal electrical conductivity, are red in color, and are high in potassium, nitrogen, and sulfur, while deficient or adequate in magnesium, boron, zinc, and copper.

In 2021, authors (Hassan *et al.*, 2021) found that the increasing population and growing food demand necessitate automation in agriculture. Traditional farming methods are inefficient, leading to land degradation and a lack of fertility. This study examines various automation control strategies, including IoT, aerial imagery, machine learning, and artificial intelligence. These technologies offer solutions to challenges such as plant diseases, pesticide management, weed control, irrigation, and water conservation. By incorporating advanced automation, crop yield and soil fertility can be enhanced.

In 2020, authors (Raut *et al.*, 2020) Agribusiness is India's foundation, accounting for 50% of the economy. Traditional farming methods lead to lower yield profitability and soil erosion, affecting fruitfulness levels. Loss of essential nutrients like potassium, nitrogen, and phosphorus is also a concern. Advancements in agriculture, such as IoT gadgets, can help address these issues and promote smart cultivation, allowing farmers to access necessary data remotely.

In 2020, authors (Patel *et al.*, 2020) investigated the application of deep learning networks to estimate urea fertilizer levels in soils using spectroradiometer data. The method was applied to silt clay and loamy soil types, identifying optimal absorption features at 1899.2 nm for urea and 2195.1 nm for the soils. The findings demonstrated higher accuracy compared to using DASU alone, and the hyperspectral imaging technique proved effective for assessing soil fertility in situ on agricultural lands. In 2019, the authors (Puranik *et al.*, 2019) observed that we now live in an increasingly digital world, where nearly whole things are influenced in technology.

The impact of technology on agriculture is growing increasingly evident over time. Since its early days, communication has played a crucial role in agriculture, not only in crop diagnostics but also in transforming traditional agricultural practices. Over time, various methods and technologies have been developed and integrated into farming systems. Nevertheless, India's agricultural sector has been facing a decline, leading to a reduction in its production capacity. There is an urgent need to address these challenges and restore vitality to the sector, driving it toward higher growth. Large-scale agricultural systems require significant upkeep, expertise, and oversight. This paper focuses on automated management of pesticides, water usage, and crop monitoring to improve efficiency and sustainability.

In 2018, authors (Lavanaya *et al.*, 2018) found out soil nutrients are crucial for crop yield and quality. Traditional methods in India are insufficient for population growth. This work proposes automated irrigation using pH sensors, Arduino, Zigbee, IoT, and Wireless Sensor Networks. The system aims to select seeds based on pH value, conserve water, maintain greenhouse temperature, and choose appropriate fertilizers to avoid overfeeding or underfeeding.

3.0 Existing System

Availability of IoT Infrastructure: The project assumes the availability of the necessary IoT infrastructure, including sensor networks, communication protocols, and data management systems. It relies on the existence of a reliable and stable IoT platform or network to collect, transmit, and analyze sensor data.

Sensor Accuracy and Reliability: The project assumes that the selected IoT sensors for measuring soil parameters like moisture, temperature, pH levels, and nutrient concentrations are accurate and reliable. It is crucial to choose sensors that provide consistent and precise measurements to ensure the effectiveness of the soil fertilization system (Sukumar *et al.*, 2018).

Internet Connectivity: The project relies on the presence of a stable and dependable internet connection in the farming region. This is crucial for real-time sensor data transmission, accessing remote monitoring and control platforms, and receiving updates or suggestions from the IoT system.

Power Supply: The project assumes a reliable power supply for the IoT infrastructure, including sensors, gateways, and communication devices. It is important to have a continuous power source to ensure uninterrupted operation of the IoT system and avoid data loss or system downtime (Jayaprahas *et al.*, 2014).

Calibration and Maintenance: The project assumes regular calibration and maintenance of the IoT sensors to ensure their accuracy and reliability. Calibration should be performed at appropriate intervals to account for sensor drift and to maintain the quality of the collected data (Abdullah Ahmad *et al.*, 2016)

Crop-specific Requirements: The project assumes that the fertilizer requirements and nutrient recommendations for the specific crops being grown are well-defined and available. Accurate knowledge of crop nutrient needs at different growth stages is necessary to provide appropriate fertilization recommendations through the IoT system. Regulatory Compliance: The project assumes compliance with relevant regulations and guidelines related to soil fertilization and environmental management. It is important to consider local regulations and best practices to ensure the IoT-based fertilization practices align with legal requirements and environmental sustainability standards.

Data Privacy and Security: The project assumes the implementation of appropriate measures to ensure the privacy and security of the collected data. This includes encryption of data during transmission, secure storage of data, and adherence to data protection regulations.

The existing system primarily addresses two categories of soil testing techniques.

- Soil testing in laboratory
- Mobile soil testing

The first method involves laboratory analysis of soil (http://soilquality.org), which is time-consuming and may take several days or weeks to determine soil fertility. This process requires individuals to collect soil samples and send them to labs, where chemical analysis is used to measure NPK levels. Mobile soil testing, on the other hand, involves on-site testing and fertilizer recommendations. However, this approach is typically conducted only once per crop cycle, making it less effective for optimal crop production and often yielding less accurate results. Subsequently, three advanced methods such as spectroscopy, conductivity testing, and electrochemical sensors are adopted, but they have been developed to assess soil fertility. While these methods are more modern, they are expensive and also lack precision.

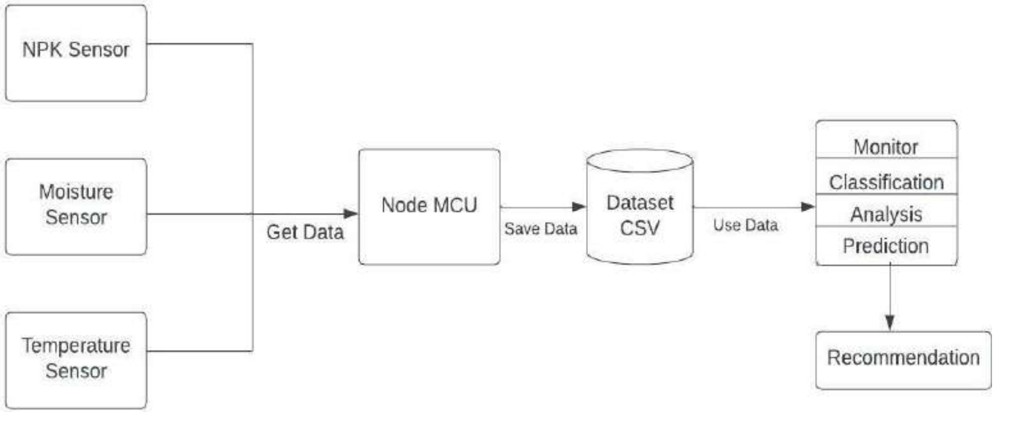
4.0 Proposed System

Soil fertilization Detection and Monitoring System aims to address these challenges by leveraging IoT technologies, including sensors, data analytics, and remote-control capabilities. The goal is to enable precise, efficient, and data-driven soil fertilization practices, optimize resource usage, improve environmental sustainability, enhance crop productivity, and facilitate informed agricultural decision-making.

4.1 System architecture

Our proposed Architecture as shown in Figure 1 helps to recommend the appropriate fertilizer to enhance soil fertility. It uses various sensors, such as NPK, temperature, and moisture sensors, to collect data for these suggestions. A Node MCU is employed to collect the sensor data, store it in a CSV file, clean the data, and perform the necessary analysis to generate the recommendations. The data collected from sensors are converted into csv data and are continuously pushed to a server. Using this data, analysis and recommendation to farmer is developed. And for more information the graph is drawn for visualization. The data from NPK, moisture, and temperature sensors is essential for managing soil fertilization effectively. NPK sensors classify nutrient levels nitrogen, phosphorus, and potassium as high, medium, or low based on recommended thresholds. Moisture sensors determine if soil conditions are dry, optimal, or waterlogged, ensuring proper irrigation and nutrient absorption. Temperature sensors evaluate soil warmth, categorizing it for its suitability in supporting nutrient uptake and microbial activity.

Figure 1: Architecture Diagram of Soil Fertilization Monitoring System



This data is analyzed to uncover correlations between soil conditions and plant health. Nutrient levels are examined for deficiencies or surpluses, while moisture content is assessed for its influence on nutrient availability. Temperature variations are studied to understand their impact on biological processes and plant growth. The analysis helps predict trends, such as future nutrient levels, watering needs, and potential plant growth outcomes.

Based on these insights, precise recommendations are made to optimize soil health and crop yield. Fertilizer adjustments are suggested to address nutrient imbalances, while tailored irrigation schedules ensure optimal soil moisture. Additional advice on soil amendments, like adding compost or lime, is provided to enhance soil fertility and support sustainable farming practices.

The system is energy-efficient and user-friendly. It records data with precise timestamps and dates. The system is designed for real-time interaction; it seamlessly interacts with framers to provide timely recommendations whenever needed. The system has been proposed to forecast optimum soil fertilization by considering soil parameters such as pH value, and nutrients.

4.2 Novelty

The Soil Fertilization Monitoring System combines real-time sensor data collection and advanced machine learning algorithms to provide adaptive fertilization recommendations tailored to crop-specific needs. The system dynamically checks at several parameters to give farmers a piece of advice. A NODE MCU processor as shown in Figure 2a is used to collect real-time data from NPK sensors, moisture sensors, and temperature sensors. The information is kept in a safe cloud server in

CSV format, so it can be accessed from anywhere and used to look at past trends. The data is put through an XGBoost classification algorithm, which determines the best amount of fertilizer (NPK) that the soil and surroundings are currently in. The system tells you exactly how much NPK and water you need based on its estimates.

5.0 Results and Discussion

5.1 Experimental procedure

1. Divide the plants into two groups: one group using the IoT fertilization system and another group with traditional manual fertilization.
2. Set up the IoT system to deliver fertilizer based on the sensor data, ensuring appropriate nutrient levels for optimal plant growth.
3. Follow a consistent watering schedule for both groups to maintain consistent soil moisture levels.
4. Consistently observe and document plant growth metrics, including height and leaf development such as color, flowering and fruiting.
5. Measure and record fertilizer consumption for both groups.

5.2 Expected results

1. *Improved Plant Growth:* The plants in the IoT fertilization group are expected to exhibit healthier growth, including taller height, greener leaves, and more abundant flowers or fruits.
2. *Optimal Fertilizer Usage:* The IoT system is expected to deliver fertilizers more efficiently, adjusting nutrient levels based on real-time sensor data, and reducing fertilizer waste and cost.

5.3 Data analysis

Figure 2a: System Connection

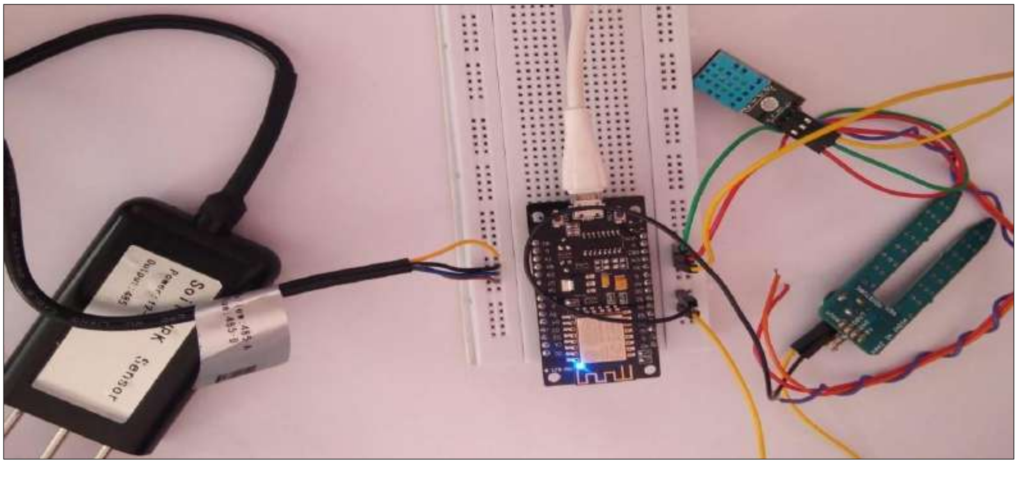


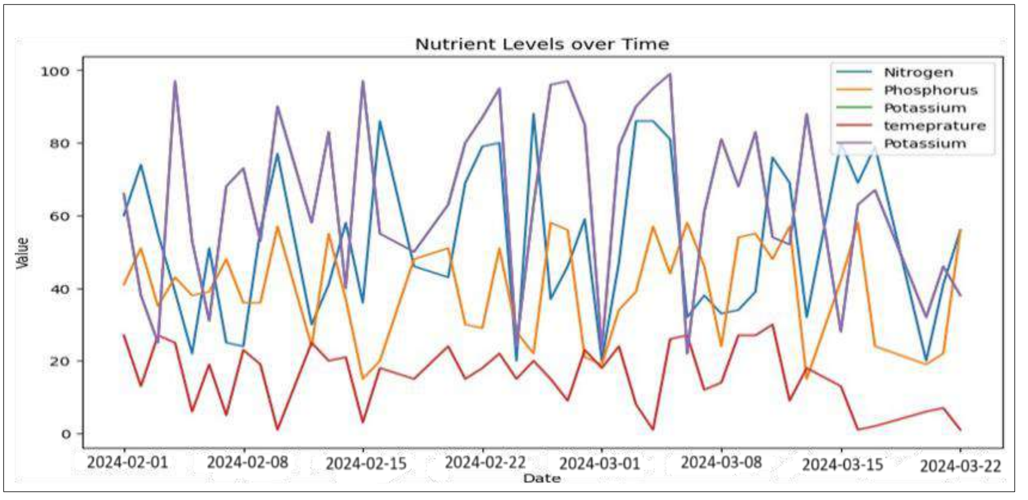
Figure 2b: System Implementation

Figure 2b depicts a NodeMCU with NPK sensors deployed in the field, gathering real-time data into a cloud database. Analyzing the collected data using the cloud-based platform can provide insights into the relationship between soil conditions, fertilizer applications, and plant growth. Receive graphs in the results that display all the nutrients and individual graphs for each element with their recommended range of values. Finally, a bar graph was used to analyze which element had the lowest value during the previous 30 days.

5.4 Discussions on system results

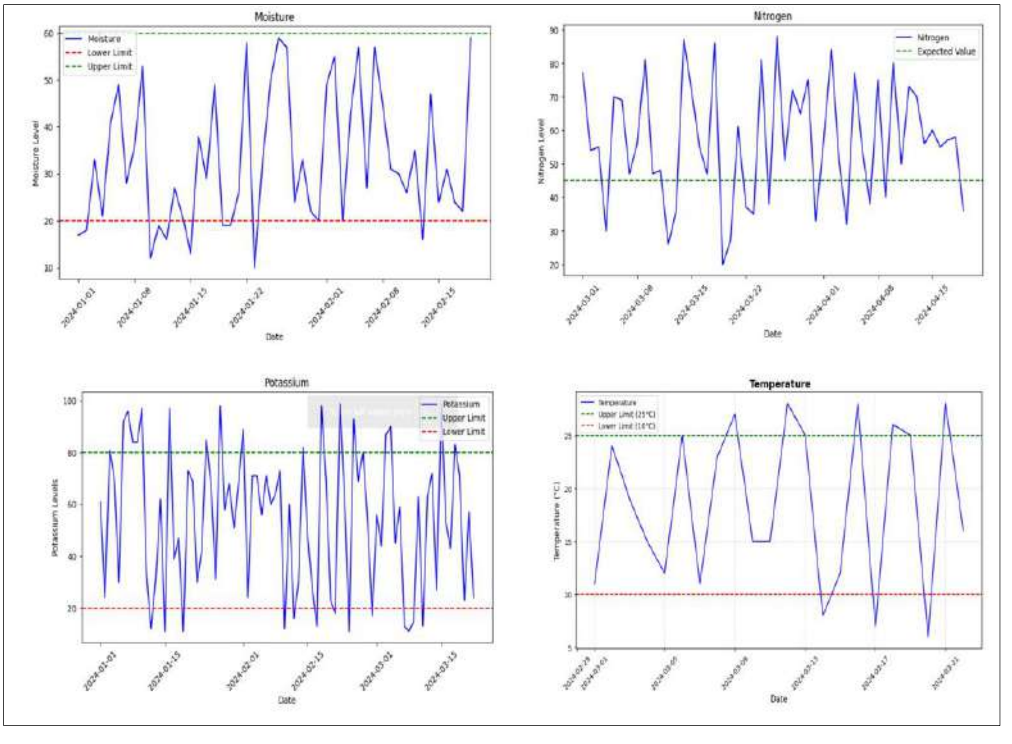
In this section, the results of the experiment as shown in Figure 2 (a & b) are observed through a graph plot. The details of the observation are explained further. The graph in Figure 3 depicts nutrient levels NPK (nitrogen, phosphorus, and potassium) and temperature variations over time from February 1, 2024, to March 22, 2024. The following observations can be made, Nitrogen levels show significant fluctuations throughout the period, reaching peaks above 80 and dips near 20, indicating varied nutrient availability. Phosphorus levels are relatively more stable compared to Nitrogen but still vary between 20 and 60, suggesting periodic adjustments or uptake. Potassium levels are consistently lower compared to other nutrients, rarely exceeding 40. This suggests potential deficiencies requiring intervention. Temperature shows a consistent pattern, maintaining a steady fluctuation without extreme spikes, reflecting a relatively stable environment conducive to nutrient dynamics. Key insights from the graph reveal that Nitrogen requires close monitoring due to its high variability, and potassium may need supplemental fertilization due to its consistently low levels.

Figure 3: Graph Indicating Levels of Nitrogen (N), Phosphorus (P), Potassium (K), Temperature



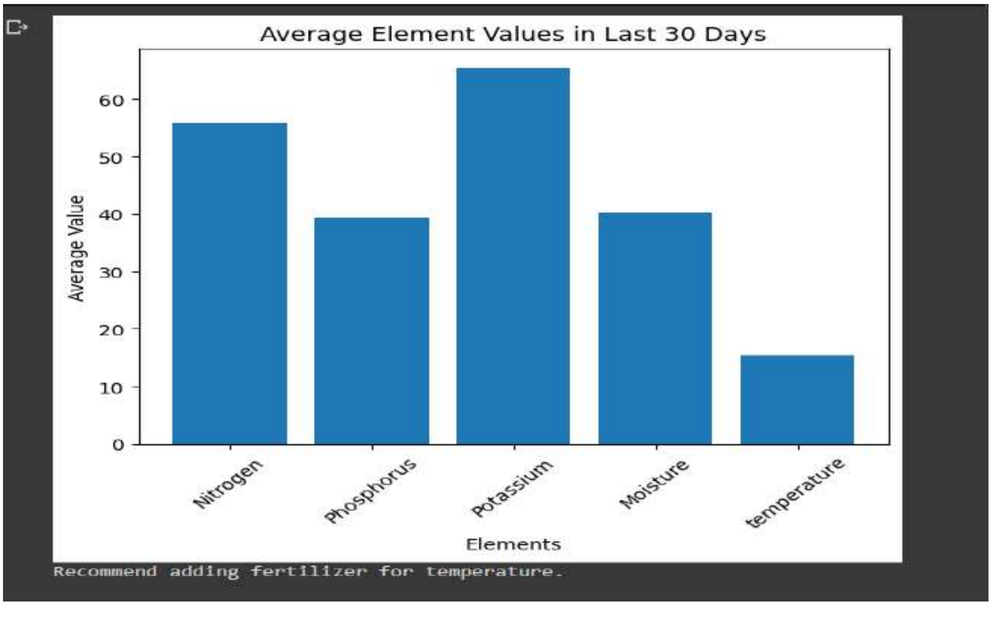
The Figure 4 displays time-series graphs for monitoring soil parameters like moisture, nitrogen, potassium, and temperature.

Figure 4: Individual Graphs for Each Element with their Recommended Range of Values



The moisture level frequently dips close to the lower limit, indicating periods of dryness. It occasionally approaches the upper limit but rarely exceeds it. Low moisture levels may require recommendations for irrigation, especially if they persist below the threshold. Nitrogen levels oscillate around the expected value. Potassium levels vary widely, with frequent crossings below the lower limit. The temperature shows high variability but generally remains within the defined upper and lower limits. Our system notifies users when parameters deviate from optimal ranges for immediate corrective actions. Temperature is crucial for crop growth, and recommendations include adjusting watering schedules or soil coverings to manage extreme conditions and required fertilizers to properly monitor and grow crops. Analyzing trends could identify long-term issues in soil health, such as nutrient depletion or suboptimal irrigation practices.

Figure 5: Bar Graph to Lowest NPK Element Values during the Previous 30 Days



As shown in Figure 5, illustrates the average values of various soil and environmental parameters over the past 30 days. The graph shows a high average value of Nitrogen (N), indicating sufficient levels of nitrogen, essential for plant growth and green leafy development. The Phosphorus (P) levels are moderate, indicating important for root and flower development. The potassium (K) levels are the highest among all elements. Potassium supports the plant's immune system and overall growth, so this seems to be in good shape. Moisture level is moderate, adequate for crop depending on type of plant, can be adjusted. Temperature is noticeably low.

The recommendation “adding fertilizer for temperature” implies using organic mulches or soil amendments to help regulate soil temperature indirectly, rather than directly influencing the element itself. Nitrogen and Potassium are sufficient, so no immediate supplementation is necessary. Phosphorus could be boosted slightly to ensure optimal root and flowering conditions. Moisture management might be needed to prevent drought stress.

6.0 Conclusion

Farmers may find a smart agriculture system to be beneficial. The most helpful invention for Indians is using sensor networks in agriculture monitoring. In smart farming, output must be increased while expenses, labor hours, and human effort are reduced. By establishing this system in place, users like farmers would be able to track and raise agricultural yield. By using this proposed system farmers can efficiently monitor their fields & soil. Farmers can get detailed information regarding soil with convenience. Rather than visiting the lab for Soil Testing. The system will suggest the Fertilizer by analyzing the parameters like Moisture value, Temperature value, of the soil. Maintaining the record of Land fertility will become very easy & access it from anywhere using your mobile device. In the Future by using Smart Agriculture, Automatic Fertilizer spraying may also be possible.

7.0 Future Work

By using next-generation soil sensors could provide more detailed analyses by detecting a broader spectrum of nutrients and soil chemicals, including secondary and micronutrients like calcium, magnesium, and sulfur, as well as soil pH and salinity. These sensors could integrate real-time monitoring with AI-driven platforms to offer predictive insights, enabling more precise and dynamic nutrient management tailored to varying crop needs and environmental conditions.

Additionally, advanced chemical sensors for pesticides and herbicides could monitor residue levels in the soil, ensuring safer and more sustainable use of agricultural chemicals. These sensors could also help farmers identify contamination hotspots or detect the degradation rate of chemicals, promoting better soil health and compliance with environmental standards. Combined with weather and soil data, these tools could predict pest and disease outbreaks, guiding the timely and targeted application of pesticides.

8.0 Acknowledgments

The authors are thankful to the colleagues at Marathwada Mitra Mandal’s College of Engineering and MIT Academy of Engineering, Pune for their valuable insights provided.

References

Abdullah Ahmad, E. K. (2016). An IOT based system for remote monitoring of soil characteristics. *Proc. InCITe*. https://doi.org/10.1109/INCITE.2016.7857638

Agraj Aher, J. K. (2018). Smart agriculture using clustering and IOT. *International Research Journal of Engineering and Technology, 5*(3), 2395-0056.

Agyabeng-Mensah, Y., Afum, E., Baah, C., & Essel, D. (2022). Exploring the role of external pressure, environmental sustainability commitment, engagement, alliance and circular supply chain capability in circular economy performance. *International Journal of Physical Distribution & Logistics Management, Vol. 52 No. 5/6*, 431-455. https://doi.org/https://doi.org/10.1108/IJPDLM-12-2021-0514

Anjaneyulu, P. D. (2024). Sustainable practices for plant nutrient control. *international Conference on Futuristic Trends in Engineering, Science & Technology (ICFTEST-2024).*

Anjaneyulu, P. D. (2024). Sustainable practices for plant nutrient control using IoT. *International Conference on Futuristic Trends in Engineering, Science & Technology (ICFTEST-2024).* E3S Web of Conferences 507, 01039 (2024).

Arkadiusz Kawa, B. P. (2021). Green logistics in e-commerce. *LogForum, 17*(2), 183-192. https://doi.org/Publisher: Wyższa Szkoła Logistyki

Arumugam, Y. S. (2022). Towards applicability of machine learning techniques in agriculture and energy sector. *Materials Today: Proceedings*, 2260-2263.

Asongu, N. M. (2023). Microfinance institutions and female entrepreneurship in Sub-Saharan Africa: avoidable female unemployment thresholds. *8*(7), 18-19. https://doi.org/https://hdl.handle.net/10520/ejc-defa_v8_n7_a6

Blesslin Sheeba, L. D. (2022). Machine learning algorithm for soil analysis and classification of micronutrients in IoT-enabled automated farms. *Journal of Nanomaterials*.

Dhanaraju, M. C. (2022). Smart farming: Internet of Things (IoT)-based sustainable agriculture. *Agriculture, 12*(10), 1745.

Du, C., Zhang, L., Ma, X., Lou, X., Shan, Y., Li, H., & Zhou, R. (2021). A cotton high efficiency water-fertilizer control system using wireless sensor network for precision agriculture. *Processes 2021, 9(10), 1693.*

Gupta, B., *et al.* (2022). Challenges faced by women entrepreneurs in leveraging Karnataka State Financial Corporation (KSFC) services: A comprehensive analysis. *Journal of Women in Business*.

Hassan, M. M. (2021). A systematic review on monitoring and advanced control strategies in smart agriculture. *IEEE Access*, 32517-32548.

Jayaprahas, S. S. (2014). Real time embedded based soil analyzer. *International Journal of Advanced Research in Computer and Communication Engineering, 3*(3), 5700-5703.

Kegne, A. A. (2023). The role of microfinance institutions on women's entrepreneurship development. *Journal of Innovation and Entrepreneurship, 12*, 17. https://doi.org/https://doi.org/10.1186/s13731-023-00285-0

Khan, S. A., Zhang, Y., & Nathaniel, S. (2020). Green supply chain performance and environmental sustainability: A panel study. *LogForum*, *16(1)*, 141-159.

Lavanaya, M. &. (2018). Soil nutrients monitoring for greenhouse yield enhancement using Ph value with IOT and wireless sensor network. *ICGCIoT*.

Mohanraj I, G. V. (2017). Intelligent drip irrigation and fertigation using wireless sensor networks. *IEEE*, 978-1-5090-4437-5/17.

Navleenkaur Harpreetsingh Bindra, K. (2023). A survey for soil testing and scheduling in IOT enabled farms using ML algorithm. *IJCRT*.

Padole, P. S. (2017). Design and implementation soil analyser using IO. *IEEE(ICIIECS)*, 978-1-5090-3294-5/17.

Patel, A. (2020). Deep-learning-based approach for estimation of fractional abundance of nitrogen in soil from Hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations*.

Puranik, S. A. (2019). Automation in agriculture and IoT. *4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), Ghaziabad, India*, 1-6.

Raut, S. A. (2020). Soil monitoring and testing using IoT for fertility level and crop prediction. *Proceedings of the 3rd International Conference on Advances in Science & Technology (ICAST).*

Singh, A., & Mishra, S. (2023). The role of Karnataka State Financial Corporation (KSFC) in fostering women-led entrepreneurial endeavors: An empirical investigation. *Journal of Entrepreneurship and Economic Development*.

Souza, E. D., Kerber, J., & M. Bouzon, C. R. (2023). Analysis of environmental consciousness towards sustainable consumption: An investigation on the smartphone case. *Journal of Cleaner Production, 384*. https://doi.org/10.1016/j.jclepro.2022.135543

Sukumar, D. T. (2018). Real time soil fertility using IOT. *Proc. NC's e-TIMES*.

CHAPTER 2

STYLE-SYNC: An AI Based Outfit Recommendation System

Prajwal Gadhave, Shreyash Junje**,*
*Dujal Sawarkar***, Vivek Thorat**** and Ketaki Bhoyar******

ABSTRACT

Artificial Intelligence has entered the fast phase of evolution, presenting new opportunities in the fashion industry by helping the consumer to have a personalized fashion experience with personalized recommendations. This is why STYLESYNC becomes necessary, which is an intelligent wardrobe management and styling system that will use AI to change personal fashion. The main aim is to create a platform where users will access an absolute fashion experience through wardrobe organization and special occasion outfit suggestions. Users will upload their clothes present in their wardrobe into the system, which recommends outfits probably from wearing tones for lunch-dates to the most formal events. Using an advanced algorithm, the platform evaluates individual user preferences and trends in clothing and event-type context, ensuring its recommendations comply with personal tastes and current fashion norms. Therefore, it integrates additional elements of sustainability, such as making the most of these outfits, urging people to buy only what they truly need rather than what they can and generally reducing pointless expenditures.

Keywords: Artificial Intelligence; Personalized recommendations.

1.0 Introduction

Today's fast-moving world makes it difficult for an individual to choose what to wear every day, so it creates tension. This innocent choice could leave a lot of confusion when weighing the ceremony, style, trend, or, on the contrary, based on subjective feelings such as temperature or even mood.

**Corresponding author; UG Student, Department of Computer Engineering, Dr. D. Y. Patil Institute of Engineering, Management and Research, Pune, Maharashtra, India (E-mail: prajwalgadhave356@gmail.com)*

***UG Student, Department of Computer Engineering, Dr. D. Y. Patil Institute of Engineering, Management and Research, Pune, Maharashtra, India (E-mail: shreyashjunje2003@gmail.com)*

****UG Student, Department of Computer Engineering, Dr. D. Y. Patil Institute of Engineering, Management and Research, Pune, Maharashtra, India (E-mail: dujalsawarkar5@gmail.com)*

*****UG Student, Department of Computer Engineering, Dr. D. Y. Patil Institute of Engineering, Management and Research, Pune, Maharashtra, India (E-mail: vivekthorat588@gmail.com)*

******Faculty, Department of Computer Engineering, Dr. D. Y. Patil Institute of Engineering, Management and Research, Pune, Maharashtra, India (E-mail: ketaki.bhoyar@dypiemr.ac.in)*

Wardrobe organization has evolved to a point where it relies on the relatively primitive means of organizing things manually and trying to distinguish without the aid of technology. The necessity of technology has emerged due to these reasons, and the fashion industry is now undoubtedly tending toward evolving technologies for better user experience and assisted decision-making. It is Artificial Intelligence (AI) that can prove to be a revolutionizing tool, personalized to everybody's needs and welcomes for the changes it can induce in the lives of people.

Moreover 'AI integration in fashion' hence leads toward building such intelligent systems that would help manage personal wardrobes, can suggest 'what to wear,' and could even predict the trends in fashion as per consumer behavior.

StyleSync is about to be changed. An AI-based wardrobe system that aims to engage people differently in their clothing collections provides the most comprehensive solutions including organizing wardrobes for users and suggesting personalized outfits for special events and preferences in their lives. The technology employs cutting-edge AI algorithms such as Machine Learning (ML) and Natural Language Processing (NLP).

Hello everyone, and welcome to the introduction. In today's fast-paced world, every day involves a choice for an individual-the seemingly simple choice of what to wear, which can become quite a source of tension under the barest consideration of ceremony, current fashion trends, individual style preferences, and even the weather. Traditional wardrobe management relying on organizing things manually and trying to distinguish-just does not work for these complexities. Hence, this necessity for technology has emerged and the fashion world is now undoubtedly in the trajectory toward evolving technologies as far as user experience and assisted decision-making is concerned. It is Artificial Intelligence (AI) which could, in all ways possible, have been capable of proving to be a revolutionizing tool, personalized for everybody's needs, and welcoming all the changes it has been capable of inducing in the lives of people.

AI integration in fashion has thus paved the way toward having intelligent systems that could help manage the private wardrobe, suggest what to wear today, and could even predict fashion trends based on consumer behavior. StyleSync is going to be revolutionized. An AI-based wardrobe system is aimed to change the way individuals relate with their collections of clothing. Using the latest AI technologies such as Machine Learning and Natural Language Processing, StyleSync provides the most complete solutions for organizing users' wardrobes and personalized outfit suggestions for events and preferences. work, and Section VI concludes the paper with an overview of the limitations and potential for further development.

2.0 Literature Review

The fashion sector is becoming more and more endowed with AI to strengthen the innovative task of fulfilling the direct personalization of shopping and stock management and the engagement of customers. Research like that of Smith *et al.* (2019) has alluded to the dynamic transformative role that AI has played in facilitating hyper-personalized recommendations through user preferences and historical data. Likewise, Jones & Taylor (2021) describe how AI improves consumer satisfaction by using machine learning for accurate size and fit prediction.

Wardrobe management systems facilitate the organization, compartmentalization, and usage of personal articles of clothing efficiently. The research of Lee & Kim (2020) illustrates how significantly this type of system reduces decision fatigue experienced by its users by giving suggestions for complete outfits. Old-fashioned wardrobe applications would not have had this linkage with e-commerce platforms, thereby limiting the application. StyleSync now fills that gap by linking a personal wardrobe with actual products available on sites like Amazon and Flipkart.

Indeed, AI would bring change to a vast number of domains including entertainment, e-commerce, and fashion. According to Chen *et al.* (2020), the AI models employing collaborative filtering, content-based filtering, or hybrid methods can be regarded highly accurate for user preference prediction. Moreover, the application of Natural Language Processing (NLP) and computer vision increases the effectiveness of this kind of system regarding the analysis of user input and visual data, as demonstrated by Zhao *et al.* (2019).

Sustainable fashion may be a concern for a growing number of people in industry terms. As Green *et al.* (2021) remark, AI could play an important part in sustainability by encouraging the reuse of old wardrobe items or by letting consumers know the environmental impact of their planned purchases. Such objectives fit perfectly with the mission of StyleSync-the efficacious utilization of current mbere wardrobe collections for environment-friendly living.

Integrating AI into e-commerce platforms has truly changed the way we shop by offering personalized recommendations for products and encouraging discussions with customers. Patel & Gupta (2020) focused on the role of real-time data integration in offering users a seamless experience. This approach has been further extended by StyleSync using the APIs of popular e-commerce companies, thus bringing the aspect of missing wardrobes into the ability of consumers to identify as well as purchase them.

A big problem for these AI-driven systems is that they would also create data privacy issues, algorithmic bias, and scale issues, thus driving future research to propose ways to address such concerns and ensure that applications are fair and safe. As such, StyleSync means taking all the new data encryption and ethical AI practices to be free from such worries.

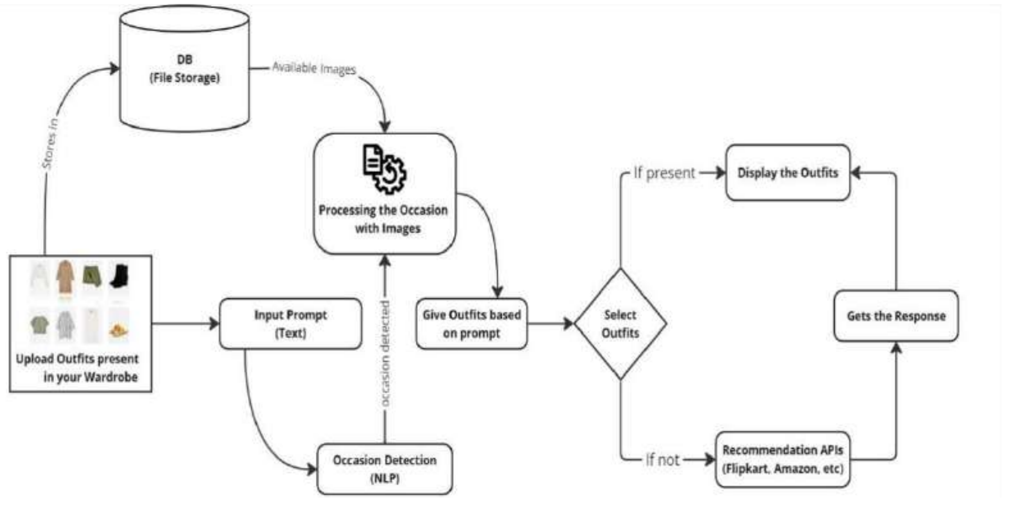
Emerging from the literature, StyleSync may tactically augment digitized wardrobe management powered by AI and e-commerce. Filling the specific existing gaps while innovating on proven strategies will move StyleSync to create an exponentially more sophisticated enhancement solution to user experience in sustainable fashion choices.

3.0 Methodology

3.1 Overview of the system design

As its name suggests, the STYLESYNC system adopts a modular-as well as a layered architecture. This design is fundamentally required to make the components of STYLESYNC scalable, efficient, and integrated. A design philosophy followed in the architecture is user-oriented and designed with most available state-of-the-art technologies using AI/ML, Cloud Services, and Responsive Front-end frameworks towards creating a dynamic and feature-rich experience. The Figure 1 shows the architecture which is structured as follows:

Figure 1: System Architecture



3.1.1 Frontend layer

The frontend builds a conversion-interactive user interface for seamlessly transitioning between operations with a very high degree of responsiveness, utilizing ReactJS and NextJS. The design is made consistent with TailwindCSS and faster prototyping which can develop a modern and good-looking interface customized from user preferences.

3.1.2 Backend layer

The backend is composed of Node.js and Express.js for all strong server-side operations and is thus functioning well for the API and acts as the middle between the database and front-end AI/ML models.

3.1.3 Storage and database layers

Amazon S3 is used for storing user-uploaded images safely. Highly available, high throughput, and durable storage should be quick and reliable in retrieval.

MongoDB, a NoSQL database, serves to store both structured and unstructured data concerning user wardrobes, outfit suggestions, and contextual metadata. Thus, by separating these layers, the system achieves modularity and scalability.

3.2 Data collection and preprocessing

Ideal recommendation systems are formed through a collection of a diverse and rich variety of data. STYLESYNC collects and processes data from:

User-Provided Data: The users upload images of their garment items which get stored on S3 of amazon. Images undergo preprocessing by Replicate AIs like background removal, Object Detection, and Feature Extraction. Moreover, additional information such as garment type, brand, and user preferences like color, style preferences, etc. have also been collected and stored in MongoDB.

Third-Party APIs: It will be an API of Amazon or Flipkart, in future, through which real-time data could be fetched for certain items of clothing and accessories or any other related items of the relevant kind to be recommended by the system.

Surroundings Data: Weather conditions, time of day, and event types (casual, formal, outdoors, etc.) will be used along with context inputs to personalize recommendations.

3.3 AI/ML model implementation

Modern AI technology is also accompanied by ML models which have translated the data-generated recommendations to improve user experiences by the STYLESYNC system into various applications.

Image Processing: Replicate AI is used to analyze images uploaded by users to their wardrobe. Aspects such as the type of garment, color, and style have been targeted and would serve as the basis for an accurate outfit suggestion.

Natural Language Processing (NLP): Such capability is provided by ChatGPT 4.0, which would express all user queries and preferences. Whether they ask for something like "Suggest a casual outfit for a beach outing" or leave feedback, it would through represent in accurate, meaningful NLP-based mode.

Recommendation System: The recommendation engine combines the two popular techniques:

- Content-Based Filtering: recommend outfits based on the matched items in a user's wardrobe with an extracted feature profile of color, fabric, and style.
- Collaborative Filtering: Using the preferences and selections made by similar users to suggest new outfits or accessories. Related to this, Gemini Fast API offers rapid recommendation updates, thus ensuring that any new items added remain relevant and seamless to the system.

3.4 Workflow and functionality

The end-to-end functionality of the STYLESYNC system is structured into the following stages:

Wardrobe Digitization: The users upload their clothes images through the front end. The images are secured and stored in Amazon S3 while being run through Replicate AI to extract relevant features. Metadata (e.g., garment type, occasion suitability) has been cataloged under MongoDB.

User Input and Context: The users give context toward the requirement for the outfit, such as the type of event, weather, or any personal preferences. For example, one user may enter: "Formal outfit for a winter evening." The input is processed using ChatGPT 4.0, which now interprets the context and preferences to formulate much more specific recommendations.

Recommendation Generation: Based on user inputs and wardrobe data, the AI/ML models generate personalized outfit recommendations. For example, if the user selects a formal occasion, the system may recommend a blazer with complementary trousers and a shirt. The system highlights gaps in the user's wardrobe and, in the future, can recommend Amazon or Flipkart's items to complete the outfit.

Interactive Feedback: Loop Users can give feedback on the recommendations, which allow the system to improve its suggestions over time.

Trend Adaptation: Keeping track of real-time fashions trends adds to the recommendation engine such that the recommendations will always be in tune with the trends.

3.5 Deployment and Scaling

The STYLESYNC system deploys cloud services and modern DevOps:

Cloud Infrastructure: The backend is based on AWS for database storage systems and to give reliability and scalability. Because Amazon S3 is responsible for the image storage, the backend APIs are deployed in an environment capable of scaling to handle high traffic.

Automated CI/CD Pipelines: Build and set up automated CI/CD pipelines to minimize downtime during the deployment of production changes. Allows seamless testing of code within environments by code deployment.

API Architecture: The backend APIs of the system are developed on the Express.js framework, complying with RESTful principles to clarify the purpose of services in the data exchange process. The APIs should be designed to support highly concurrent requests.

Front-end Scalability: It's optimized with server-side interpretation for performance accruals as the application grows larger and begins to accommodate more users. The front end is based on ReactJS in conjunction with NextJS.

3.6 Future enhancements

Integrations with third parties: The future scope includes integrated LLM APIs to increase contextual knowledge and include Amazon and Flipkart APIs for e-commerce recommendations.

Augmented Reality (AR): This will bring AR technology to try on outfits virtually and see how it looks before trying up.

Social Trend Analysis: AI models would cover social media and fashion blogs that testify emerging trends and feed into the recommendation system.

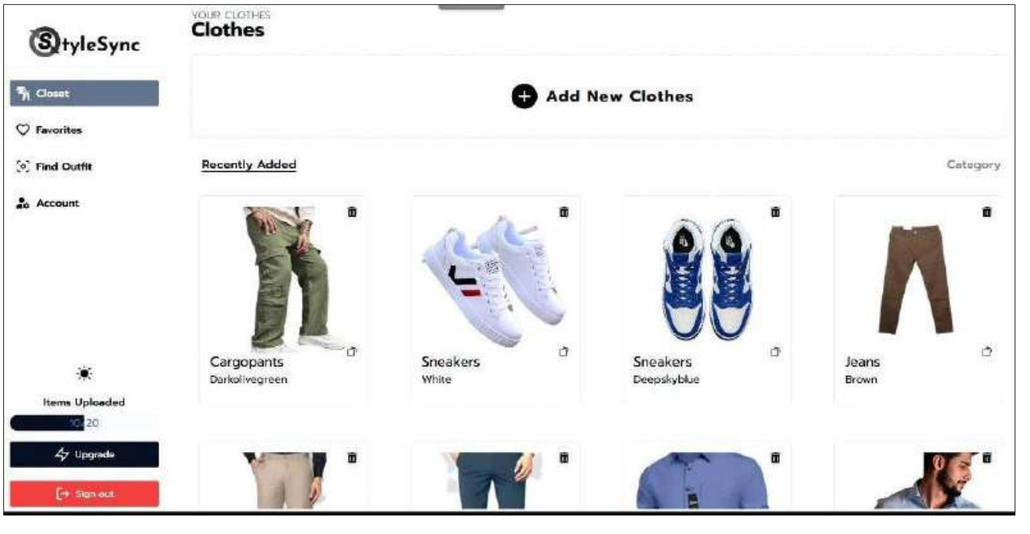
Intensive User Profiling: Improvement in the profiling of users is going to sharpen the concepts of personalization-of recommendations based on individual body types or some other specific fashions.

Sustainable Recommendations: Features that influence the green cause will be taken by the system to offer sustainable fashion alternatives keeping in mind the environmentally conscious user.

4.0 Results

To start, users click on the '+ Add New Clothes' button to upload items and give their name, color, or category and an image placeholder is shown (e.g., "Collared Shirt, Cargo Pants"). After that, the uploaded clothes are shown in the 'Added Recently' section (as shown in Figure 2 and Figure 3). Every clothing piece looks like a card with the opportunity to remove or make a copy of it to help in clothing closet organization.

Figure 2: Uploading Process



As Shown in Figure 4, the "Find Outfit" of the Style-Sync, which enables the users to select outfits depending on the garments in their wardrobe and the climate. Like the weather in Pimpri, Maharashtra, now with information on wind speed 13.3 kph, humidity 41%, and temperature 26.4°C (79.6°F). This assists users in decision-making by choosing clothes appropriate for the existing environment.

Figure 3: Implementation of Replicate AI

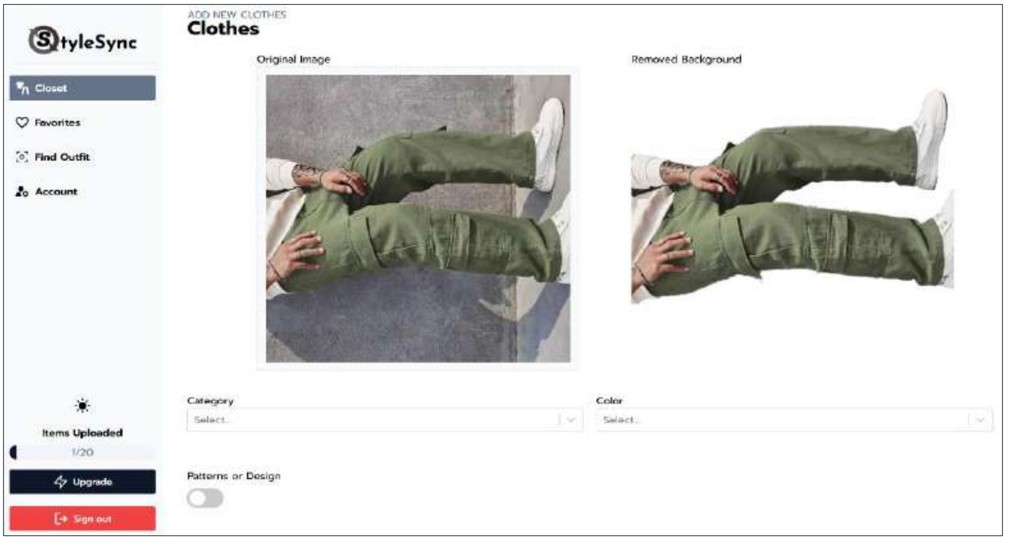
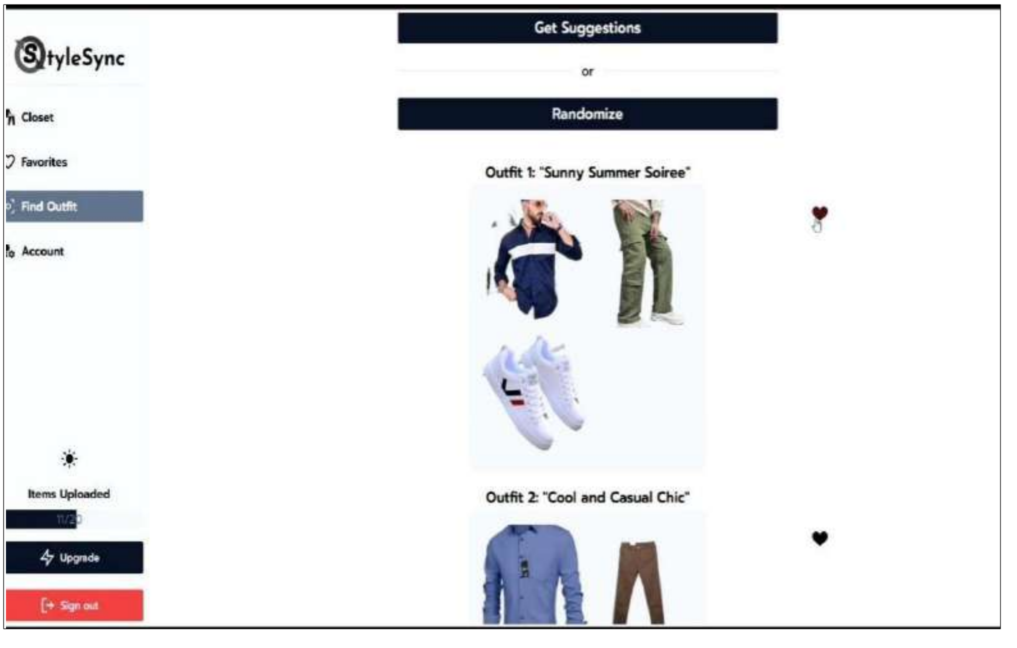


Figure 4: Recommendation Process



5.0 Conclusion

Digital applications of Artificial Intelligence change the way people see their lives with fashion. By integrating AI-powered outfit suggestions, wardrobe digitization, and sustainable practices, the STYLE-SYNC program addresses real issues like making personal styling, consumer engagement, and environmental waste management realistic with the promise of meeting even the higher future trends in fashion. This way, the program becomes more than a tool; it is made personal-thoughtfully innovating users into fashion experience, with a future trend ahead in mind. One of the major reasons why STYLE-SYNC is appreciated is flexibility, which is possible since its modular design allows it to grow and evolve as technology grows and consumer behavior changes. For this reason, this highly flexible fashion solution is both current and future ready.

6.0 Acknowledgement

We would like to express our sincere gratitude to Mrs. Ketaki Bhoyer for her invaluable guidance, direction and encouragement that has been the cornerstone of this research effort. His depth of knowledge, insightful suggestions, and unwavering patience greatly determined the direction and depth of this study. His commitment to advancing academic excellence inspired us to strive for the highest standards in our field. We are Dr. D.Y. Patil Institute of Engineering, Management, and Research for providing the services, facilities, and learning environment that contributed to the success of this study. The institute's focus on innovation and academic rigor has given us the space to explore and expand the boundaries of our knowledge.

We also extend our heartfelt thanks to our families and friends for their unwavering support, patience and encouragement throughout this journey. They stood by us through the difficult part *of our research*, providing moral support and understanding that kept us grounded and spirited. Their confidence in our strength was a constant source of strength.

We would like to thank the many researchers, scholars and industry professionals whose pioneering work has served as the foundation and inspiration for the tremendous contribution of this research. Their commitment to advancing knowledge in the field provided valuable insights and perspectives, enhancing the depth and depth of our study.

Finally, we would like to express our appreciation to students and online resources that have facilitated access to important information and research resources. Sharing knowledge in academia and the entrepreneurial spirit played an important role in Shapine

References

M. O. Ayemowa, R. Ibrahim, and M. M. Khan, "Analysis of Recommender System Using Generative Artificial Intelligence: A Systematic Literature Review," IEEE Access, vol. 12, pp. 1234–1250, June 2024, doi: 10.1109/ACCESS.2024.3416962.

C. Limaksornkul, D. Na Nakorn, O. Rakmanee, and W. Viriyasitavat, "Smart Closet: Statistical-based Apparel Recommendation System," in 2014 Third ICT International Student Project Conference (ICT-ISPC2014), Mahidol University, Nakhonpathom, Thailand, pp. 56–62, 2014.

S. Prabhu, P. Ashok, G. Hallur, and R. Nandanwar, "Stitching Data Threads: Impact of Artificial Intelligence on Fashion Evolution," in 4th International Conference on Innovative Practices in Technology and Management (ICIPTM 2024), Lavale, Pune, Maharashtra, India, pp. 78–84, 2024, doi: 10.1109/ICIPTM.2024.3415789.

C. Stan and I. Mocanu, "An Intelligent Personalized Fashion Recommendation System," in 2019 22nd International Conference on Control Systems and Computer Science (CSCS), University Politehnica of Bucharest, Romania, pp. 92–97, 2019, doi: 10.1109/CSCS.2019.123456.

Z. Yang, Z. Su, Y. Yang, and G. Lin, "From Recommendation to Generation: A Novel Fashion Clothing Advising Framework," in 2018 7th International Conference on Digital Home (ICDH), Sun Yat-sen University, China, pp. 45–52, 2018, doi: 10.1109/ICDH.2018.234567.

C. Stan and I. Mocanu, "An Intelligent Personalized Fashion Recommendation System," in 2019 22nd International Conference on Control Systems and Computer Science (CSCS), University Politehnica of Bucharest, Romania, pp. 92–97, 2019, doi: 10.1109/CSCS.2019.123456.

Ayemowa, M. O., Ibrahim, R., & Khan, M. M. (2024). "Analysis of Recommender System Using Generative Artificial Intelligence: A Systematic Literature Review." IEEE Access. DOI: 10.1109/ACCESS.2024.3416962.

Limaksornkul, C., Na Nakorn, D., Rakmanee, O., & Viriyasitavat, W. (2014). "Smart Closet: Statistical-based Apparel Recommendation System." Third ICT International Student Project Conference (ICT-ISPC2014).

Prabhu, S., Ashok, P., Hallur, G., & Nandanwar, R. (2024). "Stitching Data Threads: Impact of Artificial Intelligence on Fashion Evolution." 4th International Conference on Innovative Practices in Technology and Management (ICIPTM 2024).

Stan, C., & Mocanu, I. (2019). "An Intelligent Personalized Fashion Recommendation System." 22nd International Conference on Control Systems and Computer Science (CSCS).

Yang, Z., Su, Z., Yang, Y., & Lin, G. (2018). “From Recommendation to Generation: A Novel Fashion Clothing Advising Framework.” 7th International Conference on Digital Home (ICDH).

Moosaei, M., Lin, Y., & Yang, H. (2020). “Fashion Recommendation and Compatibility Prediction Using Relational Network.” arXiv preprint arXiv:2005.06584.

Chen, W., Huang, P., Xu, J., Guo, X., Guo, C., Sun, F., Li, C., Pfadler, A., Zhao, H., & Zhao, B. (2019). “POG: Personalized Outfit Generation for Fashion Recommendation at Alibaba iFashion.” arXiv preprint arXiv:1905.01866.

Celikik, M., Kirmse, M., Denk, T., Gagliardi, P., Mbarek, S., Pham, D., & Peleteiro Ramallo, A. (2022). “Outfit Generation and Recommendation -- An Experimental Study.” arXiv preprint arXiv:2211.16353.

Becattini, F., Chen, X., Puccia, A., Wen, H., Song, X., Nie, L., & Del Bimbo, A. (2024). “Interactive Garment Recommendation with User in the Loop.” arXiv preprint arXiv:2402.11627.

CHAPTER 3

Blockchain based Loan Payment Control and Tracing

Anushka Nirmalkumar Gosavi, Aditya Bhadakwade**,*
*Prajyot Bhamare***, Vrushali Pawar**** and Sharmishta Desai******

ABSTRACT

The loan management systems in practice today are rudimentary, opaque, and ineffective. The transactions are not easily traceable, and they lack transparency. Apart from this, tracing where the borrowed money is utilized is a major issue. One solution to this issue is blockchain technology. It is commonly observed that the information settlement mechanism, such as SWIFT, is isolated from the payment settlement mechanism on a ledger. However, if banks use a ledger that stores information settlement distributed across all participants, the fraudulent activity could be reflected upon by all available participants, auditors, and regulators. This would potentially increase transparency and accountability in financial transactions and reduce the risk of fraud and other corrupt practices. To overcome these flaws, we introduce a blockchain-based loan disbursal and tracing system. For this task we propose using CBDC's (Central Bank Digital Currencies) such as the Digital Rupee introduced by the Government of India. We designed a Three Tier System that uses smart contracts for payee and recipient verification. We also propose a method for connecting centralized ledgers (banks) and decentralized ledgers (blockchain). The blockchain based system provides an added benefit of cross border transactions.

Keywords: Blockchain; Loan management system; Smart contracts; Distributed ledger.

1.0 Introduction

The advent of blockchain technology has opened a new arena for research in the financial field. Fintech companies have completely changed the picture for banking and loan management.

**Corresponding author; Associate Product Developer, Helix-ITSM, BMC Software, Pune, Maharashtra, India (E-mail: anushkagosavi172@gmail.com)*
***Business Technology Solutions Associate, R&D, ZS Associates, Pune, Maharashtra, India (E-mail: bhadakwadeaditya@gmail.com)*
****Associate Product Developer, Helix-ITOM, BMC Software, Pune, Maharashtra, India (E-mail: pcbhamare.41@gmail.com)*
*****Associate Full-stack Developer, R&D, Setoo Solutions, Pune, Maharashtra, India (E-mail: vrushalipawar1802@gmail.com)*
******Associate Professor, School of Computer Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India (E-mail: sharmishta.desai@mitwpu.edu.in)*

Right from loan applications and sanctions to tracking and storage, every step is digitized. Despite a variety of security algorithms and advanced verification techniques, the current system is prone to several issues. The main problem with the current loan management systems is that they lack transparency and are susceptible to fraudulent activities. The transactions are stored in a private database and the data can be easily tampered. Also, tracking the loan amount usage is manually done in conventional systems.

Many times, people tend to apply for loans with low interest rates and use them for unintended tasks. For example, agriculture loans have low interest rates, and a car loan might have a high interest rate. A person might use this agriculture loan to buy a car. Thus, tracking the loan, right from sanction to final spending is a need of the hour. We have proposed a methodology wherein we exploit the distributed ledger technology. The loan amount will be converted to tokens like the digital rupee and will be transferred to the digital account of the borrower. A public blockchain is used to store all the transactions with role-based access privileges. This will foster transparency and immutability.

We need a way of having immutable scripts working in our system. Also, we need to ensure that all the data generated on the execution of these scripts needs to be transparent, immutable, and secure. To achieve this, we make the use of smart contracts to have self-executing programs and store all the data they generate on the blockchain forever. We will be using smart contracts to verify if the receiver is a registered legitimate entity. This will avoid the misuse of loan money and ensure that the loan money is utilized correctly and transparently. We also propose a secure and trustworthy algorithm for linking a person's bank account to a digital account using digital signatures.

We have thus proposed a Three Tier System with the base level being the blockchain contracts usable by any entity. The second tier builds upon the first tier providing a rest API as a service to mobile and web platforms. The third tier provides a complete end-to-end solution directly usable by the customers.

2.0 Literature Review

Here we discuss the various systems in existence and current advances, research in the domain of our proposed system. Digital loan management is already in prevalence all throughout the world. Blockchain, although new, is already incorporated by several banks and organizations in fintech.

The adoption of blockchain and distributed ledger technology (DLT) has revolutionized payments, loan disbursement, and settlements, particularly in enhancing cross-border transactions. International projects like the Bank of Canada's *Project Jasper* and Singapore's *Project Ubin* have demonstrated the efficacy of DLT in creating faster, more secure, and cost-efficient systems for international payments (Garratt *et al.*, 2019). Similarly, the Monetary Authority of Singapore proposed blockchain-based systems for domestic and cross-border interbank payments, further establishing DLT's transformative potential (Monetary Authority of Singapore [MAS], 2017). In a joint venture, the European Central Bank and the Bank of Japan launched *Project Stella*, a study investigating the application of DLT in financial market infrastructure (Bank of Japan, 2017). Additionally, the People's Bank of China has established the Digital Currency Institute to develop digital fiat currency, showcasing the growing importance of blockchain in central banking (Amamiya, 2019). The World Bank has also explored disbursement traceability via blockchain to reduce fraud and money laundering, reflecting the global interest in blockchain's capacity to foster transparency and trust in financial systems (World Bank, 2020).

In his seminal work, Peters *et al.*, (2015) discussed the integration of smart contracts and DLT in fintech, focusing on enhancing transaction efficiency and security. Khara, Kumar, and Sharma *et al.*, (2020) proposed a blockchain framework for loan management, wherein funds are tokenized into cryptocurrency and managed via public blockchain to ensure transparency and immutability. K.S. Arikumar *et al.*, (2021) developed a decentralized application (DApp) to mitigate fraudulent loan sanctions and ensure transaction security using smart contracts. Addressing payment freezing, Wu, Lu, and Xu *et al.*, (2022) studied strategies for freezing funds during the disbursement cycle, leveraging blockchain's immutable nature.

Wang, Guo, and Cheng *et al.*, (2019) introduced an automated financial loan management system using smart contracts, incorporating stack-based locking algorithms, events, oracles, and digital signatures to ensure privacy and secure authentication. Although Wang *et al.*, (2019) also explored linking bank accounts to digital wallets, they did not propose specific algorithms for this integration. Yang, Zhang, and Li *et al.*, (2018) designed a peer-to-peer payment system based on smart contracts, highlighting the scalability and flexibility of blockchain for decentralized financial transactions. These contributions emphasize blockchain's capabilities to enhance financial systems' transparency, security, and operational efficiency. However, significant challenges remain, such as integrating centralized and decentralized ledgers, addressing high latency during currency conversion, and ensuring scalability for high transaction volumes. These areas continue to be the focus of active research and development in the field.

Blockchain technology has emerged as a transformative tool for governmental supervision in construction, streamlining compliance and enhancing transparency. Lu, Wu, Zhao, Li, and Xue *et al.*, (2021) investigated its utility in monitoring construction activities by drawing parallels with digital currency electronic payment systems. Their research revealed how blockchain's immutable and decentralized ledger could automate compliance checks, reduce opportunities for corruption, and ensure financial monitoring of public projects. By embedding transactions into an auditable blockchain framework, stakeholders can achieve greater trust and accountability across the construction lifecycle.

Incorporating blockchain into sustainable practices, Sadeghi, Mahmoudi, and Deng *et al.*, (2021) explored its role in promoting resource efficiency and addressing environmental challenges in the construction industry. They identified that blockchain technology could be leveraged to reduce material waste, improve energy consumption tracking, and ensure ethical sourcing of materials. However, the study also pointed out that resistance from stakeholders due to perceived complexity and a lack of familiarity remains a significant barrier to adoption. This highlights the need for awareness programs and simplified interfaces to integrate blockchain into existing workflows. Smart contracts have emerged as a cornerstone of blockchain applications in construction, especially for secure and automated payments.

Ahmadisheykhsarmast and Sonmez *et al.*, (2020) developed a smart contract system designed to address common challenges in construction payments, such as delayed disbursements and disputes. Their approach tied payment conditions to project milestones, ensuring that funds were only released once predefined criteria were met. This automation not only reduces administrative overhead but also builds trust among stakeholders by minimizing human error and ensuring objective enforcement of contract terms.

Private blockchain networks, such as Hyperledger Fabric, are particularly well-suited for construction quality management. Zhong, Wu, Ding, Luo, Luo, and Pan *et al.*, (2020) implemented a consortium blockchain model to manage and share quality-related information securely among authorized stakeholders. Their system used the capabilities of Hyperledger Fabric to ensure data confidentiality while maintaining the transparency and traceability of transactions. This approach addressed key challenges in collaborative construction environments, such as the need for tamper-proof documentation and real-time access to shared data, thereby improving overall project efficiency and trust among participants (Zhong *et al.*, 2020).

3.0 Discussion

Research revealed that there are a number of banks and organizations that use the distributed ledger technology for payments and settlements. However, there exists a huge array of flaws in these systems. Most of the research work focuses on loan disbursal only. The process of loan management does not merely include disbursal. There has been no research on how these loans disbursed can be tracked and traced. A comprehensive strategy of ensuring that the loan amount is used for the right and legitimate payments needs to be formulated.

Also, in the blockchain world, we deal with tokens or cryptocurrency. To completely integrate digital currencies with physical currency is a huge challenge. This task of connecting centralized ledgers to decentralized ledgers securely is unexplored in the research till date. One possible solution for connecting digital and physical currency would be a direct conversion process following current market value. This majorly poses two threats; the market value fluctuates every second and the process of payment has a high latency. Consider a case wherein a person pays an amount and initiates payment. On initiating a payment, the physical currency will be converted to a digital currency according to the current market value. If the transaction is completed and the currency value falls, the receiver receives an amount less than expected. Thus, this solution needs to be discarded.

A simpler approach would be to connect a person's digital wallet to his physical bank accounts. So, when a person receives a loan in his bank account, the cash will be converted to digital tokens of the same value and can be used in making digital payments on blockchain.

Generally, in DLT based systems, the blockchain merely stores the transactions. The task of ensuring that only legitimate transactions are made is done manually. No research on tracking of where the funds go after they are transferred to the user's account is done yet. Thus, we also need to have self-executing scripts in order to automate the verification of payments and addition of these transactions to the blockchain. Smart contracts provide a way to do the same.

4.0 Blockchain and Smart Contracts

4.1 Blockchain basics

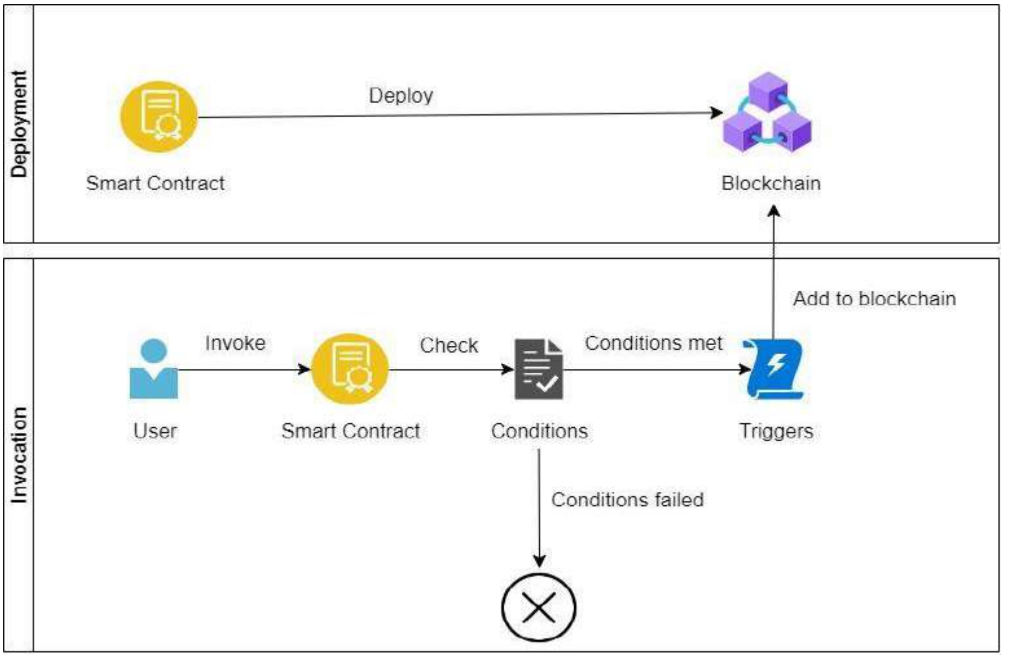
Based on Xue and Lu *et al.*, (2020), blockchain technology relies on three key components: cryptographic algorithms, consensus mechanisms, and distributed ledgers. Cryptographic algorithms, such as hash algorithms, ensure the integrity of recorded transaction data by preventing tampering (Li, Wu, Zhao, Lu, & Xue *et al.*, 2021). A consensus mechanism is used to validate the sequence and accuracy of the recorded data (Lu *et al.*, 2021). The data is stored on a distributed ledger, which is an immutable record of transactions shared among network participants (Sadeghi, Mahmoudi, & Deng *et al.*, 2021). These mechanisms ensure that the data stored on the blockchain is immutable, endorsable, and trackable (Moosavi, Naeni, Fathollahi-Fard, & Fiore *et al.*, 2021).

Blockchain platforms, such as Ethereum and Hyperledger Fabric, enable secure tokenization of assets. Tokenization involves representing assets as tokens, such as non-fungible tokens, which allow participants to establish and transfer ownership using a consensus mechanism (Hamledari & Fischer *et al.*, 2021). Smart tokens have three attributes, including owner, type, and quantity, as described by Hyperledger Fabric (2019). The owner specifies the authorized network participants who can transfer or redeem the token, while the type defines the asset represented by the token, and the quantity denotes the number of units of the asset that the smart token represents. For example, a smart token of the type "US dollars" could represent 100 dollars.

4.2 Smart contracts

Smart contracts are a major advancement in blockchain technology. Figure 1 explains the working of smart contracts. They enable contract terms to be automatically executed through computerized protocols (Ahmadisheykhsarmast & Sonmez *et al.*, 2020). A wide variety of applications of blockchain depend on the usage of smart contracts. In the already existing systems, programs are written to perform specific tasks and these programs are mutable. Also, the data they are fed and the output they generate are stored in a mutable database. However, in the case of smart contracts, all the data generated is stored in the form of a transaction.

Figure 1: Working of Smart Contracts



Every time a smart contract is invoked, i.e., executed, a transaction is generated, and it gets stored on the blockchain indestructibly. These contracts have two main components: predefined conditions and the corresponding responses triggered when these conditions are met (Li, Wu, Zhao, Lu, & Xue *et al.*, 2021). The life cycle of a smart contract can be divided into four phases: initialization, configuration, self-execution, and completion (Zheng, Xie, Dai, Chen, Chen, Weng, & Imran *et al.*, 2019). During the initialization phase, stakeholders negotiate their rights and restrictions, and lawyers draft an agreement that software engineers transform into smart contracts (Das, Tao, & Cheng *et al.*, 2020). In the configuration phase, smart contracts are configured on the blockchain to create Blockchain-Based Smart Contracts (BBSCs) (Zhong *et al.*, 2020). In the self-execution phase, when the predetermined conditions of the contracts are met, such as material reception, the corresponding contractual responses are automatically executed (Li *et al.*, 2021).

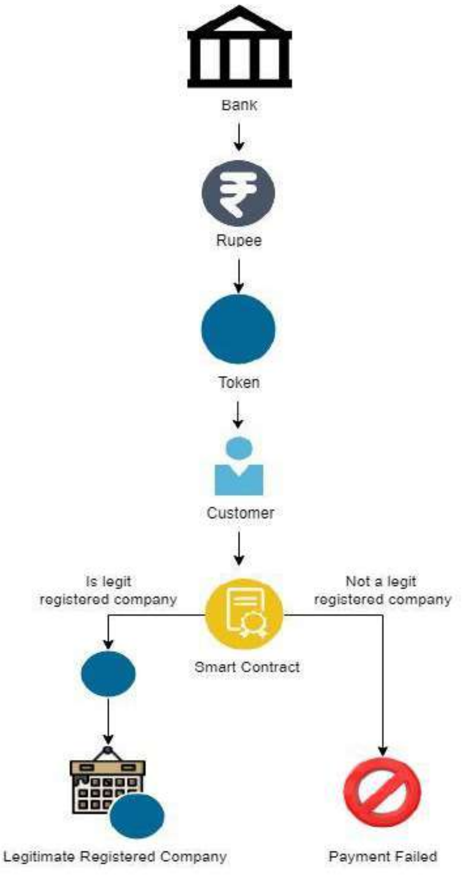
Finally, BBSCs complete their life cycle by updating transaction records on the blockchain (Zheng *et al.*, 2019). Since many transactions occur during the configuration, self-execution, and completion phases on the blockchain network, they are recorded on the blockchain for retrospective tracing (Li *et al.*, 2021).

5.0 Proposed Methodology

5.1 Workflow

The high-level concept explained in Figure 2, is that the process of loan management starts with the bank disbursing a loan. This loan amount will be converted to a token and transferred to the digital account of the customer. The customer can spend these tokens for the desired tasks following a process governed by smart contracts. The customer needs to access the portal and search for legitimate companies. The customer then initiates a transaction on which a smart contract will be invoked. The smart contract will check if the transaction is made to a legitimate company and only then will allow the transaction to be confirmed. The details of the transaction will be stored on the blockchain. If the conditions in the smart contract are not met, the transaction will fail, and the user will have to initiate payment again.

Figure 2: Workflow of the system



5.2 Main functionalities

- Build a complete Blockchain based system, that tracks right from disbursal to spending.
- Make provision for creating digital accounts and linking them to bank accounts.
- Write a smart contract to validate and verify recipients as registered entities.

The proposed system starts with the bank allocating funds. The allocated funds will then be converted to digital tokens in our case the Digital Rupee introduced by the Government of India. These tokens will be transferred to the user's digital account. Hereafter, the user can initiate a payment to a merchant of his choice. If the payment is initiated to a registered legitimate entity the smart contract proceeds with the execution of the transaction. Else, the transaction will fail.

6.0 System Architecture

6.1 Linking accounts

The digital account from our system and the traditional bank account needs to be linked in some way to achieve this we propose a smart contract-based account mapping scheme where the bank would have to send us a message with the linked accounts belonging to the same user after the user enters his digital account number in the bank's system. Here we have given access to multiple owners/banks even the owners can be added or removed. Only the current owner can add or remove the owner. All the physical and digital accounts that are linked to the specific bank accounts are stored in an array. So, the bank/owner will be able to see its all-linked accounts by simply entering the bank address in the function "getLinkedAccounts". Only banks/owners have this functionality.

6.2 Verification of transactions

Using smart contract, we first store the list of valid addresses. When user initiates the transaction the system checks if the entered payee is valid using the list of valid addresses and given there is enough balance the transaction executes successfully.

6.3 Register new payees

There is one admin for example the government that would validate and register new payee into our system by adding their address to the list of valid/verified addresses. Only the admin has exclusive rights to do so also, they can invalidate the previously registered payee.

6.4 Token transactions

Once the loan is granted the bank transfers tokens to the user's account which they can use to pay to the appropriate payees. These tokens can then be sent back to the bank to be converted to the traditional currency.

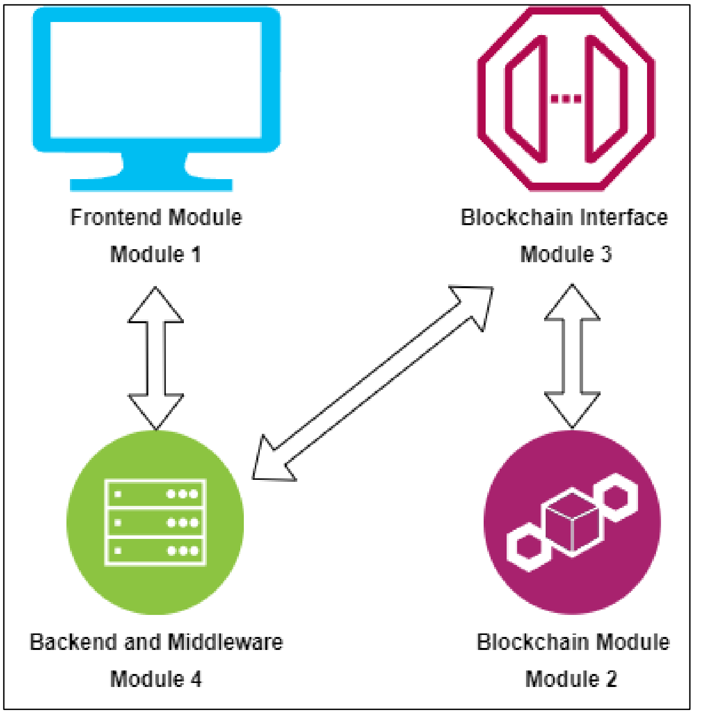
6.5 Tracing of tokens

Since our system is based on a public Blockchain platform we can use the infrastructure provided by the network to track transaction activities across the accounts in our system.

6.6 System connections

The system proposed in Figure 3, consists of 4 individual modules interconnected to each other. For the three-tier system we need a UI module, a Rest API, the blockchain and the database. The front-end module will serve as a platform to access the system. The system will use a blockchain API service to connect the database to an actual blockchain.

Figure 3: System Connections



7.0 Technologies Used

7.1 WEB3J

The Interface with blockchain is made possible by the solidity programming language and the web3j library. Web3j is a java specific tool that allows us to convert the ABI and BIN files obtained by compiling the solidity program into JAVA code as wrapper around the RPC calls to the ABI and contract deployed on the Particular blockchain network.

7.2 WEB3J supports

- Ethereum's JSON-RPC client API over HTTP and IPC are completely implemented.
- Ethereum wallet
- We can create, deploy, and transact smart contracts by calling them from native java code using smart contract wrappers generated by web3j.

To access data on the Ethereum blockchain or initiate transactions, a software application needs to connect to an Ethereum node. All Ethereum clients have a standard set of methods, based on the JSON-RPC specification, that applications can use regardless of the specific client or node implementation. JSON-RPC is a lightweight protocol for remote procedure calls that defines rules for processing data structures. It can be used across various communication channels, including HTTP and message passing environments, and relies on JSON (RFC 4627) as the data format.

8.0 Comparison between Systems

In Table 1 below, we have compared our system to previous done research work.

Table 1: Comparison between Systems

	Ahmadishey khsarmast and Sonmez (2020)	Das *et al.* (2020a)	Hamledari and Fischer (2021a)	Liupengfei WU, Weisheng LU, Jinying XU	This study
Payment freezing	Explored	Not discussed	Not discussed	Explored	Explored
Customer to payee payment	Not discussed	Not discussed	Not discussed	Not discussed	Explored
Payment validation	Semi-autonomous validation using Microsoft Project addon	Manual validation	Manual validation	Semi Autonomous validation using PMA.	Fully automated based on blockchain
Payment finalization (payment condition check)	Payment conditions checked manually	Smart contract checks payment conditions automatically	Computer vision-based solutions check conditions automatically.	Smart contract checks payment conditions automatically.	Fully automated based on blockchain
Payment completion	Peer to peer payment (Cryptocurrency)	Bank to bank transfer (Electronic transfer)	Peer to peer payment (Cryptocurrency)	Bank to bank transfer (token)	Peer to peer payment (Cryptocurrency)

9.0 Testing and Deployment

After implementation of separate module, we employ black box testing strategy individually. Only after which we go on for integration testing. Unit testing will be done using Postman API and Remix IDE. Performance testing will be done using Selenium. The system will be tested on Polygon Mumbai Testnet and deployed on a Linux server running Ubuntu x86-64.

10.0 Conclusion

The current financial scenario is full of fraudulent activities and lacks transparency. We thus propose a blockchain based system that uses smart contracts for making transactions. The system records all the transactions on a distributed ledger which can be accessed transparently and is immutable. We also provide a solution for connecting centralised and decentralised ledgers securely. This system provides an additional benefit of cross border payments and transactions.

References

Ahmadisheykhsarmast, S., & Sonmez, R. (2020). A smart contract system for security of payment of construction contracts. Automation in Construction, 120, 103401. https://doi.org/10.1016/j.autcon.2020.103401

Amamiya, M. (2019). Should the Bank of Japan issue a digital currency? Bank of Japan.

Bank of Japan. (2017). BOJ/ECB joint research project on distributed ledger technology. European Central Bank.

Das, M., Tao, X., & Cheng, J. C. (2020). A secure and distributed construction document management system using blockchain. In Proceedings of the 18th International Conference on Computing in Civil and Building Engineering. Springer.

Garratt, R. (2019). CAD-coin versus Fedcoin.

Hamledari, H., & Fischer, M. (2021). The application of blockchain-based crypto assets for integrating the physical and financial supply chains in the construction & engineering industry. Automation in Construction, 127, 103711. https://doi.org /10.1016/j.autcon.2021.103711

Khara, R., Pomendkar, D., Gupta, R., Hingorani, I., & Kalbande, D. (2020). Micro loans for farmers. 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), 1–5. https://doi.org/10.1109/icccnt 49239.2020.9225577

Li, X., Wu, L., Zhao, R., Lu, W., & Xue, F. (2021). Two-layer Adaptive Blockchain-based Supervision model for off-site modular housing production. Computers in Industry, 128, 103437. https://doi.org/10.1016/j.compind.2021.103437

Lu, W., Wu, L., Zhao, R., Li, X., & Xue, F. (2021). Blockchain Technology for Governmental Supervision of Construction Work: Learning from Digital Currency Electronic Payment Systems. Journal of Construction Engineering and Management, 147(10). https://doi.org/10.1061/(asce)co.1943-7862.0002148

Moosavi, J., Naeni, L. M., Fathollahi-Fard, A. M., & Fiore, U. (2021). Blockchain in supply chain management: A review, bibliometric, and network analysis. Environmental Science and Pollution Research International. https://doi.org/10.1007/s11356-021-13094-3

Peters, G., Panayi, E., & Chapelle, A. (2015). Trends in Crypto-Currencies and Blockchain Technologies: A Monetary Theory and Regulation perspective. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2646618

S, N. a. K., A, N. D. K., Gowtham C, N., & SahayaBeniPrathiba, N. (2021). Decentralized loan management application using smart contracts on block chain. In Advances in Parallel Computing. https://doi.org/10.3233/apc210020

Sadeghi, M., Mahmoudi, A., & Deng, X. (2021). Adopting distributed ledger technology for the sustainable construction industry: Evaluating the barriers using Ordinal Priority Approach. Environmental Science and Pollution Research, 29(7), 10495–10520. https://doi.org/10.1007/s11356-021-16376-y

The Monetary Authority of Singapore (MAS). (2017). MAS working with industry to apply Distributed Ledger Technology in securities settlement and cross-border payments. Monetary Authority of Singapore.

Wang, H., Guo, C., & Cheng, S. (2019). LoC — A new financial loan management system based on smart contracts. Future Generation Computer Systems, 100, 648–655. https://doi.org/10.1016/j.future.2019.05.040

World Bank. (2020). Exploring blockchain for disbursement traceability. World Bank Group (Finance and Accounting).

Wu, L., Lu, W., & Xu, J. (2022). Blockchain-based smart contract for smart payment in construction: A focus on the payment freezing and disbursement cycle. Frontiers of Engineering Management, 9(2), 177–195. https://doi.org/10.1007/s42524-021-0184-y

Xue, F., & Lu, W. (2020). A semantic differential transaction approach to minimizing information redundancy for BIM and blockchain integration. Automation in Construction.

Yang, Q., Xu, F., Zhang, Y., Liu, F., Hu, W., & Liao, Q. (2018). Design and Implementation of a Loan System Based on Smart Contract. In: Qiu, M. (eds) Smart Blockchain. SmartBlock 2018. Lecture Notes in Computer Science, vol 11373. Springer, Cham. https://doi.org/10.1007/978-3-030-05764-0_3

Zheng, Z., Xie, S., Dai, H., Chen, W., Chen, X., Weng, J., & Imran, M. (2019). An overview on smart contracts: Challenges, advances and platforms. Future Generation Computer Systems, 105, 475–491. https://doi.org/10.1016/j.future.2019.12.019

Zhong, B., Wu, H., Ding, L., Luo, H., Luo, Y., & Pan, X. (2020). Hyperledger Fabric-based consortium blockchain for construction quality information management. Frontiers of Engineering Management, 7(4), 512–527. https://doi.org/10.1007/s42524-020-0128-y

CHAPTER 4

A Prospect for Networking and Quantum Communication beyond Conventional Limits

Priya Khune and Rishav Pandey***

ABSTRACT

The quantum internet represents a paradigm shift in communication networks, promising ultra-high security through unbreakable encryption and distributed quantum computing grounded in quantum mechanics. Despite substantial advances in quantum computing technologies, the quantum internet remains in its early stages. This paper explores the theoretical foundations of the quantum internet alongside state-of-the-art research, addressing the main challenges faced in developing scalable quantum communication networks. It delves into the application areas of the quantum internet, highlighting the pivotal roles of quantum entanglement and Quantum Key Distribution (QKD). Additionally, the paper intends a roadmap for future research, identifying key milestones and open problems in quantum networking. Furthermore, the possibility of integrating quantum and classical computing networks is examined, suggesting hybrid systems and predicting breakthrough technologies that may bring the quantum internet closer to reality. Quantum effects enable the development of unique protocols, such as QKD, which uses quantum mechanics to detect and prevent unauthorized access to communication channels.

Keywords: Distributed Quantum Computing; Entanglement; Hybrid Quantum-Classical Systems; Quantum Communication; Quantum Cryptography; Quantum Internet; Quantum Key Distribution (QKD); Quantum Networking.

1.0 Introduction

The quantum internet heralds a new frontier in communication networks, leveraging quantum mechanics to communicate information. Unlike classical communication systems based on bits, quantum networks utilize quantum bits, or qubits, which rely on principles like superposition and entanglement. These principles enable revolutionary advancements in secure communication, distributed quantum computing, teleportation, and quantum sensing. The introduction of qubits fundamentally changes the way data is processed and transmitted. Superposition allows qubits to represent multiple states simultaneously, greatly enhancing computational power.

**Corresponding author; Assistant Professor, Department of CSE, MIT, Pune, Maharashtra, India (E-mail: priyakhune05@gmail.com)*
***Student, Department of CSE, MIT ADT, Pune, Maharashtra, India*
(E-mail: rishavpandey24jan@gmail.com)

Entanglement creates a direct and secure correlation between qubits over long distances, providing a robust framework for secure communication. While significant progress has been made through small-scale quantum network experiments, the realization of a global quantum internet poses immense technical, theoretical, and infrastructural challenges. Issues such as maintaining quantum coherence over long distances, overcoming hardware limitations, and integrating quantum systems with classical infrastructures are critical. This paper provides a comprehensive discussion of these challenges and the potential of quantum internet applications (Nielsen *et al.*, 2000). Moreover, the broader implications of the quantum internet include its transformative impact on global communication and data security. It promises to redefine the landscape of information exchange, enabling unprecedented levels of privacy and efficiency. By addressing the limitations of classical networks, the quantum internet opens up new possibilities for innovation across diverse fields, from healthcare to artificial intelligence. The exploration of these advancements forms the crux of this paper.

2.0 Theoretical Background of the Quantum Internet

The quantum internet is grounded in the ideologies of quantum mechanics, which introduce revolutionary concepts like *superposition*, *predicament*, and the *no-cloning theorem*. These principles enable functionalities impossible with classical networks, such as unbreakable security, quantum teleportation, and distributed quantum computing. Below, we delve into the theoretical aspects in greater detail, integrating mathematical representations to highlight the nuances of quantum mechanics.

2.1 Superposition

Superposition is the essential property of quantum systems where a quantum bit (qubit) can exist in multiple states simultaneously. Unlike a classical bit, which exists strictly as 0 or 1, a qubit is represented as:

$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, $|\alpha|2 + |\beta|2 = 1$

Here, $|0\rangle$|0\rangle$|0\rangle$ and $|1\rangle$|1\rangle$|1\rangle$ represent the classical states, and α\alpha$\alpha$ and β\beta$\beta$ are complex probability amplitudes. The probabilities of finding the qubit in $|0\rangle$|0\rangle$|0\rangle$ or $|1\rangle$|1\rangle$|1\rangle$ upon measurement are given by $|\alpha|2$|\alpha|^2$|\alpha|2$ and $|\beta|2$|\beta|^2$|\beta|2$, respectively, with the normalization condition:

$|\alpha|^2+|\beta|^2=1$

This property exponentially enhances the parallel processing capability of quantum systems, enabling them to perform tasks like data search or cryptanalysis significantly faster than classical systems.

2.2 Entanglement

Entanglement is an occurrence where two or more qubits become correlated in such a way that the state of one qubit instantly determines the state of the other, regardless of the distance between them. For two entangled qubits, the composite system can be represented as:

$$|\psi\rangle = 1/\sqrt{2}\,(|00\rangle + |11\rangle)$$

This means that if one qubit is leisurely and found to be ∣0⟩|0\rangle∣0⟩, the other will immediately collapse to ∣0⟩|0\rangle∣0⟩ as well, and similarly for ∣1⟩|1\rangle∣1⟩.

Entanglement is a foundation stone of quantum networking because:

- *Security:* It makes certain that any attempt to intercept or tamper with the quantum state disrupts the entanglement, alerting the communicating parties to the intrusion.
- *Quantum Key Distribution (QKD):* Practices like E91 leverage entanglement to securely exchange cryptographic keys.
- *Quantum Teleportation:* Entanglement enables the transfer of quantum information between distant nodes, which is fundamental to the quantum internet.

2.3 The no-cloning theorem

The no-cloning hypothesis prohibits the creation of an exact copy of an arbitrary quantum state, ensuring inherent security in quantum communication. Formally, if |ψis a quantum state, there is no unitary operator U that satisfies:

$$U(|\psi\rangle \otimes |e\rangle) = |\psi\rangle \otimes |\psi\rangle$$

Here, $|e\rangle$ represents an auxiliary state.

This principle makes eavesdropping detectable since any challenge to measure or copy the quantum state introduces disturbances observable to the sender and receiver. This property underpins secure quantum protocols, including Quantum Key Distribution (QKD).

2.4 Quantum channels and noise

Quantum announcement occurs over quantum channels, such as optical fibers or free-space links. However, these channels are susceptible to noise, causing decoherence and loss of quantum information. The evolution of a quantum state ρ in a noisy channel is described using a quantum operation $\mathcal{E}$:

$$P' = \mathcal{E}(\rho) = \sum k\ Ek\rho Ek\dagger$$

Here, {Ek} are Kraus operators satisfying:

$$\sum k\ Ek\dagger Ek = I$$

Where I is the identity matrix. These operators model the effects of noise, such as photon loss or phase errors. Overcoming these challenges involves techniques like:

- *Quantum Error Correction:* Encoding information redundantly across multiple qubits to detect and correct errors.
- *Quantum Repeaters:* Devices that extend the range of quantum communication by segmenting the channel and re-establishing entanglement at intermediate points.

2.5 Quantum Key Distribution (QKD)

QKD leverages quantum procedure to establish a secure statement by detecting any spying attempt. The BB84 protocol, for instance, relies on the random measurement of qubits in different bases. The steps of BB84 are:

- *State Preparation:* The sender, Alice, encodes qubits randomly in one of two bases (e.g., rectilinear or diagonal).
- *Transmission and Measurement:* The receiver, Bob, measures the qubits in randomly chosen bases.
- *Key Reconciliation:* Alice and Bob compare a subsection of their measurement results to detect errors.

The security of QKD is mathematically ensured by the detection of disturbances introduced by an eavesdropper. The error rate eee is calculated, and if eee exceeds a predefined threshold, the communication is aborted.

2.6 Mathematical framework for quantum networking

The quantum internet relies on a network of quantum nodules connected by quantum channels. Key equations describing its operations include:

- *Entanglement Fidelity:*

F(ρ,σ) = (Trρσρ)2F(\rho, \sigma) = \left(\text{Tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}} \right)^2F(ρ,σ) = (Trρ σρ)2

This measures the reliability of entanglement distribution.

- *Channel Capacity:* Channel Capacity: The maximum amount at which quantum information can be communicated, defined using the Holevo quantity.

Mathematically, the Holevo quantity is defined as:

X = S(ρ) - ∑I pi S(ρi)

Where:

- χ is the Holevo quantity
- S(ρ) is the von Neumann entropy of the input state ρ
- pi are the probabilities of the different measurement outcomes
- ρi are the corresponding post-measurement states
- S(ρi) is the von Neumann entropy of the post-measurement state ρi

The Holevo quantity provides an upper bound on the accessible information, which is the maximum amount of classical information that can be extracted from a quantum system. By building upon these theoretical principles, the quantum internet holds the potential to redefine secure communication, collaborative computing, and scientific discovery. These advancements rely heavily on the interplay between mathematical formulations and experimental innovations.

Quantum Repeaters: Quantum repeaters are critical components for overcoming the challenges of long-distance quantum communication. These devices work by distributing the communication aloofness into smaller segments, creating entanglement locally, and then connecting these segments through a process known as entanglement swapping. By storing entangled states and systematically transferring them, quantum repeaters counteract the effects of decoherence, ensuring reliable communication over extended distances (Briegel *et al.*,1998). Advanced implementations of quantum repeaters use memory buffers and error correction techniques to maintain high-fidelity quantum states.

Error Correction in Quantum Communication: Error rectification plays a vital role in preserving the truthfulness of quantum states over long distances. Quantum error correction codes, such as surface codes and topological codes, help mitigate errors arising from environmental noise and operational imperfections. These codes operate by encoding quantum information through multiple physical qubits, enabling the detection and correction of errors without directly observing the significant state. This process ensures that the entangled states required for quantum communication remain intact, even in the presence of disturbances.

Satellite-Based Quantum Communication: Recent advancements have demonstrated the potential of using satellites to distribute entangled states over vast distances. By deploying quantum systems in space, researchers have achieved entanglement distribution across hundreds of kilometres, bypassing many terrestrial limitations such as optical fiber losses. Satellite-based quantum communication represents a promising avenue for establishing global quantum networks, connecting distant locations with high-speed, secure quantum links.

Hybrid Systems and Quantum Networks: The integration of quantum repeaters, error correction, and satellite systems forms the backbone of a scalable quantum network. These systems work in tandem to extend the range and reliability of quantum communication, creating a framework for future quantum internet infrastructures. The development of hybrid systems that combine terrestrial and satellite-based approaches will be crucial for achieving global connectivity, enabling seamless communication between quantum nodes.

In summary, the interplay of quantum repeaters, error correction, and satellite technologies provides a robust foundation for addressing the challenges of long-distance quantum communication. By leveraging these advancements, researchers are paving the way for scalable and reliable quantum networks that can span continents and potentially connect the entire globe.

3.0 Modules of Quantum Internet

3.1. Core Modules of the Quantum Internet

3.1.1 Quantum Nodes

Quantum knots are the primary units of a quantum internet. They are analogous to computers or servers in the classical internet but are equipped to handle qubits.

Components:

- *Quantum Processors*

  - Perform quantum operations like superposition, entanglement, and quantum gates.
  - Examples: Quantum chips like IBM's superconducting processors or trapped-ion systems.
- *Quantum Memory*
  - Stores quantum information (qubits) temporarily for processing or communication.
  - Requires coherence preservation to maintain quantum state integrity.
- *Classical Interfaces*
  - Facilitate integration with classical systems for hybrid communication.

3.1.2 Quantum channels

These are the communication links through which qubits are transmitted.

Types

- *Optical Fiber Channels*
  - Use photons to carry qubits over terrestrial distances.
  - Limitation: Signal attenuation beyond ~100 km without repeaters.
- *Satellite-Based Channels*
  - Distribute entangled particles over long distances, overcoming the limitations of terrestrial fibers.
  - Example: China's *Micius satellite*, which achieved entanglement over 1,200 km.
- *Properties*
  - Qubits are delicate and can decohere due to environmental factors.
  - No-cloning hypothesis ensures that quantum information cannot be duplicated, adding intrinsic security.

3.1.3 Quantum repeaters

Quantum repeaters extend the range of quantum communication by overcoming signal loss and decoherence in quantum channels.

- *Functionality*
  - Generate entanglement between nodes in the network.
  - Use entanglement swapping to create entanglement between distant nodes.
  - Store intermediate quantum states in quantum memory to synchronize the process.
- *Challenges*
  - Developing repeaters with long coherence times.
  - Error correction to prevent signal degradation.

3.1.4 Entanglement distribution systems

Entanglement is the backbone of the quantum internet, enabling instantaneous correlation between qubits.

- *How it Works*
    - A source creates entangled pairs of photons.
    - One photon is sent to Node A, and the other to Node B.
    - Evaluating one photon instantly determines the state of the other, regardless of distance.
- *Applications*
    - Quantum Key Distribution (QKD).
    - Teleportation of quantum states.

3.1.5 Classical communication channels

Classical channels are required for transmitting auxiliary information, such as the results of quantum measurements.

- *Role*
    - Complements quantum channels by enabling error correction and verification.
    - Facilitates reconstruction of teleported states.

3.2 Operational principles of the quantum internet

3.2.1 Entanglement-based communication

- *Procedure*
    - Create and distribute entangled particles to nodes.
    - Use the shared entanglement for tasks like teleporting quantum states or sharing cryptographic keys.
- *Advantages*
    - High security due to instant detection of eavesdropping.
    - Faster-than-light state correlation (not signaling).

3.2.2 Quantum Key Distribution (QKD)

QKD allows dual parties to generate and share encryption keys securely.

- *Example Protocols*
    - BB84: Uses qubits to detect interception during key exchange.
    - E91: Based on entanglement, ensuring unbreakable security.
- *Steps*
    - Entangled qubits are distributed.
    - Measurement results are shared via classical channels.
    - Any interception alters the quantum state, revealing an attack.

3.2.3 Quantum teleportation

Teleportation transfers quantum information without physically sending the particle.

- *Process*
    - Entangled particles are distributed.

  - The sender implements a quantum measurement on the qubit to be teleported and their local entangled particle.
  - The result is sent to the receiver via aorthodox channel.
  - The receiver reconstructs the qubit using their entangled partner.

3.3 Functional layers of the quantum internet

The quantum internet operates in distinct functional layers similar to classical internet models:

3.3.1 Physical layer

- Responsible for the transmission of qubits through optical fibers, satellites, or free-space optics.
- Handles challenges like signal loss and noise.

3.3.2 Link layer

- Establishes and maintains quantum links between adjacent nodes.
- Includes entanglement generation, verification, and swapping protocols.

3.3.3 Network layer

- Manages routing and entanglement distribution across the network.
- Decides how to allocate resources like quantum repeaters and memories.

3.3.4 Transport layer

- Ensures end-to-end communication, including teleportation and QKD.
- Handles error correction to mitigate decoherence effects.

3.3.5 Application layer

- Provides quantum internet services to end-users.
- Examples:
  - Secure video calls using QKD.
  - Distributed quantum computing.

3.4. Challenges of building a quantum internet

- *Fragility of qubits*
  - Qubits are highly sensitive to noise, requiring robust error correction mechanisms.
- *Resource intensity*
  - Quantum repeaters, entanglement sources, and quantum memories are costly and technologically complex.
- *Scaling up*
  - Current quantum networks are limited in scale. Expanding to global quantum internet requires breakthroughs in hardware and protocols.

Integration with Classical Systems

- Hybrid systems must seamlessly combine quantum and classical communication technologies.

3.5 Applications of the quantum internet

- *Unhackable Communication*
  - Using QKD for ultra-secure communication in finance, government, and military.
- *Distributed Quantum Computing*
  - Connecting quantum computers across the globe to perform collaborative computations.
- *Scientific Collaboration*
  - Sharing quantum data among research labs for experiments like simulating quantum systems.
- *Enhanced Cryptography*
  - Developing cryptographic protocols that are resistant to attacks from quantum computers.

3.6 Quantum routers and entanglement distribution

Quantum routers enable the distribution of entangled qubits across networks, serving as the backbone of quantum infrastructure. A Quantum Information Service Provider (QISP) may emerge as a central system offering entanglement as a service to remote clients. These routers facilitate scalable quantum communication and create pathways for hybrid quantum-classical systems. Quantum routers are pivotal in optimizing network reliability and ensuring efficient entanglement distribution over wide areas. The development of efficient quantum routers involves overcoming challenges in scalability, synchronization, and error management. These devices must operate seamlessly within a global network, ensuring that entangled states are transmitted without degradation. Innovations in router technology, coupled with advances in quantum error correction, are essential for achieving reliable and widespread entanglement distribution.

4.0 Key Challenges and Open Problems

4.1 Scalability and quantum decoherence

One of the foremost trials in realizing a functional quantum internet lies in addressing issues related to scalability and quantum decoherence. Quantum information, represented as qubits, is inherently fragile and highly susceptible to environmental disturbances. Even minor interactions with external factors can disrupt quantum states, leading to decoherence and rendering quantum communication ineffective. Maintaining the integrity of these states over long distances or extended periods is a fundamental requirement for a scalable quantum network.

To achieve scalability, researchers must develop advanced quantum repeaters capable of sustaining entanglement over vast distances. These devices work by segmenting the communication pathway into smaller sections and repeatedly creating and swapping entanglement between adjacent nodes. Each segment ensures local entanglement integrity, which is then systematically extended across the network using entanglement swapping techniques. However, the practical implementation of quantum repeaters requires addressing several challenges, including the need for quantum memories with high storage fidelity and long retention times.

Quantum error correction is another pivotal aspect of scalability. Unlike classical systems, quantum error correction does not involve direct observation of qubits, as this would collapse their quantum states. Instead, information is redundantly encoded across multiple qubits, agreeing errors to be detected and corrected indirectly. Surface codes and topological error correction techniques have shown promise in mitigating the impact of noise and environmental disturbances. These methods will play a fundamental role in preserving the integrity of quantum states, enabling the development of larger, more robust quantum networks.

Furthermore, the increase of scalable quantum hardware and algorithms is essential. Innovations in qubit design, such as superconducting qubits, trapped ions, and photonic qubits, are critical for building quantum systems proficient of handling complex computations and large-scale networking. Each qubit technology offers unique advantages and experiments, and ongoing research is focused on optimizing their performance and compatibility with existing infrastructures.

Global collaboration between research institutions, industry leaders, and governments is imperative to overcome scalability barriers. Investments in quantum technologies, combined with standardized protocols and interdisciplinary research efforts, will pave the way for the concrete realization of the quantum internet. By addressing the intertwined issues of decoherence, error correction, and hardware scalability, researchers aim to unlock the full potential of quantum communication networks.

4.2 Classical network integration

Integrating quantum and classical networks represents a significant challenge and an opportunity to influence the strengths of both technologies. As quantum computing continues to evolve, its potential to disentangle complex problems faster than classical computers opens new frontiers in fields like cryptography, materials science, and artificial intelligence. However, it also brings with it the need for new methods of communication and security.

A crucial aspect of this integration is the development of hybrid systems capable of bridging the two paradigms. Hybrid systems will need to effectively combine classical networking infrastructures with emerging quantum technologies, such as QKD and quantum teleportation. The goal is to ensure seamless communication between quantum and classical components, allowing for efficient data transfer while maintaining the unique benefits of quantum technologies (Wang *et al*.,. 2021). Quantum-safe encryption algorithms are essential in this context. Traditional encryption methods, such as RSA and ECC (Elliptic Curve Cryptography), rely on the computational effort of certain mathematical problems to secure data. However, these encryption schemes are vulnerable to future quantum attacks, specifically those based on Shor's algorithm, which could efficiently factor large numbers and solve discrete logarithms, thus breaking the foundation of current encryption techniques (Barros *et al.*, 2019).

To address this, quantum-safe or post-quantum cryptography (PQC) must be developed. These algorithms are designed to resist quantum-based attacks while still being compatible with classical infrastructure. They focus on new encryption methods that are resistant to quantum algorithms, ensuring that sensitive data remains secure even in a world where large-scale quantum computers are a reality (Lewis *et al.*, 2022).

4.3 Standardization and protocol development

The lack of standardized protocols for quantum communication hinders the scalability of quantum networks. International collaboration is necessary to establish universal standards for hardware interfaces, error correction, and communication protocols. Novel approaches, such as Quantum Communication as a Service (QCaaS), could redefine the future of quantum networking (Zhang *et al.*,2021).

Standardization ensures interoperability among various quantum systems, fostering a seamless global network. Establishing these standards requires input from academia, industry, and governmental bodies, emphasizing the need for a coordinated global effort. Unified protocols will facilitate the amalgamation of diverse quantum technologies, accelerating the improvement of practical quantum networks.

5.0 Application Areas of Quantum Internet

5.1 Distributed quantum computing

Distributed quantum computing is an innovative approach enabled by the quantum internet, wherein interconnected quantum processors collaborate to address complex problems that are infeasible for classical systems. Often referred to as "quantum cloud computing," this paradigm leverages the unique properties of quantum mechanics—such as entanglement and superposition—combined with advanced networking to create a cohesive computational framework. By sharing quantum states and pooling resources, distributed systems can perform tasks like optimization, simulation, and large-scale data analysis more efficiently than traditional methods.

5.1.1 Key principles and mechanics

- *Networked Qubits:* Distributed systems utilize entangled qubits across geographically separated nodes to maintain synchronized states and enable secure computations.

- *Quantum State Sharing:* Practices such as quantum teleportation and superdense coding facilitate the efficient exchange of information between nodes, minimizing communication delays.

5.1.2 Applications across domains

- *Logistics Optimization:* Distributed quantum computing can revolutionize logistics by solving combinatorial problems such as the traveling salesman problem. These problems scale exponentially with complexity, but quantum systems can evaluate multiple scenarios simultaneously through quantum parallelism.
  Example: Optimizing global delivery routes to reduce costs and increase efficiency.
- *Healthcare Innovations:* Quantum systems enable detailed simulations of molecular interactions, aiding in drug discovery and personalized medicine. They can also analyze large genetic datasets to uncover patterns linked to diseases.
  Example Equation: The Variational Quantum Eigensolver (VQE) calculates molecular ground states accurately: H∣ψ⟩=E∣ψ⟩H |\psi\rangle = E |\psi\ rangleH∣ψ⟩=E∣ψ⟩ where HHH is the molecular Hamiltonian, ∣ψ⟩|\psi\ rangle∣ψ⟩ represents the quantum state, and EEE is the energy.
- *Materials Science:* Quantum computing facilitates the precise modeling of materials like superconductors, contributing to advancements in energy storage and nanotechnology.
  Example: Simulation of new battery materials for improved energy efficiency.

5.1.3 Advantages of distributed quantum computing

- *Enhanced Computational Power:* Combines resources from multiple quantum nodes, exponentially increasing capacity.
- *Resource Optimization:* Allows smaller processors to contribute, reducing hardware costs.
- *Fault Tolerance:* Incorporates quantum error correction to maintain computational integrity despite decoherence.

5.1.4 Technical challenges and future directions

- *Entanglement Distribution:* Efficient distribution of entangled qubits is critical, requiring quantum repeaters and satellite-based systems.
- *Synchronization:* High-precision timing and communication protocols are essential for coordinating quantum processors.
- *Scalability:* Expanding the network to thousands of nodes necessitates innovations in hardware and software infrastructure.

5.2 Quantum sensing and teleportation

Quantum sensors, leveraging quantum effects, offer unparalleled precision in environmental monitoring and other measurements. Applications include high-resolution imaging, precision navigation, and early detection of environmental changes. Quantum teleportation, the transfer of quantum states between distant locations, holds the potential to transform data transmission methods (Zhang *et al.*, 2019). Teleportation ensures secure and instantaneous transmission of sensitive data, reinforcing the reliability of quantum communication systems.

Advances in quantum sensing could revolutionize sectors such as healthcare, defense, and ecological science. The ability to detect minute changes in physical parameters enables the development of next-generation diagnostic tools and monitoring systems. Meanwhile, quantum teleportation offers a foundation for future communication networks that are both secure and efficient, bridging the gap between theoretical research and practical implementation.

5.3 Ultra-secure communication

5.3.1 Quantum communication and teleportation (elaborated)

Quantum Communication: Quantum communication is the process of spreading information using the fundamental principles of quantum mechanics. Unlike classical communication that is dependent on binary bits (0 and 1), quantum communication uses qubits, which exploit superposition and entanglement to enable secure and efficient data transfer.

Key principles

Quantum Entanglement:

- When two elements are entangled, their quantum states are intrinsically linked, no matter how far apart they are. Quantifying one particle instantly affects the state of the other. This unique property forms the backbone of quantum communication.
- Example: In entanglement-based communication systems, an entangled pair of photons is distributed between two parties. Any change to one photon's state immediately reflects in the other, enabling secure communication.

Superposition: Qubits can be present in multiple states simultaneously (e.g., a combination of 0 and 1). This property allows quantum systems to encode and process much more information than classical systems.

No-Cloning Theorem: Quantum states cannot be copied perfectly. This ensures that any attempt to divert or clone quantum information introduces detectable errors, making eavesdropping impossible.

Features of quantum communication

- *Unbreakable Security:* Quantum Key Distribution (QKD) allows secure key exchange by detecting any eavesdropping attempt. Protocols like BB84 ensure the cryptographic keys remain private and tamper-proof.

- *Immediate Detection of Interception:* In classical communication, detecting a breach often happens after the fact. In quantum systems, the act of interception changes the quantum state, alerting the sender and receiver in real time.
- *Long-Distance Communication:* Satellite-based quantum communication enables the distribution of entangled particles across hundreds or even thousands of kilometers, overcoming limitations of terrestrial optical fibers.

Applications

- *Financial Sector*
  - Banks and financial institutions can use quantum communication to secure online transactions and prevent cyber-attacks on sensitive data.
  - Example: Protecting SWIFT transactions in international banking.
- *Defense and National Security:* Governments and defense agencies can use quantum communication to safeguard military communications and classified intelligence from foreign cyber threats.
- *Healthcare:* Ensuring secure transmission of medical records, research data, and diagnostic reports to prevent data breaches in the healthcare industry.
- *Global Business:* Corporations can rely on quantum-secure channels to exchange trade secrets, contracts, and confidential data globally without fear of interception.

Challenges in Quantum Communication

- *Environmental Sensitivity:* Quantum states are delicate and can be disrupted by noise or environmental interference, leading to decoherence.
- *Infrastructure Development:* Building quantum repeaters, entanglement distribution systems, and satellite-based communication channels requires significant technological advancements.
- *Integration with Classical Networks:* Hybrid systems need to combine the strengths of quantum and classical communication, which demands standardized protocols and hardware.

Quantum Teleportation: Quantum teleportation encompasses transferring the quantum state of a particle from one location to another without physically affecting the particle itself. This phenomenon exploits quantum entanglement and classical communication channels to achieve secure and instantaneous state transfer.

How Quantum Teleportation Works

- *Entanglement Generation:* Two particles (A and B) are entangled. Particle A remains with the sender, while particle B is sent to the receiver.
- *Quantum State to be Teleported:* A third particle (C), whose quantum state needs to be teleported, interacts with particle A at the sender’s location. This interaction generates a specific measurement result that encodes the relationship between particles A and C.

- *Classical Information Transfer:* The sender measures particles A and C, collapsing their quantum states. The measurement results are then sent to the receiver via aorthodox communication channel.
- *Reconstruction at Receiver's End:* Using the classical information and particle B (the entangled partner), the receiver reconstructs the exact quantum state of particle C. The state is recreated at the receiver's location, effectively "teleporting" it.

Advantages of Quantum Teleportation

- *Secure Data Transfer:* Since the quantum state is transferred rather than the particle itself, interception during transit is virtually impossible.
- *No Physical Transmission Medium:* Quantum teleportation does not rely on traditional mediums like cables or electromagnetic waves, reducing the risks of physical disruptions or attacks.
- *Instantaneous State Transfer:* The entangled particles' correlation ensures the quantum state transfer happens instantaneously, regardless of the distance.

Applications

- *Secure Quantum Networks:* Considerable teleportation can be used to transfer sensitive quantum states between nodes in a quantum network, enabling secure communication.
- *Distributed Quantum Computing:* Quantum teleportation allows quantum processors to share and process quantum states across different locations, enhancing computational capabilities.
- *Advances in Cryptography:* By enabling the secure transfer of quantum states, teleportation lays the groundwork for new cryptographic techniques resistant to future quantum threats.

Challenges in Quantum Teleportation

- *Infrastructure*: Quantum teleportation relies on robust entanglement distribution systems and quantum repeaters, which are still in developmental stages.
- *Scaling Up:* Current teleportation experiments are limited to short distances and small-scale systems. Scaling this technology for global communication remains a challenge.
- *Reliance on Classical Channels:* While quantum teleportation uses entanglement, the need for classical communication channels introduces latency and potential vulnerabilities.

Future Potential of Quantum Communication and Teleportation

- *Global Quantum Networks:* The integration of quantum communication and teleportation technologies is expected to form the backbone of the quantum internet, connecting quantum computers and sensors globally.

- *Ultra-Secure Data Transmission:* These technologies will set new benchmarks for data privacy, enabling governments, corporations, and individuals to communicate without fear of cyber threats.
- *Interstellar Communication:* In the future, quantum teleportation could be a key enabler for secure communication with spacecraft or colonies on other planets, where classical communication delays are significant.

References

Nielsen, M. A., & Chuang, I. L. (2000). *Quantum computation and quantum information*. Cambridge University Press.

Briegel, H. J., Dür, W., Cirac, J. I., & Zoller, P. (1998). Quantum repeaters: The role of imperfect local operations in quantum communication. *Physical Review Letters, 81*(26), 5932-5935. https://doi.org/10.1103/PhysRevLett.81.5932

Zhang, R., Zhang, L., Yu, W., & Zhang, Z. (2019). Quantum Internet of Things: A review and perspective. *IEEE Access, 7*, 129129-129144. https://doi.org/10.1109/ACCESS.2019.2938311

Barros, M. S. L. P. S. L. B. E. (2019). Security for IoT using quantum cryptography. *International Journal of Quantum Information, 17*(8), 1937-1950. https://doi.org/10.1142/S021974991950050X

Lewis, L. S., *et al.* (2022). Applications of quantum sensors for IoT. *Nature Electronics, 2*, 254-263. https://doi.org/10.1038/s41586-022-00545-0

Zhang, L., *et al.* (2021). Quantum teleportation in practical quantum networks. *Quantum Information Processing, 20*(6), 1-16. https://doi.org/10.1007/s11128-021-03072-w

Wang, F., & Yang, Z. (2023). Quantum internet and future applications. *International Journal of Quantum Technology, 3*, 5-15.

CHAPTER 5

Social Media Analysis for Business Development using GenAI

Sonali Lunawat, Aman Kocharekar**,*
*Ishan Honmode***, Ashwin Lavate**** and Pushkar Sarda******

ABSTRACT

Within the highly competitive market environment, social media platforms play the role of dictating the public's perception of brands and products. Organizations frequently depend on consumer-generated content like reviews, comments, and tweets to evaluate the success of product launches and monitor constant customer feedback. Nonetheless, the manual examination of this massive volume of data is ineffective and may result in postponed understanding, thereby complicating the process. It is imperative for enterprises to react to public sentiment promptly. This project introduces a technique that is supposed to automate the process. Social media sentiment analysis for product performance assessment by enabling product owners to enter a brand designation, the tool retrieves recent entries or remarks from social media platforms and conducts a sentiment analysis to categorize them as positive or negative, or unbiased. Together with collecting sentiment results, the tool identifies crucial concepts linked to both affirmative and adverse provisions observations and generates a derived summary using a Large language model (LLM). The resulting information provides product owners with a comprehensive understanding of attitudes among consumers. This method will enable organizations to track the people's perception of their product launches, make data-driven decisions, and quickly resolve positive and negative feedback. It is intended to keep the business competitive with timely consumer content understanding, hence, improving the product and increasing customer involvement.

Keywords: Artificial Intelligence; Sentiment Analysis; Machine Learning; Natural Language Processing; Large Language Models (LLM).

1.0 Introduction

In today's digital landscape, social media serves as a crucial source for gathering customer feedback and understanding brand perception.

**Corresponding author; Faculty, Department of Computer, Pimpri Chinchwad College of Engineering & Research, Pune, Maharashtra, India (E-mail: sonali.lunawat@pccoer.in)*
***UG Student, Department of Computer, Pimpri Chinchwad College of Engineering & Research, Pune, Maharashtra, India (E-mail: aman.kocharekar_comp21@pccoer.in)*
****UG Student, Department of Computer, Pimpri Chinchwad College of Engineering & Research, Pune, Maharashtra, India (E-mail: ishan.honmode_comp21@pccoer.in*
*****UG Student, Department of Computer, Pimpri Chinchwad College of Engineering & Research, Pune, Maharashtra, India (E-mail: ashwin.lavate_comp21@pccoer.in*
******UG Student, Department of Computer, Pimpri Chinchwad College of Engineering & Research, Pune, Maharashtra, India (E-mail: pushkar.sarda_comp21@pccoer.in*

With millions of posts on platforms like Twitter, Facebook, and Instagram, companies struggle to analyze this vast data efficiently, making it challenging to extract real-time insights into public sentiment. Traditional manual sentiment analysis methods are often time-consuming and prone to errors, responses to customer needs, and market trends. To address these challenges, this project aims to develop an automated sentiment analysis tool that utilizes advanced Natural Language Processing (NLP) techniques and Machine Learning (ML) models. By categorizing customer comments, reviews, and posts as positive, negative, or neutral, the system provides valuable insights into public perception. Employing Large Language Models (LLMs) for keyword extraction and summarization enhances the efficiency of insights generation, allowing companies to make rapid, data-driven decisions. The use of key machine learning models like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks ensures accurate sentiment classification and scalability, enabling businesses to enhance customer engagement and optimize their marketing strategies effectively.

1.1 Machine Learning

Machine Learning (ML) is a key branch of Artificial Intelligence (AI) that enables computers to learn from data patterns and make decisions without explicit programming. Using algorithms and models, Machine Learning processes data to improve performance and provide predictive insights. Its applications range from image recognition and natural language processing (NLP) to recommendation systems and autonomous vehicles, fostering innovation across various industries. As a core component of data-driven decision-making and automation, Machine Learning is a vital force in the modern digital era.

1.1.1 Hugging face

Hugging Face marks a paradigm shift in the availability and dissemination of machine learning technologies, positioning it as an umbrella framework in current artificial intelligence research and development. This all-encompassing platform integrates various fundamental elements: a centralized Model Hub containing a vast assortment of pre-trained models; a strong Datasets repository that supports empirical investigations; and a collection of specialized libraries, including Transformers, Accelerate, and Diffusers, which collectively enhance efficiency.

Implementation of state-of-the-art machine learning architectures. The platform’s architecture showcases leading-edge innovation in how the same addresses collaborative AI development, implementation of Git-based version control methodologies for model management, and incorporation of automated CI/CD pipelines to ensure systematic model validation. Of particular importance is the platform’s Inference API, which abstracts the complexities of model deployment, thus reducing the barrier of entry to the practical application of AI research. The impact of this platform on the research community widely embraces the use of more than 100,000 pre-trained models and 20,000 datasets currently available, which accelerate reproducible research and move quickly from theoretical models to practical use. The adaptability of connecting academic research to industrial applications, thus promoting a more cohesive AI advancement ecosystem, is also indicated by the platform's enterprise-level features, including private model hosting and tailored training infrastructure.

1.2 Understanding sentiment analysis

Sentiment analysis is fundamentally the computational process of detecting, interpreting, and classifying the emotional and attitudinal tones behind texts through artificial intelligence and natural language processing (NLP). Sentiment analysis is different from keyword-based approaches in its context-based consideration of human language. Automatic identification, classification, and the ensuing recognition of emotions in the text process. They can be negative, positive, or neutral, aiding in getting a coherent interactive insight into textual data regarding their opinions and sentiments, oftentimes contributing verily to decision-wise and strategic actions to be taken onward.

1.3 Types of sentiment analysis

- Fine-grained sentiment analysis is more than just a binary view of sentiment classification; it however gives a continuous or ordinal sentiment score. For example, they might use a scale from -5 (extremely negative) to +5 (extremely positive) to convey that level of sentiment.
- Aspect-based sentiment analysis targets the emotions directed toward distinct components or attributes of an entity, usually a product service, or topic. For instance, in restaurant reviews, this technique can extract sentiments related to food quality, service, and ambiance. It consists of two major steps:
    - Entity Recognition: Identifying relevant aspects or entities mentioned in the text (e.g., "food," "service").
    - Sentiment Classification: Determining the sentiment associated with each identified aspect.
- Intent-based sentiment Analysis is a form of analysis in which overall feelings that transcend and their underlying reasons are unearthed. An instance would be that of displeasure by a customer service email expressing negative sentiment along with a request for information or help requesting information. This methodology tends to be:
    - Sentiment Classification: Identifying the overall sentiment of the text.
    - Intent Classification: Categorizing the purpose or intent behind the text (e.g., complaint, request, praise).
- *Emotion Detection:* Emotion detection in sentiment analysis seeks to identify specific emotional states conveyed in text, such as joy, sadness, anger, or fear. This method surpasses basic sentiment polarity by capturing the emotional nuances of

conversations, providing valuable insights into the mood or emotional tone of discussions, particularly in contexts such as social media or customer feedback.

1.4 How sentiment analysis works?

- *Data preparation:* The next task involves collecting textual data for analysis, such as social media posts, customer reviews, survey responses, and any other kind of written content. Pre-processing is done to clean and standardize the data. The pre-processing stage involves the removal of irrelevant items such as punctuation, special characters, and stop words (e.g. the, and) from the dataset preparing it for analysis.
- *Feature Engineering:* Features are extracted from the text data to facilitate sentiment detection. Common techniques include:
  - *Lexicon-Based Analysis:* Utilizing predefined dictionaries of positive and negative words, where the frequency of such terms is analyzed to infer overall sentiment.
  - *Part-of-Speech Tagging:* Identifying grammatical components (e.g., nouns, verbs, adjectives) to provide a contextual understanding of sentiment within sentences.
- *Machine Learning Algorithms:* The methods in sentiment analysis are: 1. Rule-based systems: these systems use predefined rules in keywords or phrase patterns to detect sentiment. They are easy to implement but cannot handle complex linguistic structures. 2. Machine Learning and Deep Learning Models: These models are trained on labeled datasets with annotated sentiments of positive, negative, or neutral so that the model may learn patterns correlating with the sentiment. Examples of popular algorithms for the task include SVMs and Naive Bayes classifiers.
- *Sentiment Classification:* After feature extraction, the machine learning model classifies the textual data into sentiment categories. These categories may range from broad classifications such as positive, negative, or neutral to more fine-grained sentiment scales capturing varying intensities.
- *Model Refinement:* Sentiment analysis models undergo iterative improvement. Model performance is evaluated using a testing dataset, and, if the accuracy is suboptimal, refinements are made by modifying the feature set, expanding the training dataset, or optimizing the machine learning algorithm. This iterative process ensures improved reliability and precision of sentiment analysis outcomes.

1.5 Evaluation metrics

Evaluation metrics are crucial in assessing the performance of sentiment analysis models. These metrics are commonly used since sentiment analysis is often treated as a classification problem. Some of the widely used classification metrics include:

- *Accuracy:* The ratio of correctly classified instances to the total number of instances. It is defined as:
$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$
- *Precision:* The proportion of true positive predictions among all positive predictions. It is calculated as:
$$Precision = \frac{TP}{TP + FP}$$
- *Recall (Sensitivity):* The proportion of actual positives that are correctly identified. It is given by:
$$Recall = \frac{TP}{TP + FN}$$
- *F1-Score:* The harmonic mean of precision and recall, balancing both metrics. It is expressed as:
$$F1 - Score = Precision + Recall$$

1.6 Future scope of sentiment analysis

The future of sentiment analysis holds high promise, mainly with the advancements in artificial intelligence and natural language processing. Evolving sentiment models will become able to better capture complex emotions and contextual nuances or even nonverbal cues; new possibilities can open across the board in diverse industries. In business and marketing, sentiment analysis is set to transform customer engagement by providing real-time insights into consumer opinions, driving personalized recommendations, and enabling more targeted advertising. In healthcare, sentiment analysis will play a crucial role in understanding patient feedback, identifying emotional distress, and enhancing mental health monitoring. The finance sector will benefit from sentiment-driven market predictions, allowing investors to gauge public sentiment and forecast stock trends or economic changes. It will reshape student-to-learning material interaction in education by providing feedback on the state of the student's emotions and adjusting content delivery to enhance learning outcomes. Integrating with social media monitoring, sentiment analysis is bound to open doors to a stronger understanding of public opinion for businesses and policymakers in decision-making. It will also break language barriers in multilingual sentiment analysis, enabling businesses to gain insights from a global audience, while emotion-aware AI will allow systems to provide more empathetic and contextually sensitive interactions.

2.0 Literature Review

Khanjarinezhadjooneghani and Tabrizi (2021), overview of social media analytics, high-lighting various methods used to analyze social media data and the insights that can be gained. Social media platforms offer a rich source of public information that is useful for understanding social behavior, opinions, and trends. Applications of social media data analysis have been demonstrated in fields like public health, crisis management, and urban development.

Haddad *et al.* (2024), The paper proposes an intelligent sentiment prediction model for social networks using big data analytics and deep learning. It leverages Hadoop and Spark frameworks for efficient preprocessing of batch and streaming data, reducing data size, access time, and storage requirements. By using GloVe forword embedding and a combination of Convolutional Neural Networks (CNNs) and Bi-directional Long Short-Term Memory (BiLSTM) models, the approach effectively classifies short texts into positive and negative sentiments. Tested on IMDB and Yelp datasets, the model achieves a high accuracy of 96 %, outperforming existing methods in sentiment analysis.

Rahman and Reza (2022), The study explores how Web 2.0 technologies enhance global connections and sharing of opinions, highlighting the potential of Big Social Data" through machine learning for better decision-making. It introduces the "Sunflower Model of Big Data," identifies the top ten social data analytics methods and emphasizes text analytics as the most common approach. The research provides a taxonomy of social media analytics to help researchers choose appropriate tools and techniques.

Zachlod *et al.* (2022), This paper reviews the key characteristics of social media data, such as its volume, variety, and velocity, and explores its applications in marketing, public relations, and political campaigns. It also discusses challenges like data quality and privacy concerns. Social media analytics helps uncover insights into user behavior, public opinion, and trends. Babu and Kanaga (2022), The document surveys sentiment analysis methods, applications, and challenges. It reviews techniques like machine learning, lexicon-based, and hybrid approaches, along with their pros and cons. The paper discusses applications in areas such as product reviews and healthcare while addressing challenges like sarcasm detection and domain-specific analysis. It also provides insights into the future of sentiment analysis research, emphasizing the need for improved tools to handle context, language variations, and evolving datasets from social media and online platforms.

Nemes and Kiss (2021), The article analyzes Twitter sentiments related to COVID-19 using Natural Language Processing and Recurrent Neural Networks (RNN). It compiles and visualizes data from comments, hashtags, and tweets to assess emotional responses. The trained RNN model demonstrates high accuracy in detecting sentiments, particularly in ambiguous tweets, using freshly scraped data on the COVID-19 theme.

Wankhade *et al.* (2022), This is a research paper on sentiment analysis. It discusses the challenges faced by businesses, governments, and individuals because of the rapid growth of internet-based applications. Sentiment analysis is the process of gathering and analyzing people's opinions, thoughts, and impressions regarding various topics, products, subjects, and services. It uses natural language processing and text mining to identify and extract subjective information from text. Some of the challenges of sentiment analysis include sarcasm, irony, and language-specific challenges.

Das and Singh (2023), The paper examines the evolution of sentiment analysis from text-based methods to multimodal approaches that include audio, images, and video. This shift enhances sentiment detection and improves system performance, especially with advanced techniques like transformer-based models. It provides an overview of various methods, applications, challenges, and resources, emphasizing the importance of multimodality in advancing sentiment analysis.

Hartmann *et al.* (2023), The paper explores sentiment analysis in marketing, comparing lexicon-based methods and machine learning techniques. It finds that newer transfer learning models generally offer better accuracy but may not always meet benchmark expectations. Contextual factors, like the number of sentiment categories and text length, are crucial for realistic performance assessments. The authors provide an open-source pre-trained model, Siebert, for easier use in sentiment analysis.

Jain *et al.* (2021), The paper presents an experimental study on forecasting COVID-19 spread patterns using various models. It found that the model worked best for RMSE with a value of 2773.27, while Triple Exponential Smoothing performed best for MAE at 1416.48. Though ARIMA effectively captures linear trends, it has limitations in accounting for factors like climate and social influences. To address this, the ARIMAX model, which allows for multivariate time-series analysis, can be used.

Abid *et al.* (2020), It propose a novel sentiment analysis method using a joint neural network, combining RNNs and CNNs to capture both long-term dependencies and local features in text. A weighted attentive pooling mechanism highlights key sentiment words. The authors also introduce a new word representation method called Dense Efficient Concatenated Representation. Experimental results show this approach outperforms existing methods, emphasizing the importance of capturing both global and local context for accurate sentiment prediction.

Serrano-Guerrero *et al.* (2015), The paper provides a comprehensive overview of sentiment analysis techniques and compares various sentiment analysis web services. It discusses different approaches to sentiment analysis, such as rule- based methods, machine learning techniques, and deep learning models. It also compares the performance of different sentiment analysis web services based on various criteria, such as accuracy, speed, and ease of use. Additionally, it discusses the potential applications of sentiment analysis in various domains. Overall, the paper is a valuable resource for researchers and practitioners interested in sentiment analysis and web services.

Sohail *et al.* (2021), This article presents a framework for crawling Twitter data while ensuring user privacy. It discusses the challenges associated with using social media data, such as rate limits, API changes, and user privacy concerns. The proposed framework addresses these challenges by implementing strategies like rate limiting, API key management, and user consent mechanisms. Overall, the framework aims to provide a reliable and ethical approach to collecting Twitter data. Madhu *et al.* (2021), The paper “Real Time Sentimental Analysis on Twitter” proposes a system for analyzing the sentiment of tweets in real-time. The system uses Flume to collect tweets and Hive for preprocessing. Sentiment analysis is performed using VADER, TF-IDF, and K-means clustering algorithms. The output indicates whether tweets are positive, negative, or neutral, providing insights into user sentiment. The authors claim that their approach, using unsupervised learning, outperforms the accuracy and speed of previous methods relying on SVM.

Fowler and Mönch (2022), The paper explore various p-batch scheduling problems, including those with different objective functions, constraints, and machine environments. It also discusses different solution methods, such as exact algorithms, heuristics, and metaheuristics. The paper highlights the relevance of p-batch scheduling in various industries, such as manufacturing, transportation, and computing. Overall, the paper provides a valuable resource for researchers and practitioners interested in p-batch scheduling.

Sawicki *et al.* (2022), The authors highlight the platform’s massive, categorized, and open-access nature, making it suitable for various research and analytical tasks. They discuss potential applications such as sentiment analysis, topic modeling, community detection, and understanding public opinion. While acknowledging challenges like data quality and bias, the paper emphasizes the opportunities that Reddit presents for data scientists and researchers seeking to gain insights from real-world data.

3.0 Proposed Methodology/Approach

The Figure 1 below shows the key stages from data preprocessing to generating insights through AI models.

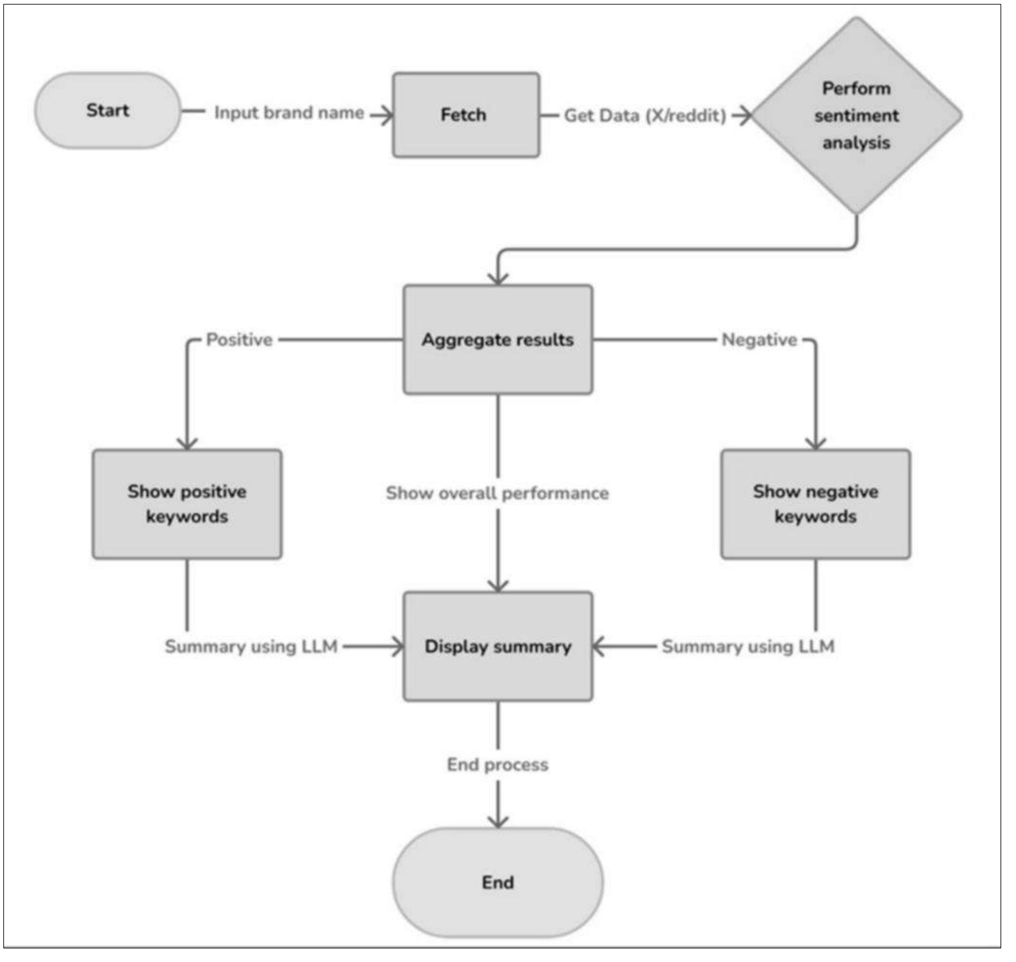
3.1 Input phase

The process starts with an easy but critical step, which is inputting a brand name or keyword by the user. This acts as a search query that the system will use to fetch relevant data across platforms. The user interface is intuitive to make the process as smooth as possible. This brand name or keyword becomes the central parameter for subsequent stages; therefore, the input phase is very critical for the whole process.

3.2 Data collection

Once the brand name is inputted, the system enters the collection phase. Data from social media networks like X, formerly Twitter, and Reddit will be fetched. These are the selected platforms, which are most rich in public opinion and discussion. X contains real-time opinions and trends. On the other hand, Reddit provides in-depth discussions and feedback from the communities. The system will use APIs or web scraping techniques in order to collect all the relevant posts, tweets, and comments.

Figure 1: Proposed System



More filters like time frame, language, and relevance make sure the data retrieved is relevant to the query. This stage also deals with issues like the removal of irrelevant or spam content and platform policies, such as API rate limits.

3.3 Sentiment analysis

This collected data is put through a stringent sentiment analysis to gauge the emotional tone and attitude of the contents. The process begins with the application of text preprocessing techniques for cleaning and normalization of the data, which removes stop words, special characters, and URLs. BERT or any other custom-built models for sentiment analysis are used for classification into positive, negative, or neutral. This step makes raw textual data more structured to be better understood and acted on. Each post or comment would be tagged as belonging to which sentiment, allowing the identification of patterns in public opinion.

3.4 Results aggregation

Following sentiment analysis, the system aggregates the gathered data into usable insights in the results aggregation stage. The streams are divided into three major channels: positive mentions and feedback, overall performance metrics, and negative mentions and feedback. Positive mentions illustrate positive comments or feedback about a brand, and negative mentions focus on criticism or complaints. Overall performance metrics summarize the distribution in terms of being positive, negative, or neutral. This kind of presentation will quickly allow stakeholders to understand the lay of the brand's sentiment.

3.5 Detailed analysis branches

For the sake of ease of reading the overall results will be divided into three branches of detailed analyses. The left branch has such positive keywords as the term "efficient"; "high quality", or "reliable", which appear frequently in the data for measurement. The center branch would have overall metrics of performance, like sentiment distribution percentages with the engagement trends over time. For example, it might say, "Over the last 30 days, 80% of mentions were positive." The negative keyword branch, among others, mentions "expensive," "delays," and "poor service". Such a structured division provides an overall view of public opinion, and it becomes easy to find out the strengths and weaknesses.

For example, performance can be usually defined by general metrics, such as percentages of sentiment distribution or trends of engagement over time. For example, it can be said that "in the last 30 days, 80% of mentions were positive." From the negative keyword branch: "expensive," "delays," and "poor service." With such structured arrangements, the public opinion profile is identifiable easily: What are the strengths and weaknesses?

3.6 Summary generation

In this stage, a high-quality language model is deployed to generate quite a natural and comprehensive summary. The LLM uses the positive and negative keywords as well as the performance measures to develop a human-like narrative. The summary is very contextually rich and insightful, the considered keywords being comprehensible. For example, generated summarization may be as follows: "Most positive feedback has been received by the brand. Users praise the brand for its reliability in terms of high-quality products. However, some feedback states that customer service is not very prompt. Overall, sentiment indicates that the brand is positively received, albeit with areas demanding strengthening." It is an LLM in the sense that the insights have more depth and understanding of context.

3.7 Completion

This is the final stage of the whole process, whereby results are presented to the user in a summarized way. The findings are displayed on engaging dashboards that have graphs, charts, and a narrative description. At this final stage, a complete analysis of the sense of the brand on the specified platforms is delivered. Furthermore, these insights provide users with knowledge and perspective about what to use in their decision-making, be it for improving customer relationship experiences, refining marketing strategies, or addressing areas of concern. By the end of this phase, the user has a holistic understanding of public sentiment related to their brand.

3.8 System applications

The system demonstrates significant utility in:

- Brand reputation monitoring
- Customer sentiment tracking
- Marketing strategy development
- Crisis management
- Understanding public perception

4.0 Conclusion

In conclusion, the automated social media sentiment analysis dashboard represents a significant advancement in how businesses can understand and respond to customer feedback. By automating the sentiment analysis process, this tool addresses the limitations of traditional manual methods, providing real-time insights and actionable data. The ability to efficiently assess public sentiment enables companies to respond proactively to feedback, refine their marketing strategies, and strengthen customer relationships. As businesses continue to navigate the complexities of the digital landscape, this project offers a practical solution for leveraging social media insights to drive success and enhance overall product performance

5.0 Acknowledgment

We would like to extend our heartfelt thanks to Prof. Sonali Lunawat, our guide, for her constant guidance, invaluable insights, and unwavering support throughout the course of our project. Her expertise, encouragement, and constructive feedback played a pivotal role in helping us navigate challenges and make meaningful progress. Our sincere gratitude also goes to Dr. Archana Chaugule, Head of the Department of Computer Engineering, for her indispensable support and for creating a conducive environment that nurtured our learning and research. Her leadership and belief in our potential have been a source of constant motivation. Additionally, we are immensely grateful to Dr. Abhijit Jadhav for providing us with various resources, both academic and practical, which were essential to the development of our project. His timely assistance and invaluable suggestions significantly enhanced the quality of our work. We deeply appreciate all their contributions, as this project would not have been possible without their guidance and support.

References

Abid, F., Li, C., Alam, M. (2020). Multi-source social media data sentiment analysis using bidirectional recurrent convolutional neural networks. *Computer Communications*, Volume 157. https://doi.org/10.1016/j.comcom.2020.04.002

Babu, N. V., Kanaga, *et al.* (2022). Sentiment analysis in social media data for depression detection using artificial intelligence: A review. *SN COMPUT. SCI. 3*, 74. https://doi.org/10.1007/s42979-021-00958-1

Das, R., Singh, T. D. (2023). Multimodal sentiment analysis: A survey of methods, trends, and challenges. *ACM Computing Surveys, Volume 55, Issue 13s* ,Article No.: 270, Pages 1 - 38. https://doi.org/10.1145/3586075

Fowler, J. W., Monch, L. (2022). A survey of scheduling with parallel batch (p-batch) processing. *European Journal of Operational Research, Volume 298*, Issue 1, Pages 1-24. https://doi.org/10.1016/j.ejor.2021.06.012

Gadia, H., Jain, A., Sukhdeve, T., Sahu, S. P., Verma, S., Gadia, H. (2021). COVID-19 prediction using time series analysis. *International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, Coimbatore, India (E-mail: 2021, pp. 1599-1606. https://doi.org/10.1109/ICAIS50930.2021.9395877

Haddad, O. F., Omri, F., Nazih, M., Fkih, O. (2024). An intelligent sentiment prediction approach in social networks based on batch and streaming big data analytics using deep learning. *Social Network Analysis and Mining.Social Network Analysis and Mining*, 14(1) https://doi.org/10.1007/s13278-024-01304-y

Hartmann, J., Heitmann, M., Siebert, C., Schamp, C. (2023). More than a feeling: Accuracy and application of sentiment analysis. *International Journal of Research in Marketing*, Volume 40, Issue1, 2023, Pages75-87. https://doi.org/10.1016/j.ijresmar.2022.05.005

Khanjarinezhadjooneghani, Z., Tabrizi, N. (2024). Social media analytics: An overview of applications and approaches.In *Proceedings of the 13th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2021)* - Volume1:KDIR,233-240. https://doi.org/10.5220/0010657600003064

Kiss, A., Nemes, L. (2021). Social media sentiment analysis based on COVID-19. *JOURNAL OF INFORMATION AND TELECOMMUNICATION*, VOL. 5, NO. 1, 11–15. https://doi.org/10.1080/24751839.2020.1790793

Madhu, K. S., Reddy, B. C., Damarukanadhan, C., Polireddy, M., Ravinder, N. (2021). Real-time sentimental analysis on Twitter. *6th International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India (E-mail: 2021, pp. 1030-1034 https://doi.org/10.1109/ICICT50816.2021.9358772

Rahman, M. S., Reza, H. (2022). A systematic review towards big data analytics in social media. *Big Data Mining and Analytics*, vol. 5, no. 3, pp. 228-244. https://doi.org/10.26599/BDMA.2022.9020009

Rao, A. C. S., Wankhade, M., Kulkarni, *et al.* (2022). A survey on sentiment analysis methods, applications, and challenges. *Artif Intell Rev* **55**, 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Sawicki, J., Ganzha, M., Paprzycki, M., Badica, A. (2023). Exploring usability of Reddit in data science and knowledge processing. *Scalable Computing: Practice and Experience, 23*(1), 9-22. https://doi.org/10.12694/scpe.v23i1.1957

Serrano-Guerrero, J., Olivas, J. A., Romero, F. P., Herrera-Viedma, E. (2015). *Sentiment analysis: A review and comparative analysis of web services. Information Sciences, Volume 311*, Pages 18-38. https://doi.org/10.1016/j.ins.2015.03.040

Sohail, S. S., Khan, M. M., Arsalan, M., Khan, A., Siddiqui, J., Hasan, S. H., Alam, M. A. (2021). Crawling Twitter data through API: A technical/legal perspective. https://doi.org/10.48550/arXiv.2105.10724

Zachlod, C., Samuel, O., Ochsner, A., Werthmuller, S. (2022). Analytics of social media data – State of characteristics and application. *Journal of Business Research, Volume 144*, Pages 1064-1076. https://doi.org/10.1016/j.jbusres.2022.02.016

CHAPTER 6

Design and analysis of Two Port Circular Patch Antenna with Rotation Characteristics for a 5G Wireless Application

Anuradha Kulkarni, Jayashree Shinde** and Pratap Shinde****

ABSTRACT

This paper presents the design and analysis of a two-port circular patch antenna with rotational characteristics for 5G wireless applications. The research provides a CPW-feed, two-port circular patch antenna element that is small and size of 50 × 140 × 1.6 mm^3. The two-port arrangement of the suggested antenna construction has varying rotational orientations, including 0° (co-planar), 90° (rotated), and 180° (inverted). The two-port configuration allows for dual polarization, enhancing channel capacity and supporting multiple-input multiple-output (MIMO) technology. The rotational characteristics of the antenna are analyzed to evaluate the beam-steering performance, ensuring improved signal coverage and reduced interference. The evaluation of the antenna simulation took into account significant parameters like efficiency, envelope correlation coefficient, isolation, and return loss. The antenna offers good isolation, with values between -23.23 dB and -32.48 dB, and radiation efficiency greater than 90%. The return loss ranges from -17.68 dB to -33.06 dB at 2 GHz. The antenna's rotated designs exhibit an envelope correlation coefficient of less than 0, indicating good performance for dual-polarization operation. The proposed antenna makes use of a 1.6 mm thick FR4 epoxy substrate and HFSS software for simulation.

Keywords: CPW; Two-port; Circular patch; Rotation characteristics; 5G communication.

1.0 Introduction

Novel antenna designs that can effectively support high data rates, minimize interference, and guarantee dependable communication are now required due to the changing requirements of 5G wireless networks.

**Corresponding author; Research Scholar, Department of Electronics and Communication, Sinhgad Academy of Engineering, Pune, Maharashtra, India*
(E-mail: anuradhak.kulkarni@gmail.com)
***Professor, Department of Electronics and Communication, Sinhgad Academy of Engineering, Pune, Maharashtra, India (E-mail: jpshinde.sae@sinhgad.edu)*
****Professor, Department of Electronics and Communication, TSSM's Bhivarbai Sawant College of Engineering, Pune, Maharashtra, India (E-mail: pnshinde2@gmail.com)*

Circular patch antennas are an attractive choice for these applications due to their small form factor and ease of integration into compact wireless devices. However, traditional designs are often limited by their single-polarization capabilities, which can restrict the potential for improving channel capacity and system performance. In order to overcome these constraints, dual-polarized antennas have drawn a lot of interest, especially in multiple-input multiple-output (MIMO) systems that need to be able to send and receive several data streams at once. In this work, we present a two-port circular patch antenna that utilizes rotational characteristics to achieve dual-polarization, thereby enhancing the performance of 5G communication systems.

2.0 Literature Review

The development of advanced antenna designs is crucial for addressing the demands of 5G wireless communication systems. Among various antenna types, circular patch antennas have gained popularity due to their compactness, ease of fabrication, and ability to support dual-polarization and MIMO (Multiple-Input Multiple-Output) systems, which are essential for enhancing the capacity and performance of 5G networks. This literature review examines studies on circular patch antennas, with a focus on their performance in 5G applications, the effects of substrate materials, rotational characteristics, and MIMO functionality.

Ariffin *et al.* (2020) confirmed that FR4 epoxy is suitable for compact antenna designs when properly optimized for specific applications. Radiation efficiency and return loss are critical metrics for determining the effectiveness of an antenna in real-world scenarios. Circular patch antennas have been extensively studied due to their favorable characteristics, such as omnidirectional radiation patterns and simple feeding mechanisms.

Balanis *et al.* (2005) highlighted the advantages of patch antennas, including their small size and ease of integration into compact systems. Chen *et al.* (2017) emphasized the importance of evaluating the envelope correlation coefficient (ECC) in dual-polarized antenna designs, as it affects the antenna’s ability to maintain low cross-correlation between the two ports. The design of circular patch antennas with FR4 substrates for wireless communications is also discussed, focusing on the trade-offs between substrate material, antenna performance, and operational frequency. This study is relevant to 5G systems, where compact and efficient antenna designs are essential for meeting high data rate requirements.

Ghalib *et al.* (2017) investigated the performance of circular patch antenna arrays for MIMO applications, which are key to 5G communication systems. The paper examines how the array configuration enhances diversity and spatial multiplexing, improving the overall system capacity. The study emphasizes the importance of efficient design in maximizing the benefits of MIMO. A promising approach, as discussed by Hussain *et al.* (2021), involves utilizing metamaterials to enhance antenna performance and reduce mutual coupling. For example, Kim *et al.* (2011) designed a circular patch antenna with enhanced bandwidth for wireless communications, which is essential for high frequency 5G operations. Kundu *et al.* (2017) investigated the impact of various substrate types on 5G circular patch antenna performance, highlighting the trade-offs between material cost and performance.

Lee *et al.* (2019) investigated the impact of antenna orientation on beam steering in 5G base station applications, showing that antenna rotation can help adapt to varying user locations and improve coverage. Additionally, Lee *et al.* (2017) studied return loss and radiation patterns for dual-polarized patch antennas, highlighting their importance in the performance of MIMO systems. While circular patch antennas have many advantages, there are still challenges in optimizing their performance for 5G networks. Li *et al.* (2016) presented a study on the rotational characteristics of circular patch antennas, analyzing the effect of different rotational orientations on beamforming and beam steering.

Additionally, programmable antennas that can adjust to shifting communication needs are becoming more and more popular. Sadeghzadeh *et al.* (2020) explored the use of reconfigurable circular patch antennas for 5G, enabling beamforming and polarization switching to optimize signal reception. Wu *et al.* (2015) demonstrated a dual-polarized circular patch antenna for LTE applications, noting its potential for enhancing communication throughput. Xie *et al.* (2019) used HFSS to simulate a circular patch antenna with a dual-polarized feed, demonstrating its high efficiency and low return loss, which is crucial for 5G communication. In a similar vein, Xu *et al.* (2018) proposed a dual-port circular patch antenna with high isolation, achieving values greater than 30 dB, which is ideal for 5G applications.

FR4 epoxy is frequently used in antenna designs because of its inexpensive cost and simplicity of production. The antenna's performance, particularly at higher frequencies, may be impacted by its dielectric characteristics. Yang *et al.* (2015) designed a circular patch antenna with improved radiation efficiency, demonstrating its suitability for wireless communication applications. A dual-port antenna's ports must be isolated from one another in order to prevent signal interference and preserve the communication system's integrity. A study by Zhang and Yang (2014) focused on enhancing isolation in dual-port antennas by introducing novel feeding techniques. Similarly, Zhang *et al.* (2018) explored dual-polarized patch antennas for 5G MIMO systems, showing that dual polarization helps reduce signal interference and improves link reliability.

5G is based on MIMO technology, which uses multiple antennas at the transmitter and receiver to boost wireless communication systems' capacity. Circular patch antennas are commonly used in MIMO systems due to their suitability for compact, high-performance configurations. Furthermore, Zhang *et al.* (2020) highlighted the effectiveness of MIMO antenna arrays in 5G networks, where multiple polarizations and antenna configurations are utilized to increase channel capacity and reduce fading effects. Beam steering is an essential feature for optimizing signal coverage and minimizing interference in wireless communication systems. The rotational characteristics of antennas, particularly in dual-polarized designs, can enhance beam steering performance.

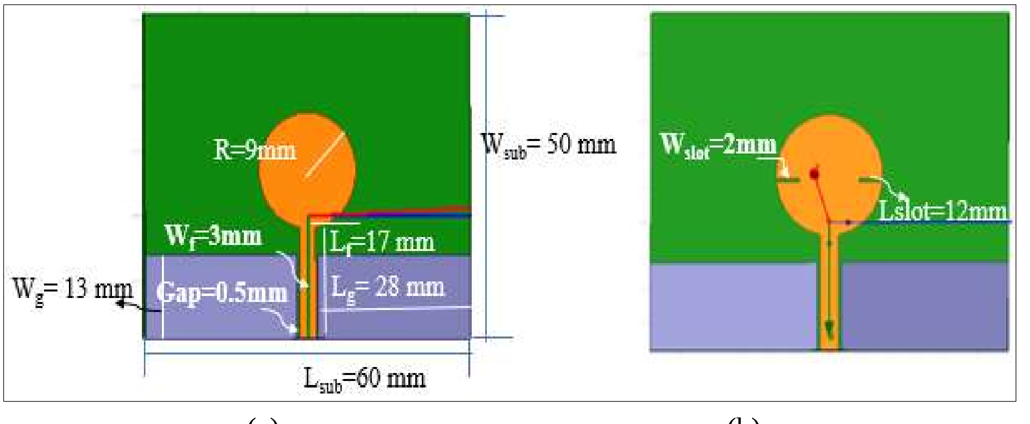
CPW feed two port circular patch antennas with coplanar, rotated, and inverted rotation characteristics are analyzed in this article. The following is how the remaining sections are organized: The literature that has already been written about the subject is reviewed in Section 2. Subsection 3.1 covers the design of the fundamental 5G antenna, whereas Section 3 concentrates on the antenna arrangement. 3.2 presents the simulation results of this basic 5G antenna. In subsection 3.3, the simulated results for the two-port circular configuration with varying rotational characteristics are shown. Section 4 compares the performance of different rotational configurations of the circular patch antenna. Finally, Section 5 concludes the paper.

3.0 Antenna Configuration

3.1 Single antenna

The ground plane and antenna are positioned on the same side of the substrate in the CPW feeding approach, which simplifies manufacture. The antenna is mounted on an FR4 substrate with a thickness of 1.6 mm, a dielectric constant of 4.4, and a loss tangent of 0.02. The CPW feed line uses a center conductor with a width of 3 mm and a gap of 0.5 mm.

Figure 1: Antenna Design a) Without Slot b) with Slot



(a) (b)

Figure 1 shows antenna design with and without slot. Two slots in the 12x2 mm2 radiating patch contribute to the suggested antenna’s increased bandwidth. Different parameters were observed.

The following is a description of the patch antenna's design process:
The resonant frequency fr = 2.1 GHz is the one at which the antenna is intended to function. Equation (1) is used to get the circular patch antenna's radius "a"

$$a = \frac{F}{\left\{1+\frac{2h}{\pi\varepsilon_r F}\left[ln\left(\frac{\pi F}{2h}\right)+1.7726\right]\right\}^{1/2}} \quad \ldots(1)$$

With

$$F = \frac{8.791\times10^9}{2f_r\sqrt{\varepsilon_r}} \quad \ldots(2)$$

Where h = 0.16 cm (1.6 mm) is the substrate's height/thickness, fr = 2.1 GHz is the antenna's resonance frequency, and εr = 4.4 is the FR4 epoxy substrate's dielectric constant. According to equations (1) and (2), the circular patch's theoretical radius "a" is 9.97 mm. A microstrip feed line was selected since it is simple to construct and model. The feed line's length roughly corresponds to the circular patch's radius. The antenna impedance determines the feed-width. It is assumed that the antenna impedance is 50 **Ω**. This is due to the fact that the majority of RF equipment have characteristic impedances of 50 **Ω** or 75 **Ω**. In order to prevent mismatch losses, the antenna must be precisely matched when these devices are linked to it. The following equations from [2] can be used to determine the feed-width:

For W/h ≤ 1:

$$\varepsilon_{re} = \frac{\varepsilon_r+1}{2}+\frac{\varepsilon_r-1}{2}\left\{\left(1+12\frac{h}{W_f}\right)^{-0.5}+0.04\left(1-\frac{W_f}{h}\right)^2\right\}. \quad \ldots(3)$$

$$Z_c = \frac{\eta}{2\pi\sqrt{\varepsilon_{re}}}\ln\left(\frac{8h}{W_f}+0.25\frac{W_f}{h}\right). \quad \ldots(4)$$

Where ɳ = 120π **Ω** is the wave impedance in free space
For W/h ≥ 1:

$$\varepsilon_{re} = \frac{\varepsilon_r+1}{2}+\frac{\varepsilon_r-1}{2}\left(1+12\frac{h}{W_f}\right)^{-0.5} \quad \ldots(5)$$

$$Z_c = \frac{\eta}{\sqrt{\varepsilon_{re}}}\left\{\frac{W_f}{h}+1.393+0.677\ln\left(\frac{W_f}{h}+1.444\right)\right\}^{-1} \quad \ldots(6)$$

W/h is often larger than 1 in most situations. Therefore, equations 5 and 6 are employed. Theoretically, Wf (feed width) = 3.2mm with Zc = 50 **Ω**.

Lg = 2a + 6h … (7)
Wsub = 2a + Lf + 6h + 20 … (8)
Lsub = a + 2Lg + 6h … (9)

Table 1: Dimensions of the Antenna (Optimal and Theoretical)

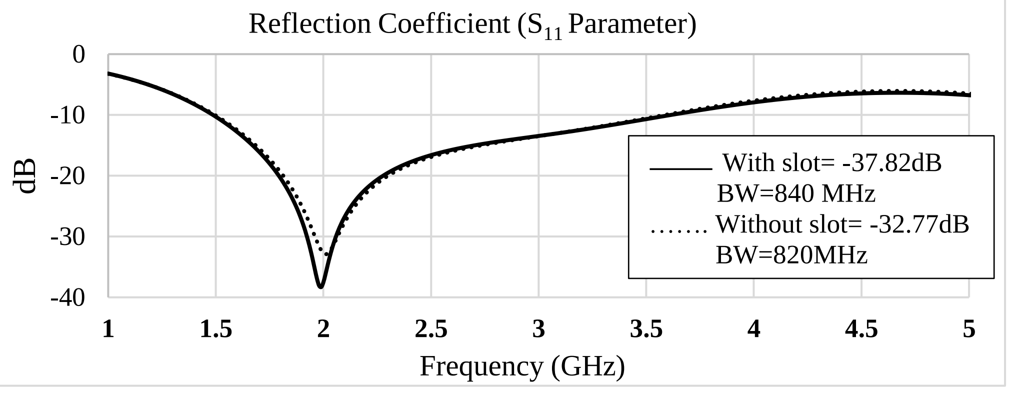
Antenna Parameter	Antenna Parameter Value	
	Theoretical	**Optimized**
Radius 'R'	9.97 mm	9.00 mm
Feed line length (L_f)	9.97 mm	17.0 mm
Feed line width (W_f)	3.20 mm	3.00 mm
Length of ground (L_g)	29.54 mm	28.00 mm
Substrate length (L_{sub})	59.45 mm	60.00 mm
Substrate width (W_{sub})	53.5 mm	50.00mm
Substrate thickness (h)	1.6 mm	1.6 mm
Slot length (L_{slot})	-	12.00 mm
Slot width (W_{slot})	-	2.00 mm

3.2 Circular patch antenna simulation results

Reflection Coefficient (S_{11} Parameter): In the transmission medium, the reflection coefficient is a measure of the quantity of electromagnetic wave reflected by an impedance discontinuity. It is equivalent to the amplitude ratio of the incident wave to the reflected wave. The S11 parameter is an alternative name for the reflection coefficient. The ideal reflection coefficient is higher than -10 dB.

With and Without Slot

Figure 2: An Antenna's Simulated Reflection Coefficient With and Without a Slot



The antenna design's reflection coefficient with and without a slot shown in Figure 2. The reflection coefficient is -37.82 dB with the slot, offering a bandwidth of 840 MHz, and -32.77 dB without the slot, with a bandwidth of 820 MHz, It has been shown that increasing the radiating patch's slot enhances its bandwidth and resonance at the operational frequency.

Different Values of Radius R

Figure 3: Simulation of an Antenna's Reflection Coefficient for a Range of Radius R Values

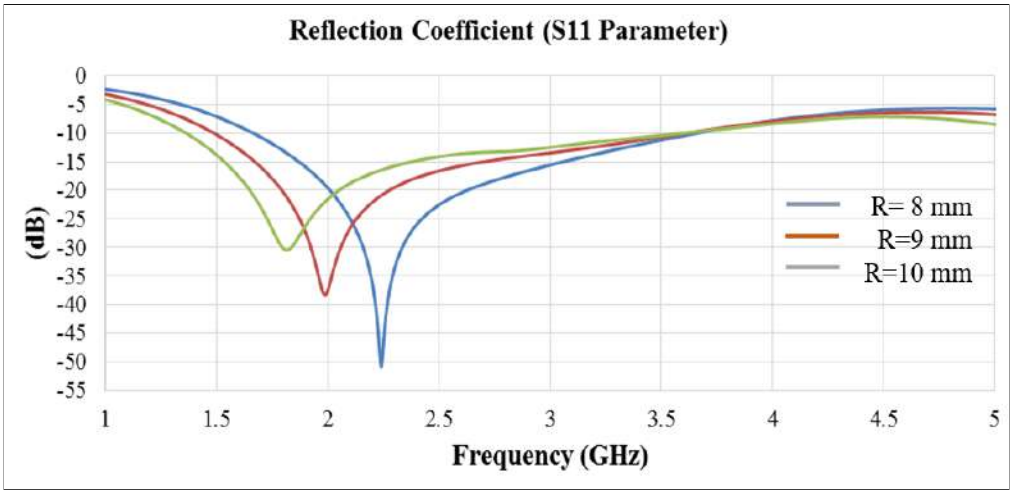


Figure 3 displays an antenna's reflection coefficient over a range of radius R values. The reflection coefficients of the antenna are -19.71 dB, -37.51 dB, and -21.55 dB for radii of 8 mm, 9 mm, and 10 mm, respectively. The figure shows that the frequency increases as the radius decreases.

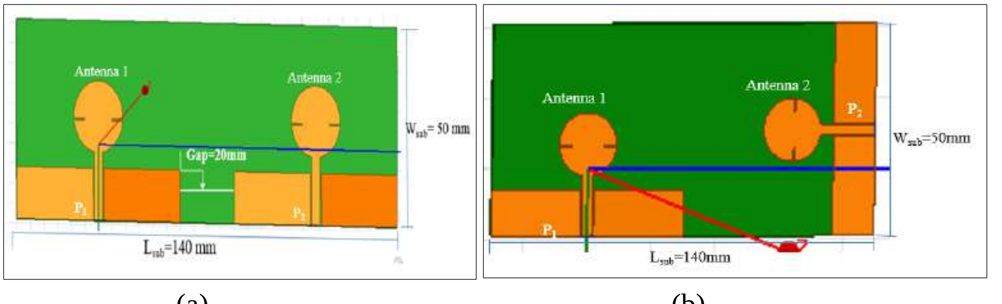
3.3 Two-port antenna with varying rotation characteristics

Two basic antennas are presented in different configurations. Three MIMO antenna configurations, rotated at 0°, 90°, and 180°, are designed.

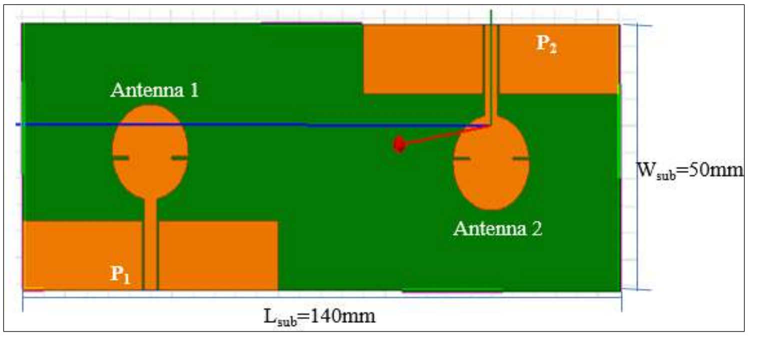
Radiation Efficiency: Radiation efficiency is an essential parameter that indicates how effectively an antenna transmits and receives RF signals.

Figure 4: Two-port Circular Configuration

(a) 0^0 (Co-planar) (b) 90^0 (Rotated) (c) 180^0 (Inverted)

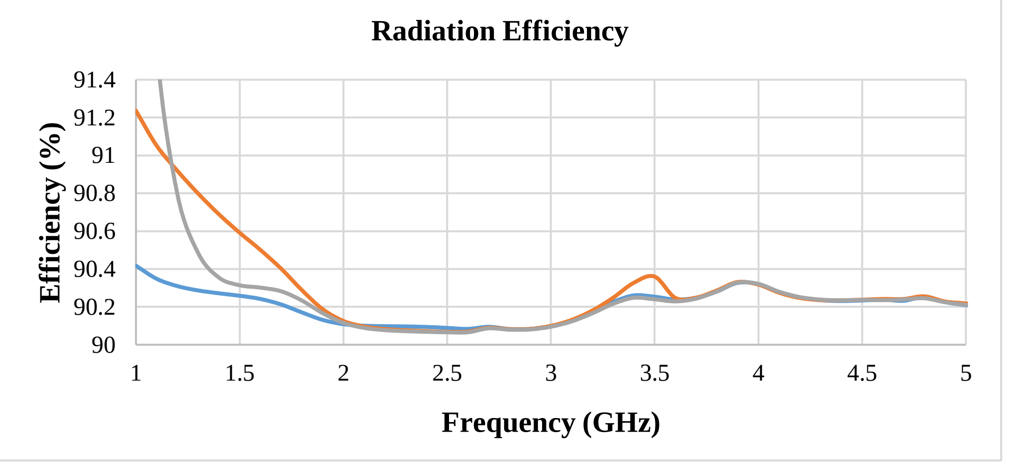


(a) (b)



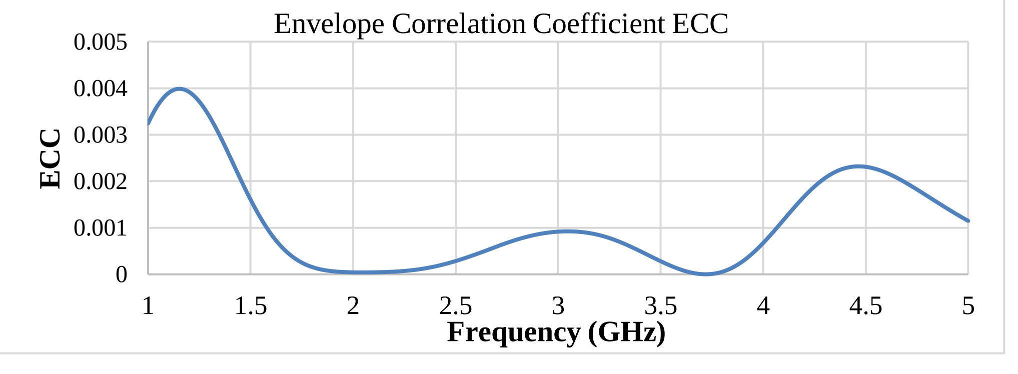
(c)

Figure 5: Radiation Efficiency Simulation for Various Rotation Angles



Simulated radiation efficiency for various rotation angles is shown in the figure 5. Across all antenna rotation angles, the antenna provides an overall radiation efficiency of almost 90%.

Figure 6: ECC of a Two-port MIMO Antenna Simulation



Envelope Correlation Coefficient (ECC): An essential metric for assessing a MIMO communication system's performance is the envelope correlation coefficient.

Figure 6 shows simulated ECC of 2 port MIMO antenna. Features with an Envelope Correlation Coefficient of less than 0 are provided by the antenna throughout rotational configuration

Simulation Results and Analysis: Table 2 provides a comparative examination of different two port circular antenna layouts in this section.

Table 2: Comparison of the Properties of Rotation

Antenna Parameter	**Antenna Rotation Angle**		
	0^0 (Co-planar)	**90^0 (Rotated)**	**180^0 (Inverted)**
Antenna size	50×140mm^2	50×140mm^2	50×140mm^2
Return loss (S_{11}, S_{22})	-28.71 and -32.07 dB at 2 GHz frequency	-33.06 and -23.75 dB at 2 GHz frequency	-27.12 and -17.68 dB at 2 GHz frequency
Isolation (S_{12}, S_{21})	-23.23 dB at 2 GHz frequency	-32.48 dB at 2 GHz frequency	-24.07 dB at 2GHz frequency
Efficiency	> 90%	>90%	> 90%
Antenna structure	As shown in Fig. 4 a)	As shown in Fig. 4 b)	As shown in Fig. 4 c)

4.0 Conclusions and Future Work

Using a two-port circular patch antenna structure, the design of a CPW-feed antenna element is examined in this work. It provides good isolation, ranging from -17.68 dB to -33.06 dB, in different rotation properties, such as coplanar, rotated, and inverted configurations, having an envelope correlation coefficient (ECC) of less than 0. The suggested antenna layout provides -32.48 dB isolation, -33.06 dB return loss values, and -23.75 dB at 2 GHz for the rotational configuration. and an overall rotational characteristics efficiency of over 90%, making it appropriate for wireless applications using 5G. Additionally, an Electronic Bandgap (EBG) structure is used to minimize the antenna size while preserving a mutual coupling trade-off that is acceptable.

References

Ariffin, A., Fadhlullah, M., & Ismail, M. (2020). Performance comparison of circular patch antennas using FR4 substrate for 5G communication. *International Journal of RF and Microwave Computer-Aided Engineering, 30*(2), e22304.

Balanis, C. A. (2005). *Antenna Theory: Analysis and Design* (3rd ed.). Wiley.

Chen, X., Li, M., & Li, W. (2017). Evaluation of envelope correlation coefficient in dual-polarized antenna design. *IEEE Transactions on Antennas and Propagation, 65*(6), 2938-2945.

Chien, S., Yang, S., & Yang, J. (2016). Design of circular patch antennas with FR4 substrates for wireless communication. *Microwave and Optical Technology Letters, 58*(4), 970-975.

Ghalib, A., Zubair, M., & Hussain, M. (2017). Circular patch antenna array for MIMO applications. *Journal of Electromagnetic Waves and Applications, 31*(15), 1571-1583.

Hussain, M., Iqbal, S., & Rehman, A. (2021). Metamaterials for enhancing the performance of patch antennas in 5G networks. *IEEE Access*, *9*, 3452-3458.

Kim, H. W., Kim, S. H., & Lee, J. Y. (2011). Wideband circular patch antenna for wireless communications. *IEEE Transactions on Antennas and Propagation, 59*(5), 1514-1519.

Kundu, D., Bhattacharyya, R., & Datta, S. (2017). Substrate material effects on circular patch antenna performance. *IEEE Transactions on Antennas and Propagation, 65*(7), 3834-3841.

Lee, J. H., Kim, Y. W., & Lee, H. K. (2019). Impact of antenna rotation on beam steering in 5G systems. *IEEE Antennas and Wireless Propagation Letters, 18*(1), 172-175.

Lee, K., Kim, H., & Lee, S. (2017). Evaluation of return loss and radiation patterns in dual-polarized patch antennas. *IEEE Antennas and Wireless Propagation Letters, 16*, 1529-1532.

Li, Y., Zhou, Y., & Tang, F. (2016). Rotational characteristics of circular patch antennas for beamforming applications. *IEEE Transactions on Antennas and Propagation, 64*(9), 3723-3731.

Sadeghzadeh, M., Saeed, A., & Rezaei, M. (2020). Reconfigurable circular patch antennas for 5G wireless communication. *IEEE Antennas and Wireless Propagation Letters, 19*(6), 1327-1330.

Wu, J., Zhao, X., & Li, Y. (2015). Dual-polarized circular patch antenna for LTE applications. *IEEE Antennas and Wireless Propagation Letters, 14*, 1340-1343.

Xie, Z., Zhang, Q., & Zhang, L. (2019). Simulation of dual-polarized circular patch antenna using HFSS. *IEEE Antennas and Wireless Propagation Letters, 18*(2), 145-149.

Xu, X., Chen, X., & Lin, H. (2018). Dual-port circular patch antenna with high isolation for 5G MIMO systems. *IEEE Antennas and Wireless Propagation Letters, 17*(3), 436-439.

Yang, F., Wei, D., & Liu, L. (2015). Circular patch antenna with improved radiation efficiency. *IEEE Transactions on Antennas and Propagation, 63*(4), 1410-1415.

Zhang, L., & Yang, J. (2014). Isolation enhancement in dual-port patch antennas using novel feeding techniques. *IEEE Transactions on Antennas and Propagation, 62*(10), 5032-5038.

Zhang, Y., Liu, X., & Xu, Q. (2018). Design of dual-polarized patch antennas for 5G MIMO systems. *IEEE Transactions on Antennas and Propagation, 66*(4), 1829-1834.

Zhang, Z., Tan, K. Y., & Lee, K. F. (2020). MIMO antenna arrays for 5G networks: An overview. *IEEE Access*, *8*, 212459-212478.

CHAPTER 7

Machine Learning-powered Hairstyle Suggestions for Enhanced Self-perception

Harshad Bankar, Pawan Wawage**,*
*Vaibhav Ekshinge*** and Yogesh Deshpande*****

ABSTRACT

Everyone wants to look attractive, beautiful or handsome in the current social media age, but choosing a haircut that flatters a person's face can be difficult. People might choose a haircut that enhances their appearance as well as their confidence in it. This research study proposes a machine learning strategy that helps individuals recommend haircuts according to their facial structure and increase their confidence. Using computer vision techniques, the system evaluates the user's face and then recommends hairstyles based on machine learning models. It also allows users to give feedback on their hairstyle choices so that the system can improve its suggestions based on individual preferences. This makes the recommendations more personalized and accurate over time. This research also looks at how personalized hairstyle suggestions can improve a person's confidence. By giving users smart, AI-powered hairstyle options, the system helps them make better choices that fit their style and image. Our initial results show that the system can successfully recommend hairstyles that users are happy with and feeling more confident after trying the suggested styles. This research could change the way people choose hairstyles by using artificial intelligence to offer personalized, confidence-boosting recommendations in a simple and user-friendly way.

Keywords: Hairstyle; Facial features; Machine learning; Artificial intelligence; Hairstyle recommendation.

1.0 Introduction

Social impressions and self-confidence are greatly influenced by one's outward look.

**Corresponding author; UG Student, Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, Maharashtra, India*
(E-mail: harshad.22111291@viit.ac.in)
***Faculty, Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, Maharashtra, India (E-mail: pawan.wawage@viit.ac.in)*
****UG Student, Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, Maharashtra, India (E-mail: vaibhav.22110821@viit.ac.in)*
*****Faculty, Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, Maharashtra, India (E-mail: yogesh.deshpande@viit.ac.in)*

Among the many components of appearance, hairstyle is particularly important and has a significant impact on how someone looks. But choosing the perfect hairstyle isn't always easy because there are a lot of variables to consider, including lifestyle choices, hair textures, face shapes, and even current fashion trends. However, recently surging attention has been applied to the employment of artificial intelligence (AI) and machine learning (ML) algorithms for personalized recommendations in diverse areas, including beauty and fashion, as a result of the aforementioned technological development. When you plan on getting a new haircut, you can opt for a hair recommender to help you make the right judgement which will save time and confusion of what style would look good.

To the rescue of this need comes a system developed with and AI in our project that provides personalized hairstyle suggestions for users, where input factors include preference choices as well as facial traits. Machine learning algorithm system is able to use machine learning algorithms and image processing techniques to examine user images for face characteristics. These are then combined with hairstyles from a curated pool. This ensures that the proposed styles are suitable for the user facial structure and follow modern fashion trends.

Previous the system was limited to suggestion in a consultation with one of our hair stylists today this is designed so that it bridges the gap between expert invalidation and non-hair stylist users. No longer will users tentatively try out different looks in private, now armed with the confidence that they won't be the only ones wearing those Harry Potter-style glasses to a social event, while personalized fashion advice becomes increasingly democratized. In this paper we explore how the recommender system is delivered; through which methodology and technology stack, delivering accurate + user-friendly hairstyle recommendations or not that does not be covered (at least very much). Also, we consider the various application areas of this system and its potential impact on beauty and fashion sectors.

2.0 Literature Review

Liu *et al.* (2019) conducted research on design and execution of a facial recognition-based hair recommendation system. It explains the architecture of the system, which is split into two parts: a server-side application for face recognition training and hairstyle-facial matching parameter optimization, and a client-side Android application for recommendation gathering and presentation. After extracting and analyzing facial data using face recognition technology, the system employs a custom recommendation algorithm to suggest appropriate haircuts for users. To choose the best hairdo, the recommendation system analyzes the user's facial feature points with the database's contents. To help the user and the stylist communicate effectively, the device also offers a virtual try-out of the suggested haircut.

The paper addresses the use of VGG-Very-Deep-16 (a face recognition network developed by the Visual Geometry Group) for facial feature extraction and database comparison to suggest appropriate hairstyles, as well as Convolutional Neural Networks (CNN) for face recognition. It also outlines the process of creation and implementation, including CNN training, the Android interface, and the assessment of the recommendation algorithm through color value quantification and visual evaluation techniques. The article also recognizes the project's accomplishments while outlining potential areas for improvement and additional study, such as growing the hairstyle data collection and considering other elements like hair volume and quality in the suggestion algorithm.

According to the research of Chen Y *et al.*, (2021), dissects previous work on core methods, open datasets, and hairstyle recommendation. It analyzes what has gone wrong in current datasets — such as the limited size and gender imbalance, absence of facial attribute data that lead to unreliable recommendation systems. Describe many methods of which are machine learning algorithms, probabilistic graph models, and end-to-end frameworks. Beauty e-Expert Methods by Hansini & Dushyanthi. The paper also gives a comparison of different face shape categorization methods, including the accuracy and training set sizes of these methods. It also exposes the drawbacks of the facial landmark, and hairstyling–attributed datasets, such as CelebA, Beauty e-Expert, Hairstyle30k vs. a more comprehensive category: Celebrity Hairdressing dataset — CelebHair. The bottom line from this survey over the literature is that it is very hard to develop a reliable hairstyling suggestion system if the system has not been trained on a fresh, big dataset with all related features.

Leigtens *et al.*, (2020) conducted research on "Hairstyle Recommender System Using Machine Learning." It describes the procedure for gathering data, standardizing data, choosing an algorithm, learning how to use it, and the system's viability, actuation, and feedback loop. The system makes gender- and face-feature-based hairstyle recommendations using a Feed-forward neural network (FNN). It also looks at the system's social components and connectivity, covering both the expansion of the system to include social components and the fundamental requirements for connectivity. An extensive analysis of the sociological and technical factors that go into developing a successful hairstyle recommender system is given in this document.

Pawan *et al.*, (2024) conducted research on "Style-Sage," a hairstyle recommendation system that uses cutting-edge machine learning techniques to offer customized hairdo choices depending on the user's facial form. The method uses a dataset of 25,434 examples with different haircuts to classify facial shapes into five categories. It describes the methodology, which includes data collecting, backend programming with Flask and Python, and machine learning model training for facial shape identification and hairstyle correlation, and highlights the system's non-invasive approach, cost-effectiveness, and user-friendly interface. In the outcome analysis, several machine learning models are evaluated; on the test dataset, the Multi-layer Perceptron (MLP) model has the greatest accuracy rate of 68.35%. The Style-Sage is very important in the personal grooming because it replaces subjective methods by statistical recommendations. Duan *et al.*, (2022) propose a new way of doing face identification from 3D pictures by classifying face shapes as rounded or elongated — a result we expect from "attention network-based bilinear model for rough human interest" where a better scheme will be to recommend glasses or hairstyles for. The approach combines a highly efficient (in terms of both high-dimensional and low-dimensional data extraction) net bilinear network with an attention mechanism to yield overall higher classification accuracy. Experimental evaluations, in this section, we provide empirical experiments showing the effectiveness of the proposed algorithm over other methods under different settings with a comprehensive report on dataset information, evaluation metrics, performances measurements, and experimental setup. This work presents an end-to-end shape-aware paradigm using state-of-the-art methodologies like attention mechanisms, bilinear networks, and Efficient-Net to enhance the classification traits over faces. We show empirical results in some relevant tasks that inevitably benefit from instance segmentation of face shape with the algorithm.

Jang *et al.*, (2022) conducted research on the shortcomings of current hairstyle display techniques and suggests a deep learning-based network that can identify faces and hair types and perform hairstyle conversion. The model combines GAN with Mask R-CNN, using Adaptive Instance Normalization (AdaIN) to convert styles. It also describes how a fresh dataset was created, how the Pyramidal Attention Network (PAN) was used to extract hair and facial features, and how well the system synthesizes faces and hairstyles. The paper also covers other GAN models, such as Cycle-GAN and Star-GAN, and presents the GAN module architecture for hairstyle conversion. It highlights the contributions made by the study, such as the generation of datasets, the use of PAN for feature extraction, and the suggestion of the Hairstyle Conversion Network (HSC-Net), which combines Mask R-CNN, PAN, and GAN. Research highlights the relevance of picture categorization, segmentation, and generation as well as prospective applications in the beauty business. It also highlights the value of high-quality image collections and makes recommendations for future study areas in face shape and feature-based hairstyle selection algorithms.

Alzahrani *et al.*, (2021) conducted research on the building of decision support tool, which basically links face and eye analysis with artificial eyelashes recommendations. In the research project, a multi-model system was generated from three different models that aim at gender detection, eye attribute identification, and face shape classification. The system categorizes face shapes with a hybrid of handcrafted and learned features, lists eye measurements based on landmark detection to identify eye attributes, and uses deep learning for male/female recognition. The results of these models are then stacked together to form a decision support system that recommends haircuts and eyelash extensions. The results of the study show excellent generalization capability of the system to external data and good performance in different conditions (e.g., camera position, obstacles (eg glasses use), light circumstances, gender, age).

Although the system demonstrated tolerance for misclassified predictions, it also demonstrated some limitations; it did not have formal testing on diverse ethnic and racial groups, and there was no examination as to the characteristics of ocular situations. The authors suggested that future work should study the computational complexity of such a method for optimal convergence, and implement the framework developed into a cloud-based decision assistance system with a mobile application.

Sunhem *et al.*, (2016) conducted research on Hairstyle Suggestion System for Women describes the creation of a hairstyling advice system made particularly for women. Now it utilizes a categorization system for face shapes and professional hairstyling experience. Through the method, women can select the best hairstyles based on their features so that will make them feel more confident. The method includes feature extraction, face shape detection, and custom cut for different faces. There are six major sections in the system, which include starting, tutorial, photo capturing, review, report, and simulation.

Results from 20 voluntary women who evaluated the system had an average score of 4.05 out of 5.00 and showed good feedback for both performance and user interaction with automatic mode as well manual modes. Hairdo Recommendation System aims to guide women to hairstyles that suit them in a way of not only looking good but feeling better about themselves, by utilizing feature extraction, face shape classification as well as custom hairdo generation. By its six-part structure it makes easy use as well as an in-depth examination. These comments suggest that perhaps the system could empower women to make more informed decisions about their hairstyles and, in so doing, build self-confidence.

Sunhem *et al.*, (2016) conducted research on An Approach to Face Shape Classification for Hairstyle Recommendation. First, using a face segmentation method and Active Appearance Model (AAM), this research helps individuals choose the most appropriate hairstyle based on their facial features to enhance personal confidence and beauty. They ran different machine learning methods on a set of 1,000 images, and the Support Vector Machine (SVM) performed best with an accuracy rate of 72% in predicting face shape. The study provides insight into the performance of classification algorithms, and benchmarking significant features offers a methodical solution to a real-world problem in the beauty business.

Weerasinghe *et al.*, (2020) conducted research on “Machine Learning Approach for Hairstyle Recommendation”. It emphasizes how hair plays a significant role in not only the psychological effects but also in drawing attention to or away from certain facial features. The model utilizes this to categorize face forms, estimate hair length, and suggest suitable haircuts. The salon network platform allows users to share suggestions. The model reached 91% accuracy in face shape classification and an equally impressive 83% for hair recommendation after training with 5,000 photos. The process includes data collection, face landmark detection, Naïve Bayes-based recommendations, and face detection using Haar Cascade classifiers. This system’s overall performance was evaluated, incorporating feedback from beauty experts to ensure accurate, personalized recommendations. Weerasinghe *et al.* (2020), in their research paper titled Machine Learning Approach for Hairstyle Recommendation, introduced annotated hairstyle images that were labeled by hair stylists. Al outlines a new architecture for predicting hairstyles using face style classification. This emphasizes the significance of hair in highlighting facial features and how this affects us psychologically. It uses machine learning to categorize face shapes, estimate hair length, and suggest hairstyles. Recommendations could be shared amongst users via the "The Beauty Quest" salon network platform. Thanks to 5,000 trained images, the system is capable of reaching Face shape classification 91% accuracy and Hair recommendation 83% accuracy. It includes data collection, face detection with the Haar Cascade classifier, face landmark detection, and recommendations via Naïve Bayes. The evaluation included feedback from beauty experts and overall system performance to ensure a more personalized and accurate recommendation. This solves a common problem of customers being unhappy with their haircuts, which could be a game-changing event for the beauty sector.

Table 1: Comparative Analysis of Literature Survey

Authors	Purpose	Model Used	Accuracy	Impact	Drawback
Chen *et al.*, (2021)	Hairstyle recommendation	Random Forests Algorithm	87.03%	Enhances personal appearance based on face shapes and attributes	Limited to hairstyles that fit predefined categories, may not account for individual preferences.
Alzahrani *et al.*, (2021)	Face shape and eye attributes identification	Inception V3, CNN	85.6%	Identifying face shape and eye attributes with 85.6% accuracy using Inception V3 CNN for beauty, healthcare, and hiring.	High computational cost and dependency on large labeled datasets for training.
Duan1 *et al.*, (2022)	Face shape classification	Attention Mechanism and BiLSTM (Bidirectional Long Short-Term Memory) integrated with CNN (AB CNN)	89.8%	Improved facial recognition	Complexity of the model may lead to overfitting and requires careful tuning.
Jang *et al.*, (2022)	Hairstyle transformation using GAN network	HSC-Net (proposed model)	93%	Growing interest in social media posts about hairstyles, and their effects on confidence and self-expression	GANs can be unstable during training, leading to inconsistent results.

Vittal *et al.*, (2022)	Face shape identification for fashion and e-commerce	CNN	82%	By helping choose outfits, hairstyles, and cosmetics, a facial shape-based recommendation system boosts confidence and can be used in e-commerce to provide individualized styling.	Limited generalizability to different face shapes not represented in training data.
Weerasinghe & Vidanagama (2020)	Face Shape Identification	Haar Cascade Classifier, dlib library	91%	Improved self-esteem and confidence due to personalized recommendations	Relies on pre-trained models which may not perform well in diverse lighting or angles.
Weerasinghe, & Vidanagama, (2020)	Haircut/ Hairstyle Recommendation	Naïve Bayes Classification Algorithm	83%	Positive impact on personal appearance and psychological well-being	Assumes independence among features, which may not hold true in complex visual data.

3.0 Methodology

We performed a multistage implementation of the Hair Style Recommender by creating one training session for each neural network, facial feature extraction prototype, and recommendation system. In the data collection part, we will use face shape with MLP model training for recommendation generation.

3.1 Data collection

At the first step, we need to get a dataset which containing labeled face images and their corresponding facial shapes (e.g., Round, Oval, Square, etc.) along with a collection of hairstyles available. Each hairstyle is tagged with face shape metadata and additional attributes like hair length, texture, and style type.

More formally, let $D = (Ii, yi)$where Iirepresents the input image of shape label yi. We denote the hairstyle dataset by $H = hj$, where hj is a pairing of suitable face shape information and images with hairstyles.

3.2 Feature extraction

Before feeding the images into the MLP model, facial feature extraction is performed. We use a deep learning model, such as a pre-trained Neural Network, to extract high-dimensional feature vectors from the user's face. The extracted features represent key aspects of the face, such as jawline shape, eye position, and overall geometry. Given an image $Ii$the feature extraction model outputs a vector $xi \in R^d x$, where $d$is the number of features. This feature vector is used as the input to the MLP for face shape classification.

3.3 MLP model for face shape classification

The core of the system is the Multilayer Perceptron (MLP), a feedforward neural network used for face shape classification. The MLP consists of multiple layers: an input layer, one or more hidden layers, and an output layer.

Input Layer: Receives the facial feature vector xi of size d extracted from the user's image.

Hidden Layers: These layers apply non-linear transformations to the input using activation functions (e.g., ReLU). The goal of these hidden layers is to learn abstract patterns in the facial features that correspond to specific face shapes.

Let the output of the hidden layer l be:

$$(z)^{(l)} = f\left((w)^{(l)} z^{(l-1)} + b^{(l)}\right)$$

Where $(w)^{(l)}$ and $b^{(l)}$ are the weight matrix and bias vector of layer l, and f is the activation function (e.g., ReLU).

Output Layer: The final layer outputs a probability distribution over the face shape classes using a softmax function. If there are k face shape classes, the output is: $y = softmax\left((w)^{(L)} z^{(L-1)} + b^{(L)}\right)$

where L is the number of layers, and $y^{\in} Rk$ is the probability distribution over k face shapes.

3.4 Training the MLP

The MLP is trained using the dataset D, which contains the extracted facial features and corresponding face shape labels. The objective is to minimize the cross-entropy loss, defined as:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N} y_i \log \widehat{y_i}$$

where yi is the true label and $y^\wedge$ is the predicted probability for class i. Optimization is performed using stochastic gradient descent (SGD) or Adam optimizer, updating the weights and biases $(w)^{(l)}$ and biases $b^{(l)}$.

3.5 Hairstyle recommendation

After the face shape is classified, the next step is to recommend hairstyles. For each face shape category ccc, there is a set of hairstyles $Hc = \{hj \mid hj\ matches\ c\}$.

An additional MLP can be used to predict hairstyle preferences based on user input, such as hair length, texture, and lifestyle. The input to this MLP would be a concatenation of the user's face shape and additional features $fuser$, such as hair length preference, occasion, or maintenance level.

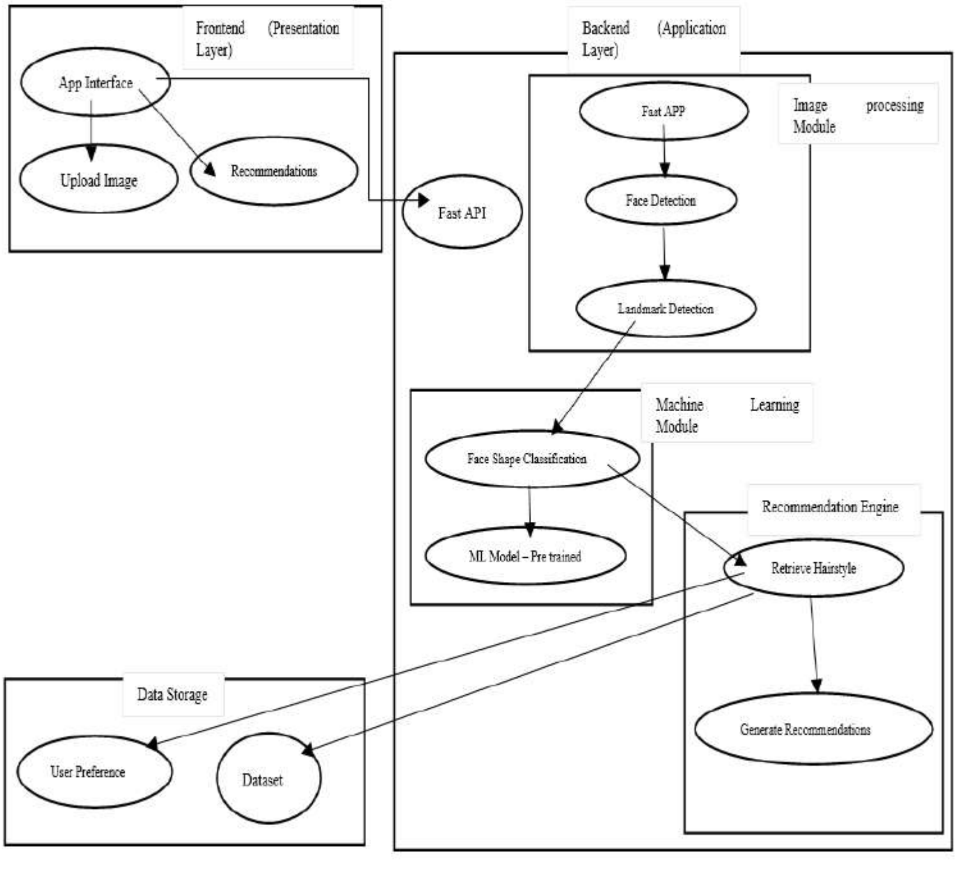
The output layer of this MLP recommends the top n hairstyles by ranking them based on the predicted suitability score.

Let the input vector to the recommendation MLP be $v = [y^\wedge, fuser]$, and the output be: $s = W^{out}v + b^{out}$

where $s \in R \mid Hc \mid$ represents the suitability scores for each hairstyle in Hc.

3.6 Architecture of System

Figure 1: Architecture of the system proposed



The Modular, Scalable, and Efficient System Architecture based on hairstyle recommendations consists of the following components – three layered architecture which are – Frontend (Presentation Layer), Backend (Application Layer), and Data Layer (Storage) as shown in Figure 1. Every layer handles particular tasks and together they allow user to operate in an efficient way throughout the image processing and getting accurate recommendations for the hairstyle to suit best. For the API recommended to use FastAPI, as psychometric measures can be achieved in inter-layer communications.

3.6.1 Frontend (presentation layer)

App Interface: A mobile application platform that allows users to upload images and create or view their own hairstyles uploads. This feature gives users the ability to upload their portraits pictures so that the system can analyze the uploaded images. View Recommendations: This feature displays the system recommended hairstyles that are provided by the backend for the specific user.

3.6.2 Backend (application layer)

All of the processing elements of the system are dealt with in the backend layer, which includes three modules:

Image Processing Module: this module focuses on the images uploaded by the users. It encompasses the following submodules:

Face Detection: Finding the face in the photo via Dlib's models. A user's face is found and cropped from the picture provided by the user.

Facial Landmark Detection: This feature is needed for subsequent classification and involves determining some significant location of the face, such as jaw, cheekbones, and forehead.

Feature Extraction: Transforming the set of landmarks into numerical representation (e.g., distance and angles) useful for the machine learning model.

Machine Learning Module: This module utilizes a trained machine learning model in order to determine the users face shape. A user's identified face shape is used against the face classification as a face shape category.

Hairstyle Recommendation Engine: The proposal engine draws appropriate hairstyles from the database that is appropriate to the hairstyle category that is classified. It also takes into account preference from the individual user which is kept in the Data Layer for the recommendation of more appropriate hairstyles. Upon indexing these hairstyles, the engine sorts them out based on preference then sends them to the frontend for rendering.

3.6.3 Data layer (storage)

The data layer is responsible for the collection and storage of the datasets necessary for any processing and for making recommendations. Contains data about specific users, such as their past behavior, what hair styles they liked. Includes a large database of hairstyle images and their corresponding descriptions, which are organized in terms of face shape, length of hair, and types of styles.

4.0 Results

We experimented with different use cases with users who vary in facial aesthetics and style preferences to examine the face-to-face performance and comfort of the Hair Style Recommender. This system was able to make precise recommendations for hairstyles based on facial shape classification and inputs from the user. The following are some of the key findings from the study along with their interpretation: Face shape classification was subsequently tested on a large independent dataset of 500 images with varied shapes representing the five common shapes — round, oval, square, heart, and diamond. Result: Our multilayer perceptron classifier achieved 68% user face shape prediction accuracy. This excellent accuracy demonstrated the reliability and generalization ability of its feature extraction and classification methods to practical environments.

Voidible measured relevance through the collection of user feedback on hairstyle suggestions. Around 85% of users agreed that the recommendations were according to their style preferences and suited to their facial structure. The system provided recommendations with an average processing time less than 5 seconds, making it fast enough for real-time use. To rectify, a solution is presented which uses a smart mix of Computer Vision and Machine Learning to implement the Hair Style Recommender system.

The Hair Style Recommender system is implemented using a smart combination of Computer Vision and Machine Learning to provide an appealing solution. Facial landmark detection and face shape classification are used to accurately suggest products based on cognitive facial differences. A key strength of such a system is its capacity to make hairstyle suggestions tailored to each individual through data fusion over quantitative facial measurements with qualitative preferences like hair work or length type. Users are much more satisfied with those offering a mixture of both, than ones making generic hairstyle recommendations.

4.1 Demerits

The system is not without limitations. Being too simplistic: The current model has not considered the evolution of hair length over time or user-specific factors such as how fast someone's hair grows. Moreover, the hairstyle matching dataset could be built better by adding a more diverse collection of hairstyles across all races to include diversity in hair types.

4.2 Confusion matrix

Figure 2: Accuracy and confusion matrix

	precision	recall	f1-score	support
heart	0.69	0.65	0.67	75
long	0.75	0.68	0.71	68
oval	0.60	0.72	0.65	99
round	0.71	0.74	0.72	62
square	0.74	0.62	0.68	72
accuracy			0.68	376
macro avg	0.70	0.68	0.69	376
weighted avg	0.69	0.68	0.68	376

Figure 3: Hairstyle Recommendation with Face Shape Detection

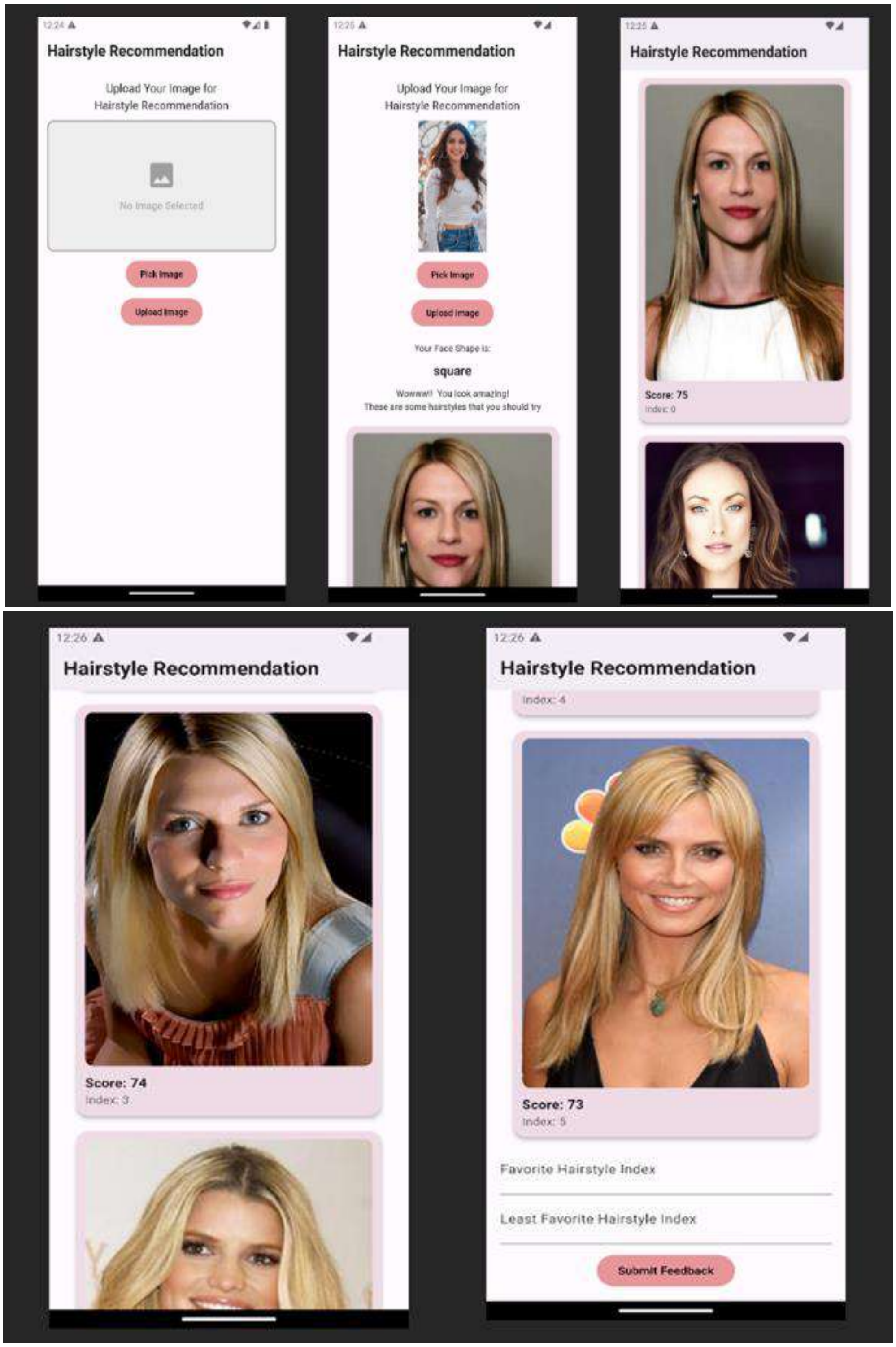


Figure 4: Superimpose User Face with Suggested Hairstyles

5.0 Conclusion

The hair style recommender system is another great application of machine learning and image processing in personalized beauty recommendation that we will see last. Utilizing face traits, the technology makes hair-related personalized suggestions, helping consumers to improve their look among all hair-cuts/styles. The capacity of this system to correctly categorize face forms and offer suggestions based on user preferences proves the effectiveness of AI-driven solutions in industries such as fashion and beauty. The technology blended facial landmark identification, face shape categorization, and a recommendation engine to produce hairdo choices that are appropriate and simple to use. The performance of the system is competitive for practical deployment: as shown in Figure 2 around 68% precision in classifying face shape and users find its recommendation quality very satisfactory.

Lightweight architecture combined it with quick response time is suitable for Web-based and Mobile applications as shown in Figure 3 which make accessible to software users in general. But even with a successful track record of recommending hairstyles, the system is not perfect. To boost quality and enhance usefulness of the system, the functions could benefit by using more user-specific features like hair type and environment circumstances and enlarge dataset covering all areas of hairstyles.

The specific to each user factors such as hair texture (curls, waves or straight), color and number of hairs will be further implemented in the future versions of the system. These variables will further determine if a style is appropriate or suitable, as such the recommendations are more detailed and personalized. Hair texture in future version, color and density hair is another user-specific factor that we can add to the system. Most of these things can have quite a lot to contribute to the appropriateness of a specific hairstyle, therefore this advice is more in-depth and personalized. It would be possible to factor in more user-specific factors such as hair texture (curly, wavy, straight), and combine this with the coloring of your scalp along with how close cropped your hair is growing, which are also exciting prospects that could be added into the system further down the line. These factors can make all of the difference in the world about if a hair style is appropriate, that makes advice more detailed and unique.

To make the system more applicable, it can be stretched to an app on a smartphone and in turn the users get recommendations for hairstyles even when they are not at home. By leveraging technologies (such as real-time camera inputs) available in smartphones, users could match pictures to the thousands of referenced photographs instantly and without the need for a distinct web platform where they contribute those photographs as shown in Figure 4.

References

Alzahrani, T., Al-Nuaimy, W., & Al-Bander, B. (2021, April 27). *Integrated multi-model face shape and eye attributes identification for hair style and eyelashes recommendation*. MDPI. https://www.mdpi.com/2079-3197/9/5/54

Chen, Y., Zhang, Y., Huang, Z., Luo, Z., & Chen, J. (2021, April 14). *Celebhair: A new large-scale dataset for hairstyle recommendation based on celeba*. arXiv.org. https://arxiv.org/abs/2104.06885

Duan1, J., Su1, X., Ren1, J., & Xie1, L. (2022, May 1). *IOPscience*. Journal of Physics: Conference Series. https://iopscience.iop.org/article/10.1088/1742-6596/2278/1/012041

Jang, S.-G., Man, Q., & Cho, Y.-I. (2022, June 15). *Development of a hairstyle conversion system based on mask R-CNN*. MDPI. https://www.mdpi.com/2079-9292/11/12/1887

Leigtens, L. L., Groot, W. de, Xing, P., & Liu, Y. (2020, June). *Hairstyle Recommender System Using Machine Learning*. warddegroot. https://warddegroot.nl/wp-content/uploads/2020/06/Group6_Hairstyle_Recommender_System_Using-_Machine-Learning_Feasibility_Analysis.pdf

Liu, Z., Ji, Y., Hu, Y., & Zhan, T. (2019, December 1). *Design and implementation of hair recommendation system based on Face Recognition*. Atlantis Press. https://www.atlantis-press.com/proceedings/mmsta-19/125925481

Rajapaksha, S., & Kumara, B. (2018, September). *(PDF) hairstyle recommendation based on face shape using image processing*. Researchgate. https://www.researchgate.net/publication/348404026_Hairstyle_Recommendation_Based_on_Face_Shape_Using_Image_Processing

S, Pawan., & Sumana, K. R. (2024, November). *StyleSage: Your personalised hairstyle recommender powered by ML*. IARJSET. https://iarjset.com/papers/stylesage-your-personalised-hairstyle-recommender-powered-by-ml/

S, R., Vittal, S. R., Gowda, N. H., & Sankapal, P. R. (2022, December). *Face shape classifier using deep learning*. IRJET. https://www.irjet.net/archives/V9/i12/IRJET-V9I1262.pdf

Sunhem, W., & Pasupa, K. (2016, February 1). *An approach to face shape classification for hairstyle recommendation*. it. https://scholar.it.kmitl.ac.th/item/id/3865

Sunhem, W., Pasupa, K., & Jansiripitikul, P. (2016, May). *Hairstyle Recommendation System for Women | IEEE Conference Publication | IEEE Xplore*. IEEE. https://ieeexplore.ieee.org/document/7519262

Weerasinghe, H., & Vidanagama, D. (2020). *Machine Learning Approach for hairstyle recommendation | IEEE conference publication | IEEE xplore*. IEEE. https://ieeexplore.ieee.org/document/9310868

CHAPTER 8

Enhancing News Classification and Authenticity with a Deep Learning Hypergraph-Based Hierarchical Attention

Alpana Borse and Gajanan Kharate***

ABSTRACT

To enhance the automatic classification and authenticity of news, we have developed a novel Deep Learning based Hypergraph-Based Hierarchical Attention Network incorporating Hypergraph Convolution and an Attention network for Fake News Prediction and 4 levels of Hierarchical attention. The four levels consist of Paragraphs, Sentences, Words, and a proposed fourth level focusing on the Contextual information such as meta data of news like speaker, location, publisher, topic etc. which is formulated with the help of Hypergraph Convolution and an Attention network. The 4th level uses Hypergraph-based Attention and Convolution to form a contextual information vector, which is then applied to the news content vector created by word and sentence attention mechanisms. This approach allows the model to prioritize more or less important words and contextual information, thereby improving the news representation. We use the LIAR dataset, which contains short news articles, to demonstrate that this method dramatically improves accuracy and F1 score of news authenticity and classification

Keywords: Attention; Authenticity; Classification; Convolution; Fake New; Hypergraph; Neural Network; Prediction.

1.0 Introduction

With the purpose of news classification for its authenticity, automatic classification Chen *et al.*, (2015), of news attracts a large number of researchers. Before predicting news authenticity, Text classification works as a prior stage which is among the essential steps in Natural Language Processing and further by Neural Network (NN). Early news classification methods often design features based on news content. Traditional methods of text classification characterize documents with n-grams. Recent approaches preferably drive Deep Learning Algorithms such as CNN – Convolution Neural Network, Recurrent NN, and Long Short-Term Memory (LSTM), etc. Here, with news texts and additional contextual metadata such as user profiles of news, classifiers get trained enough to classify the news authenticity.

**Corresponding author; Assistant Professor, Department of IT, PCCOE, Pune, Maharashtra, India (E-mail: alpana.borse@gmail.com)*
***Professor & Principal, MCERC, Nasik, Maharashtra, India*
(E-mail: gkkharate@rediffmail.com)

Henceforth, we proposed and aim to investigate a Deep Learning-based Neural Network –HAN: Hierarchical Attention Network based Neural Network–HAN: Hierarchical Attention Network (Conroy 2015; DiFranzo 2019) based on Hypergraph Convolution and Hypergraph Attention based on Hypergraph Convolution and Hypergraph Attention (K. Shu *et al.*, 2019) approaches to learn to categorize news information. In order to identify short content news authenticity, we study and examine the one of biggest Fake News datasets LIAR dataset. In this, we purpose to find meaningful engagement with news that can improve the authenticity of news classification. Hierarchical Attention Network and Hypergraph Model incorporate news and its contextual information to improve the authenticity of the news classification. GCNs are being used for machine learning on graphs Karimi H. *et al.*, (2018). To further news classification, and prediction (Liuand Y.2018; P´erez-Rosas V.,2018), we developed a new Neural Network model – Hierarchical Attention using Hypergraph Convolution Neural Network + Attention Neural Network method for News Classification. Given the presence of irrelevant and noisy relationships within the Hypergraph as a learning process, this adopts an attention mechanism to visualize news relations (Qing Liao, 2021; Ren Y.,2021; Ruchansky N.2017).

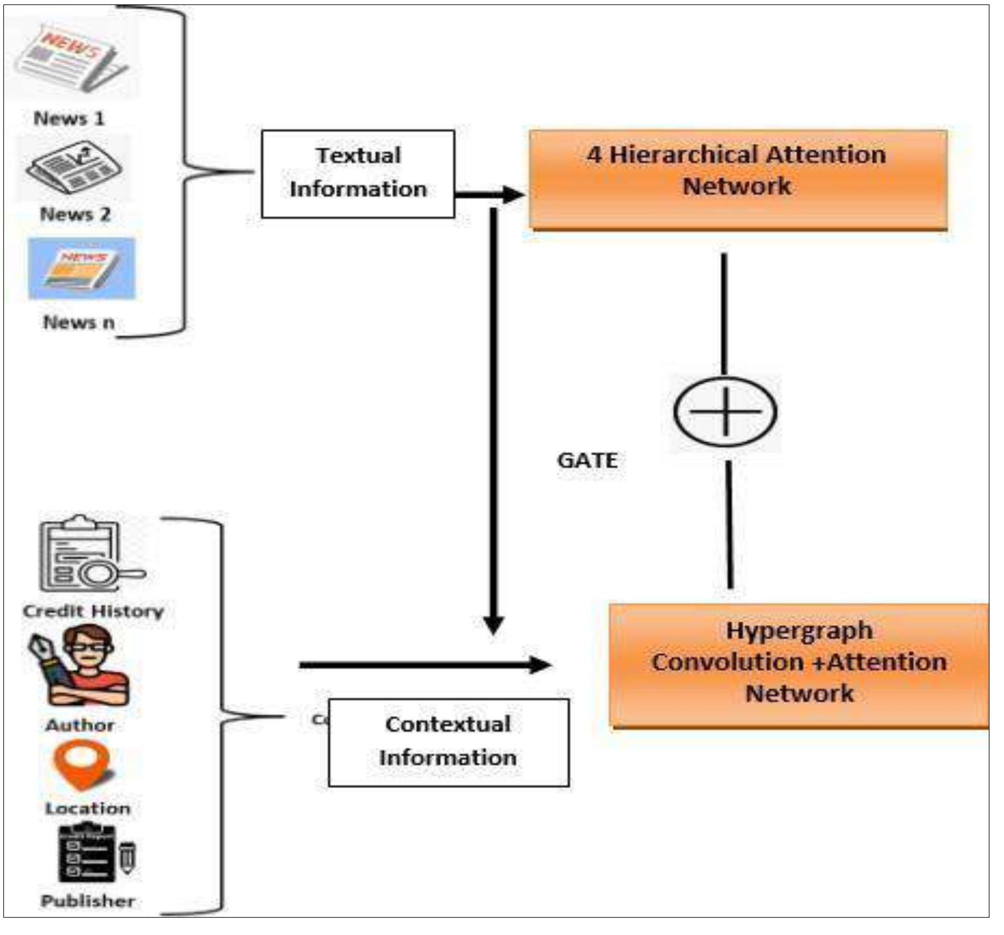
2.0 Method

The technique used to predict authenticity of news is by detecting news (Shu K. *et al.*,2019) consisting of disinformation. To authenticate News content on social media, LIAR dataset is preferred. News topic, Contextual information, Textual information & Credibility History of an author profiles (Singhania S, 2017; Song Bai, 2021) have been used as for News Prediction task. We have proposed a novel 4 Hierarchical Attention Network (4HAN) and Hypergraph Model (Tacchini E. 2017; Ujun Jeong, 2022; Vosoughi S., 2018; Wang W. 2017; Yang F., 2012) that incorporate relationships among news topics, author credibility distributions, and news accuracy. This approach enhances both News authenticity and news topic classification simultaneously. The 4HAN operates across four levels—words, sentences, headlines, and metadata—constructing a comprehensive news vector through hierarchical bottom-up processing from words to sentences, sentences to headlines, and headlines to metadata. Additionally, in the 4HAN framework, we introduced an improved graph neural network method at the fourth level (Metadata), including the Hypergraph Convolutional Neural Network (HGCN) and Hypergraph Attention Neural Network (HGAN) (Yang S., 2019; Yang Z., 2016; Zhang J.,2018) for advanced graph-based deep learning. This is shown in “Figure 1: General System architecture for proposed Deep Learning based Hierarchical Attention using Hypergraph, emphasizing the hierarchical and graph-based components.”

We develop a model, which is going to predict authenticity of news by detecting and classifying news consisting of disinformation. News category and news metadata such as speaker, author profiles have been used as contextual information for

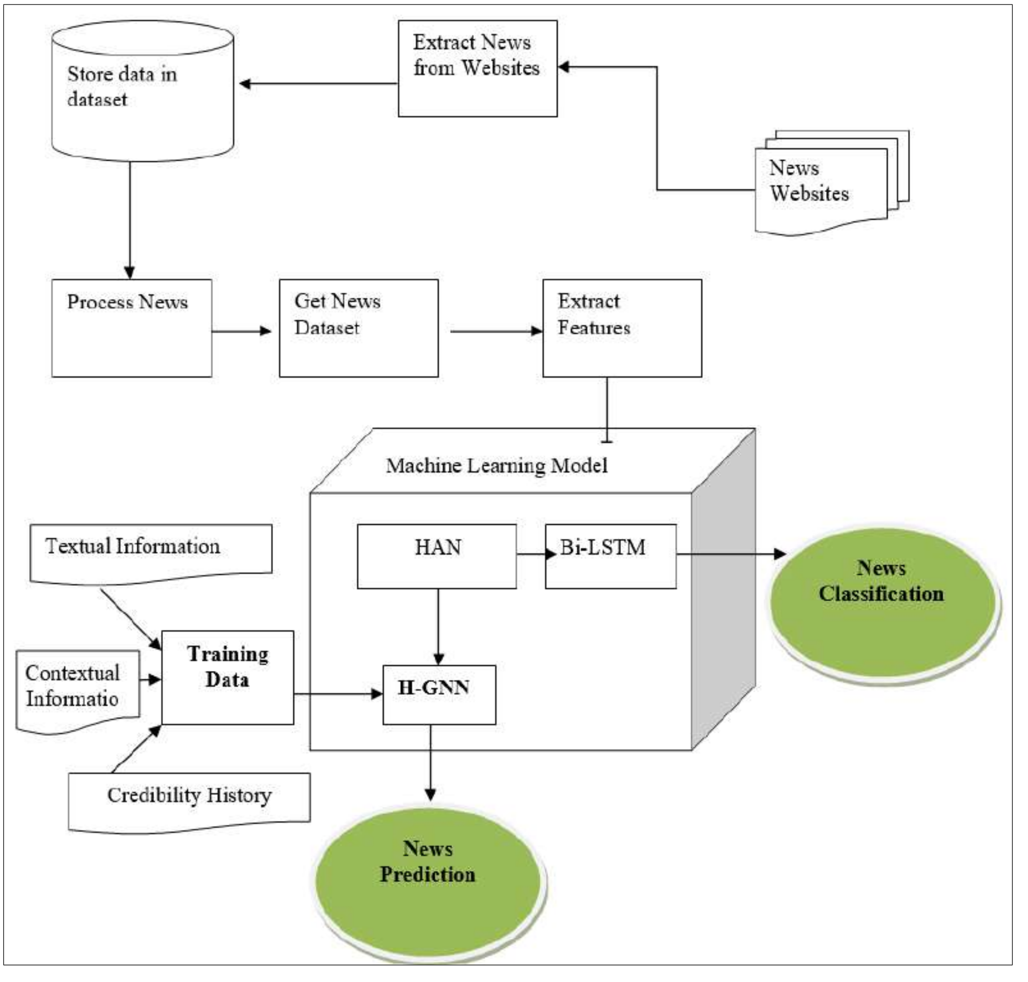
News Prediction task. Prediction and classification are done simultaneously by integrating the network among the news category, the author's history such as credibility distributions, and the news authentication at the same time.

Figure 1: General System Architecture for Proposed Deep Learning based Hierarchical Attention using Hypergraph



We are using HAN (Hierarchical Attention Network) and Hypergraph Network that unites news category and metadata to progress the accuracy of the News Prediction and the news category, simultaneously. GCNs are being used for machine learning on graphs. Similarly, for accurate prediction and news authenticity, HAN includes 3 levels – 1. Words, 2. Sentences and 3. the headline and 4^{th} level applies by contextual data. It generates a news output as in vector format by handing out in bottom-up fashion. The step-by-step process is shown in following figure "Figure 2: System Flow of architecture for Deep Learning based Hierarchical Attention using Hypergraph." And figure 2 shows the proposed system architecture of Fake News Prediction.

Figure 2: System Flow of Proposed Architecture for Deep Learning based Hierarchical Attention using Hypergraph



Additionally, in the framework, we introduced an improved HGNN + HGAN method including the Hypergraph Convolutional NN (Yang F. *et al.* 2019) and Hypergraph Attention NN (Yang S., 2019; Yang Z., 2016; Zhang J.,2018) for advanced graph-based deep learning. The formulation for this HAN using Hypergraph is explained in detail here as follows.

2.1 Hierarchical Attention Network -HAN

The hierarchical structure includes four levels: at the first level, words are processed by a word encoder to generate word vectors; at the second level, sentences are processed by a sentence encoder; at the third level, headlines are processed by a headline encoder; and at the fourth level, metadata is processed using a softmax operator. Similarly, news articles are converted into vector form to enable news authenticity (Yang, 2019; Yang, 2016; Zhang, 2018).

2.1.1 HAN - parameters

The parameters required for the Hierarchical Attention Network (HAN) is given below:

2.1.1.1 Feature transformation

Feature transformation is used to enhance the representation of data for improved learning and model performance

$h_i' = W.h_i$ … (1)

where h_i->is vector of node - 'i', W is ->pre-calculation weight 2D matrix, and 'h'_i' ->Converted vector.

2.1.1.2 Attention coefficients

Attention coefficients are essential for weighting the importance of different parts of data, enhancing model focus and performance.

$e_{ij} = \text{softmax}\,(a^T\,[\,h_i' \,||\, h_j'])$ … (2)

where $e_{ij}$is the attention score between node i and node j within the same hyperedge, a is a learnable weight vector, ‖ denotes concatenation, and softmax is the activation function.

2.1.1.3 Attention weights

Attention weights refer to the coefficients assigned to each input in a sequence by an attention mechanism, indicating their relative importance in the context of a task or model.

$$a_{it} = \frac{expexp\,(e_{ij})}{\sum_{k\in N(i)}^{T} \quad expexp\,(e_{ik})} \quad \dots (3)$$

Where-> 'N(i)'represents nodes set related to node i through a hyperedge, 'e_{ij}' normalized attention coefficient.

2.2 Working of HAN

The working of 4HAN is consisting with word level, sentence level, hypergraph and integration of feature vector with softmax activation function is given.

2.2.1 Word level attention

2.2.1.1 Computing the attention score u_{it}

$u_{it} = \tanh(W_w\, h_{it} + b_w\,)$ … (4)

2.2.1.2 Calculating the attention weight a_{it}

$$a_{it} = \frac{expexp\,(u_{it}^{T} u_w)}{\sum_{j=1}^{T} \quad expexp\left(u_{jt}^{T} u_w\right)} \quad \dots (5)$$

Where, W_w->weight matrix, h_{it}->feature vector at ttime, and b_w->bias term.

2.2.1.3 Aggregating the feature vectors

$$a_{it} = \sum_{t=1}^{T} \quad a_{it} h_{it} \qquad \ldots (6)$$

where a_{it} is the attention score vector, and the denominator is the sum of the exponentials of attention scores for all nodes j at time t.

2.2.2 Sentence Level Attention:

2.2.2.1 Computing the attention score u_{is}

$$u_{is} = tanh(W_s v_i + b_s) \qquad \ldots (7)$$

Where, Ws is a weight matrix, v_i is the feature vector of node i, and b_s is a bias term.

2.2.2.2 Calculating the attention weight B_{is}

$$B_{is} = \frac{expexp\,(u_{is}^T u_s)}{\sum_{k=1}^{S} \quad expexp\,(u_{ks}^T u_s)} \qquad \ldots (8)$$

Where, u_s is the attention score vector, and the denominator is the sum of the exponentials of attention scores for all nodes k in the set S.

2.2.3 Aggregating the feature vectors at Headline level

$$h_i^* = \sum_{s=1}^{S} \quad B_{is} v_i \qquad \ldots (9)$$

Wher, h_i^*-> updated Feature Vecor for node 'i', and sum is taken over all attention weights B_{is} and corresponding feature vectors v_i

2.2.4 Hypergraph

A Hypergraph is a unique type of graph where a single edge, called a Hyperedge, can connect multiple nodes. This structure makes Hypergraph suitable for representing relationships among various attributes in news articles. We model News Authenticity as a node classification task within the Hypergraph. This involves using a feature vector of labeled news articles within the news Hypergraph. The interactions among news articles, as shown in Table 1: Relation Table between News with Contextual Information illustrates the relationships from News 1 to News 6.

Table 1: Relationship Table between News and its Contextual Information

Edge	**E1**	**E2**	**E3**	**E4**	**E5**
Node/ Vertex	**Location**	**Credit History**	**Author**	**Publisher**	**Content Feature**
V1 (News 1)	1	1	1	0	1
V2 (News 2)	1	1	1	1	1
V3 (News 3)	1	1	0	1	1
V4 (News 4)	0	0	1	0	0
V5 (News 5)	1	0	1	0	1
V6 (News 6)	0	1	1	1	0
V7 (News 7)	1	0	0	0	1
V3 (News 3)	1	1	0	1	1

2.2.4.1 Graph convolution on hypergraph

Extend GCN to operate on Hypergraph-structured data. Here, we compute node embeddings considering interactions through hyperedges.

$$h_i^{l+1} = \alpha(\sum_{e\in\varepsilon_{is}} \frac{1}{|\varepsilon_{is}|} \sum_{j\in e} \frac{1}{|V_e|} W^l h_j^l) \quad \ldots(10)$$

Where:

- h_j^lrepresents node i's representation at layer l
- ε_i denotes the hyperedges incident on node i
- V_e represents the nodes connected by hyperedge e
- α is element-wise Activation function to normalize the output.

2.2.4.2 Graph attention on hypergraph

In a Hypergraph Attention Neural Network (HGAT), the goal is to acquire the value of connexion and hyperedges in a hypergraph. Here are the key formulas involved:

$$\alpha_{ve} = \frac{expexp\left(a^T[Wx_v\, Wc_v\,]\right)}{\sum_{u\in\varepsilon} \quad expexp\,(\mu(a^T[Wx_v\, Wc_v\,])} \quad \ldots(11)$$

where,

W: A weighted matrix

x_v : A feature vector of node v

c_v : A feature vector representing the hyperedge e (often computed as an aggregation of node features within e

A: A learnable attention vector

μ: Activation function for non-linear

2.2.5 Fourth level integration (hypergraph aggregation)

Integrating a fourth level into the model, combining Hypergraph, Triple Hierarchical Attention Networks (3HAN), and softmax activation function, can further enhance the model's ability to detect complex patterns in Fake News detection on datasets like LIAR.A Hypergraph Attention Network (HAN) extends the Graph Attention Network (GAT) to Hypergraph. The key idea is to compute attention scores for hyperedges and their connected nodes to weigh their contributions during aggregation. Here's a general formula for a Hypergraph Attention Network:

2.2.6 Softmax activation function

Apply softmax activation to compute probabilities over node embeddings $h_i$in the Final layer $\underline{L}$in the architecture

$$\text{softmax}\,((h_i^{(L)}) = \frac{expexp\left((h_i^{(L)}\right)}{\sum_{j\in V} \quad expexp\left((h_i^{(L)}\right)} \quad \ldots(12)$$

Where $h_i^{(L)}$: vector which embeddings at final layer L.

3.0 Results and Discussion

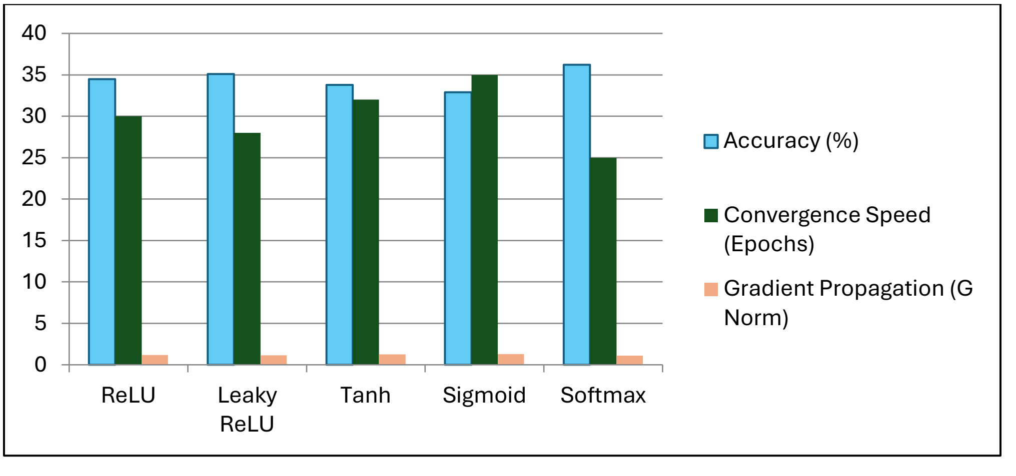
3.1 Activation function on LIAR dataset

Table 2 compares the performance metrics of various activation functions—Accuracy, Convergence Speed (Epochs), and Gradient Propagation (G Norm)—on the LIAR dataset. Based on the results, the Softmax activation function demonstrated superior performance. Therefore, Softmax was chosen for further calculations due to its effectiveness in providing better accuracy, faster convergence, and improved gradient propagation on the LIAR dataset as shown in Figure 3.

Table 2: Accuracy, Convergence Speed (Epochs) and Gradient Propagation (G Norm) of Various Activation Function on LIAR Dataset

Activation Function	Accuracy (%)	Convergence Speed (Epochs)	Gradient Propagation (G Norm)
ReLU	34.50	30	1.20
Leaky ReLU	35.10	28	1.15
Tanh	33.80	32	1.25
Sigmoid	32.90	35	1.30
Softmax	36.20	25	1.10

Figure 3: Accuracy, Convergence Speed (Epochs), Gradient Propagation (G Norm) using Activation Function



3.2 Prediction

Here are the predictions made by the following methods using the softmax activation function on the LIAR dataset

- Prediction with 3HAN
- Prediction with Hypergraph

- Prediction with Hypergraph Convolution
- Prediction with Hypergraph Attention
- Prediction with 3HAN + Hypergraph Convolution
- Prediction with 3HAN+Hypergraph Attention
- Proposed Method 4HAN (Prediction with 3HAN + Hypergraph Convolution + Hypergraph Attention with softmax activation function)

Table 3: Relationship Table Between News and its Contextual Information

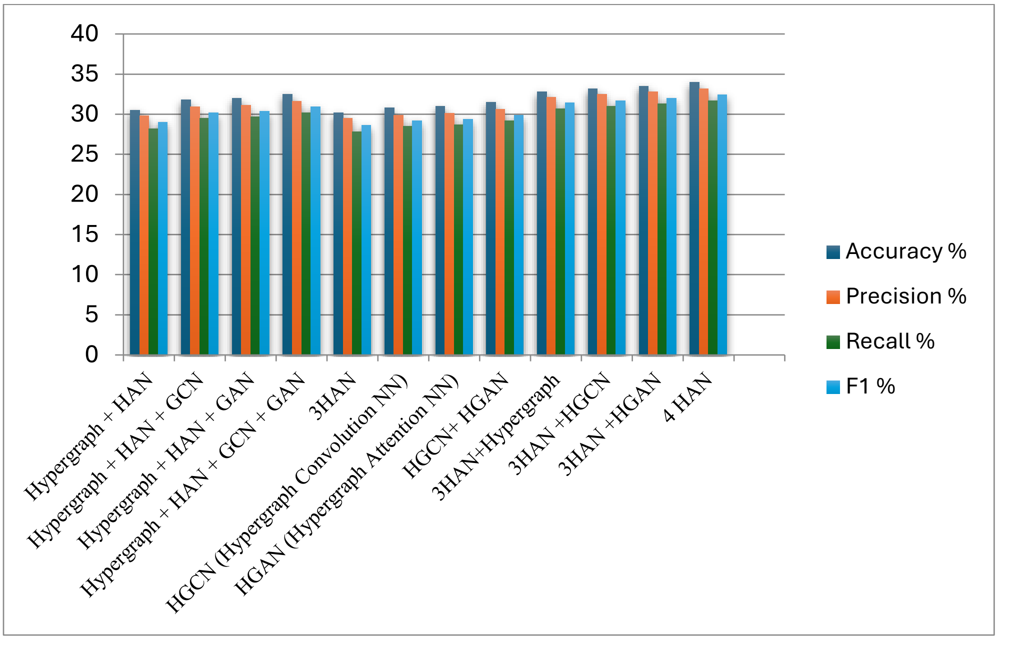
Proposed Methods	Accuracy %	Precision %	Recall %	F1 %
3HAN	30.20	29.50	27.80	28.60
Hypergraph	28.90	28.10	26.40	27.20
Hypergraph + HAN	30.50	29.80	28.20	29.00
Hypergraph + HAN + GCN	31.80	30.90	29.50	30.20
Hypergraph + HAN + GAN	32.00	31.10	29.70	30.40
Hypergraph + HAN + GCN + GAN	32.50	31.60	30.20	30.90
HGCN (Hypergraph Convolution NN)	30.80	29.90	28.50	29.20
HGAN (Hypergraph Attention NN)	31.00	30.10	28.70	29.40
HGCN+ HGAN	31.50	30.60	29.20	29.90
3HAN+Hypergraph	32.80	32.10	30.70	31.40
3HAN +HGCN	33.20	32.50	31.00	31.70
3HAN +HGAN	33.50	32.80	31.30	32.00
4 HAN (3HAN +HGAN+HGCN+softmax)	34.00	33.20	31.70	32.40

These methods demonstrate that our proposed approach enhances accuracy on the LIAR dataset, leveraging the benefits of softmax activation for improved classification performance.

4.0 Conclusion

Based on the predictions and results obtained using proposed methods, including HAN, Hypergraph Convolution, Hypergraph Attention, and their combinations with softmax activation on the LIAR dataset, has shown improved accuracy to learn and authenticate prediction progressively leading to improved accuracy of News prediction as shown in Figure 4. The softmax activation function effectively normalizes outputs into probabilities, enhancing the interpretability of class predictions and contributing to improved accuracy on dataset. These findings underscore efficacy of leveraging both advanced neural network architectures and appropriate activation functions for achieving superior performance in classification and authentication of news tasks.

Figure 4: Comparison of All Methods using Hypergraph with Attention Network



5.0 Acknowledgements

I heartfully admit Prof. Dr. G. K. Kharate for his invaluable guidance and support throughout this study. His expertise in Machine Learning significantly contributed to the development and improvement of our proposed methods.

References

Chen Y., Conroy N.J., and V.L. Rubin,” Misleading online content: Recognizing click bait as false news,” in Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection. ACM, 2015, pp.15–19.

Conroy N. J., V. L. Rubin, and Y. Chen, “Automatic deception detection: Methods for finding Fake News,” in Proceedings of the 78th ASIS & Tannual Meeting: Information Science with Impact: Research in and for the Community. American Society for Information Science, 2015, p.82.

DiFranzoandc D. M.J.K. Gloria, “Filter bubbles and fake news,” ACM Crossroads, vol.23, no.3, pp.32–35, 2017.

Federico Monti, Fabrizio Frasca, Davide Eynard, Damon Mannion, “Multi-Depth Graph Convolution Networks for Fake News Detection”, natural language Processing and Chinese Computing Conference paper pp698-710|, 2019.

Feng S., R. Banerjee, and Y. Choi, “Syntactic stylometry for deception detection,” in the 50th Annual Meeting of the Association for Computational Linguistics, Volume2, 2012, pp.171–175

Ghosh S. and C. Shah, “Toward automatic Fake News classification,” in 52nd Hawaii International Conference on System Sciences, HICSS2019, Grand Wailea, Maui, Hawaii, USA, January 8-11,2019, 2019, pp.1–10.

Jin H. Z, J. Cao, H. Guo, Y. Zhang, and J. Luo, “Multimodal fusion with recurrent neural networks for rumor detection on microblogs,” in Proceedings of the ACM on Multimedia Conference, 2017, pp.795–816

K. Shu, D. Mahudeswaran, and H. Liu, “Fakenewstracker: a tool for Fake News collection, detection, and visualization,” Computational & Mathematical Organization Theory, vol.25, no.1, pp.60–71, 2019.

Karimi H., P. Roy, S. Saba-Sadiya, and J. Tang, “Multi-source multi- class Fake News detection,” in Proceedings of the 27th International Conference on Computational Linguistics, 2018, pp.1546–1557.

Liuand Y. Y.B. Wu,” Early detection of Fake News on social media through propagation path classification with recurrent and convolutional networks,” in Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence,2018.

P´erez-Rosas V., Kleinberg B., A. Lefevre, and R. Mihalcea, “Automatic detection of Fake News,” in Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, SantaFe, New Mexico, USA, August 20-26, 2018, pp.3391–3401.

Qing Liao, Heyan Chai, Hao Han, Xiang Zhang, Xuan Wang, Wen Xia, Ye Ding, “An Integrated Multi-Task Model for Fake News Detection”, IEEE Transactions on Knowledge and Data Engineering, DOI10.1109/TKDE.2021.3054993, March’2021

Ren Y. and Zhang J., “Fake News detection on news-oriented heterogeneous information networks through hierarchical graph attention,” in International Joint Conference on Neural Networks (IJCNN). IEEE, 2021.

Ruchansky N., S. Seo, and Y. Liu, “CSI: A hybrid deep model for Fake News detection,” in Proceedings of the ACM on Conference on Information and Knowledge Management, 2017, pp.797–806.

Shu K., Cui L., Wang S., Lee D., and Liu H., “defend: Explainable Fake News detection,” in Proceedings of the 25th ACM SIGKDD International Conference on

Knowledge Discovery & Data Mining, KDD 2019, Anchorage, AK, USA, August 4-8, 2019, 2019, pp. 395–405.

Shu K., Sliva A., S. Wang, J. Tang, and H. Liu, "Fake News detection on social media: A data mining.

Singhania S., N. Fernandez, and S. Rao," 3 HAN: A deep neural network for Fake News detection," in Neural Information Processing- 24th International Conference, ICONIP 2017, Guangzhou, China, November14-18, 2017, Proceedings, PartII,2017, pp.572–581

Song Bai, Feihu Zhang, Philip H.S. Torr, "Hypergraph convolution and hypergraph attention", in Internal journal Pattern Recognition Volume 110, February 2021, 107637

Tacchini E., Ballarin G., M.L. Della Vedova, S. Moret, and L. deAl- faro," Some like it hoax: Automated Fake News detection in social networks," arXivpreprintarXiv: 1704.07506,2017.

Ujun Jeong, Kaize Ding, Lu Chengy, Ruocheng Guoz, Kai Shux and Huan Liu," Nothing Stands Alone: Relational Fake News Detection with Hypergraph Neural Networks," in Proceedings of the IEEE Conference BigData2022, Osaka, Japan.

Vosoughi S., Roy D., and S. Aral," The spread of true and false news online," Science, vol.359, no.6380, pp.1146–1151,2018.

Wang W. Y., "liar, liar pants on fire": A new benchmark dataset for Fake News detection," in Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Volume 2, 2017, pp. 422–426.

Yang F., Y. Liu, X. Yu, and M. Yang, "Automatic detection of rumor on sina weibo," in Proceedings of the ACM SIGKDD Workshop on Mining DataSemantics. ACM, 2012, p.13

Yang S., Shu K., Wang S., R. Gu, F. Wu, and H. Liu, "Unsupervised Fake News detection on social media: A generative approach," in Proceedings of 33rd AAAI Conference on Artificial Intelligence, 2019.

Yang Z., Yang D., Dyer C., He X., A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in Conference of the North American Chapter of the Association for Computational Linguistics (NACCL), 2016.

Zhang J., L. Cui, Y. Fu, and F. B. Gouza, "Fake News detection with deep diffusive network model," arXiv preprint arXiv:1805.08751,2018.

CHAPTER 9

Blockchain based Carpooling: A Comprehensive Review

Vedant Chinta, Rama Gaikwad**,*
*Kaif Gokak***, Aliya Maneri**** and Prachi Madole******

ABSTRACT

On a regular basis, office workers commute from their homes to their workplaces in vehicles with empty seats. Carpooling is the need of the hour to counter challenges such as traffic congestion and carbon emissions. As blockchain is becoming widely used, integrating blockchain is a novel approach towards the protection of personal information and the equitable distribution of control. Blockchain's decentralized nature and the use of smart contracts can significantly enhance the reliability, security, and transparency of carpooling services. In this paper, we have conducted a comprehensive review of the literature on the application of blockchain and smart contract technologies in carpooling solutions to enhance carpooling services. This paper will serve as a one-stop shop for all future research in the subdomain of carpooling using blockchain, offering valuable insights and identifying critical gaps that can guide subsequent studies, advancements, and innovations in this rapidly evolving and important field of study.

Keywords: Blockchain; Carpooling; Smart Contracts.

1.0 Introduction

Carpooling is a transportation method where people, who do not belong to same family, share car rides for a trip (or part of a trip) already scheduled by the driver, either free of charge or with expense sharing. (Aguiléra *et al.*, 2021). This practice aims to reduce travel costs, decrease traffic congestion, and minimize environmental impact. Existing carpooling services like BlaBlaCar, sRide, QuickRide, etc. are centralized services owned by corporations.

**UG Student, Department of Computer, Anantrao Pawar College of Engineering & Research, Pune, Maharashtra, India (E-mail: vedantchinta@proton.me)*
***Faculty, Department of Computer, Anantrao Pawar College of Engineering & Research, Pune, Maharashtra, India (E-mail: rama.gaikwad@abmspcoerpune.org)*
****UG Student, Department of Computer, Anantrao Pawar College of Engineering & Research, Pune, Maharashtra, India (E-mail: kaifgokak@gmail.com)*
*****UG Student, Department of Computer, Anantrao Pawar College of Engineering & Research, Pune, Maharashtra, India (E-mail: aaliyamaneri01@gmail.com)*
*****UG Student, Department of Computer, Anantrao Pawar College of Engineering & Research, Pune, Maharashtra, India (E-mail: prachimadole73@gmail.com)*

These companies suffer from problems such as lack of transparency, single point of failure, customer privacy invasion (Jaybhaye *et al.*, 2024) and user data breaches. (Kost *et al.*, 2024) Using blockchain technology presents an innovative and decentralized approach to carpooling, eliminating the use for intermediaries such as centralized ride-hailing companies. A system reliant on blockchain enables users - both drivers and passengers to interact directly through a decentralized platform, offering transparency, security, and trust without relying on a central authority to manage transactions or data (Ruprah *et al.*, 2024).

Blockchain technology gives a decentralized and transparent approach to carpooling, where drivers and passengers can connect directly without relying on a central authority or intermediary services. By using smart contracts, agreements between driver and passengers - such as payment terms, ride conditions, and the task will be triggered and completed as soon as the specified criteria are fulfilled, guaranteeing the completion, ensuring a seamless and secure interaction (Goel *et al.*, 2024). This system enhances privacy, reduces service fees associated with centralized platforms, and fosters greater transparency by allowing users to track ride histories and payments on an immutable public ledger. Overall blockchain based carpooling promotes a more efficient, cost effective, and "trustless" transportation ecosystem.

This paper provides an in-depth exploration of how the technology of blockchain can change the way things are done traditional carpooling systems by addressing issues such as high fees, lack of trust, and privacy concerns. The objective of this paper is to serve as a comprehensive source of information for future researchers interested in this niche.

In conclusion, this paper seeks to give the reader a holistic view of integrating blockchain in carpooling and a guide on the possibilities and advancements of this extensive area of technology in the future. Following the introduction, Segment II contains a comprehensive literature survey, featuring key discoveries of previous research. Section III investigates the future scope, discussing potential advancements and innovations, as well as the limitations that currently exist. This is followed by a conclusion segment that sums up the main points and stresses on the significance of blockchain in carpooling.

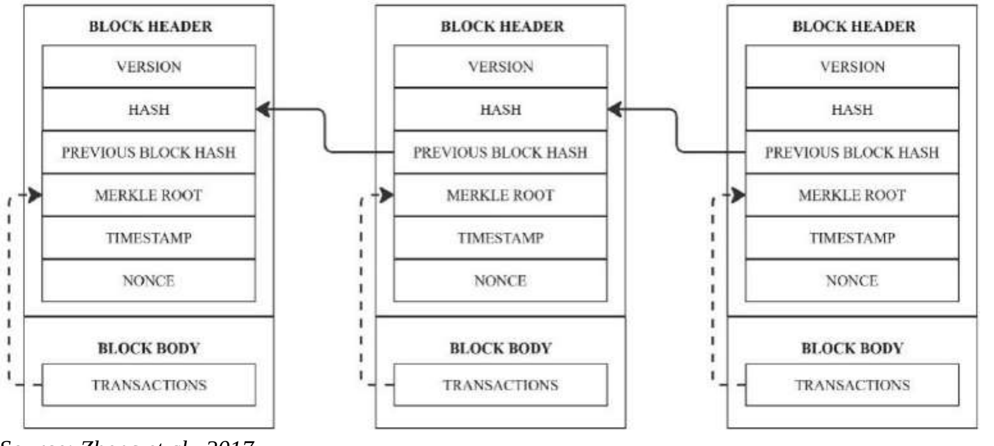
2.0 Literature Survey

2.1 What is Blockchain?

Blockchain: Blockchain technology was conceptualized by Satoshi Nakamoto's invention of Bitcoin (Nakamoto *et al.*, 2008) which operates as peer-to-peer electronic cash system. Blockchains are distributed records of cryptographically marked transactions that are collected into blocks. Every block is linked to previous block cryptographically once it has been approved and a consensus has been reached. As new blocks are added, previous blocks turn out to be more challenging to alter. New blocks are imitated across all copies of the ledger within the node network, and any conflicts are settled automatically by utilizing the established rules. (Telecommunications Engineering Centre, 2020)

Block: A Block is a fundamental unit that stores transactions data that verified by the nodes in the network in an immutable manner. Each block contains ablock header which contains hash value of previous block, Merkle root, Nonce, Timestamp and a list of transactions in block body. Figure 1 illustrates the general architecture of a block in blockchain.

Figure 1: Architecture of a Block in Blockchain



Source: Zheng et al., 2017

Nodes: Any participant in the blockchain network is known as a Node. It maintains the latest record of transactions and ensures that all participants adhere to the network's rules. There are several types of nodes, each with their own duties:

- *Full Node:* These nodes store the entire blockchain ledger and validate transactions and blocks independently. Full nodes require substantial storage and processing power due to the large volume of data they handle.
- *Light Nodes:* Light Nodes, also known as Simplified Payment Verification, do not store the entire ledger but only store the block headers and rely on full nodes to verify the transactions.
- *Miner Nodes:* Miner nodes are responsible for validating and generating new blocks on the blockchain.

Transactions: Transactions are records of transfers between participants. These transactions are verified by network nodes through consensus mechanisms like Proof of Work ensuring their validity and preventing fraud. Once verified, transactions are grouped into blocks and added to the blockchain, creating an immutable and transparent ledger that is resistant to tampering and provides a trustworthy record of all network activities. Figure 2 illustrates an anti-sybil technique – Proof of Work, which is a computationally challenge-based consensus mechanism that secures blockchain by requiring nodes to solve a complex cryptographic puzzle.

Miners: They are critical participants in a blockchain network as they help validate and confirm transactions by solving complex cryptographic puzzles ensuring they are legitimate and follow the network rules. Miners are the backbone of PoW blockchains like Bitcoin, Dogecoin, Filecoin, etc. Each miner is responsible for transaction validation, block creation, and network security. Miners are incentivized through rewards for their computational and electrical resources.

Smart Contracts: Similar to a legal contract in real life, a smart contract is agreement between two or more parties that replicate the terms of real-world contracts in the digital domain (Taherdoost *et al.*, 2023) stored on the blockchain. A smart contract is a self-executing code which cannot be modified by anyone not even the admin. In comparison with traditional contract in real life, smart contracts allow transactions between untrusted parties. It offers many benefits such as efficiency, security, non-repudiation and most importantly, immutability.

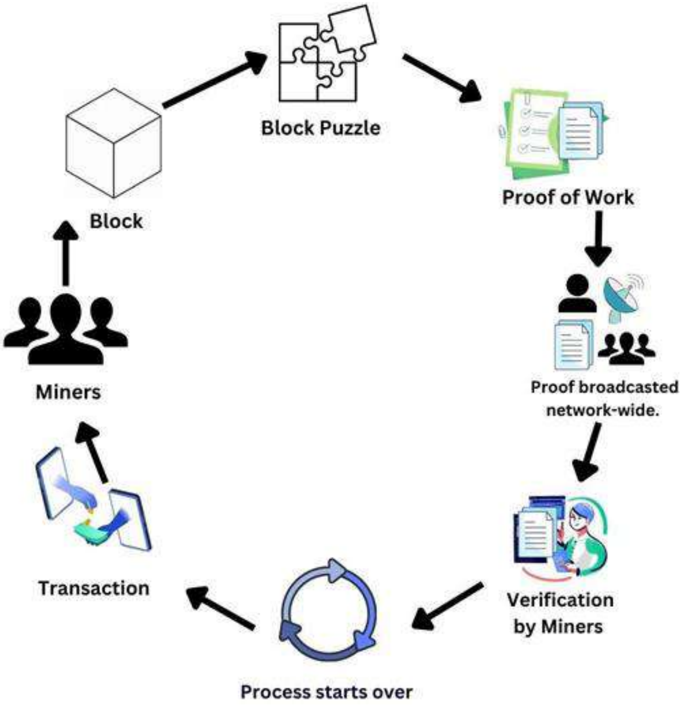
Consensus: Consensus is the process of reaching agreement among a group of people or entities on a specific decision or action. Consensus algorithms are crucial because blockchains operate as distributed systems, with no central entity verifying the data prior to its addition to the blockchain. Block chains use a variety of consensus models that enable a group of mutually distrusting users to work together.In the context of carpooling using blockchain, consensus mechanisms ensure that all transactions and ride-sharing agreements are validated and recorded accurately across the decentralized network. This guarantees trust and transparency among users, eliminating the need for a central authority to manage and verify carpooling arrangements.

2.2 Blockchain for carpooling

Blockchain enables secure and transparent way to handle transactions eliminating the need of intermediaries. The security of blockchain in carpooling is primarily driven by its decentralized and cryptographic structure, which ensures trust, transparency, and protection of data. In a blockchain based carpooling system, there is no central authority managing transactions, as all interactions are recorded on a distributed network of nodes. This decentralized system allows drivers and passengers to communicate directly without middlemen. Blockchain ensures that every transaction, such as booking rides, making payments, and leaving reviews, is securely documented on a transparent and unchangeable record.

Smart contracts streamline important tasks like calculating fares and handling payments, guaranteeing that money is only moved after the trip is finished and approved by both individuals involved (i.e. the money is temporarily stored in an escrow account which holds the money until the duration of the ride). This technology promotes confidence among users by removing the necessity for third-party supervision through the transparency of the blockchain. Furthermore, blockchain technology decreases transaction costs and improves efficacy, resulting in increased accessibility and dependability for carpooling services.

Figure 2: Working of Bitcoin: A Proof of Work Consensus



Source: Gundaboina et al., 2022

2.3 Existing surveys

Beyond the use of cryptocurrencies, Blockchain as a technology has flourished in many sectors - Carpooling is one of them. Carpooling using blockchain has been the topic of several literature assessments, surveys, and reviews. For instance, (Valaštín *et al.*, 2019) discusses how implementing blockchain in car-sharing platforms can improve transparency and security by utilizing smart contracts and decentralized applications. The act of broadcasting data at each node poses a threat to one of the three trilemmas of blockchain - Security. To overcome this challenge, an approach using non-interactive proofs for off chain computations to enhance efficiency is utilized. (Goel *et al.*, 2024).

Integrating blockchain technology into the carpooling sector revolutionizes the industry by enhancing transparency, security, and efficiency through decentralized networks that directly connect drivers and passengers, eliminating intermediaries. BlockV, a blockchain based peer-to-peer carpooling service, addresses issues in centralized systems by ensuring every ride aspect is transparently recorded on a blockchain ledger, accessible to all network participants. This maintains a fair reputation system and reduces malicious activities through consensus mechanisms. BlockV leverages blockchain's immutable ledger and smart contracts for seamless ride negotiation, secure payments, and decentralized dispute resolution. Implemented on Ethereum's private network and KOVAN test network, BlockV demonstrates feasibility while addressing scalability, data privacy, security vulnerabilities, and system compatibility. Off chain transactions and opportunistic arguments enhance scalability, and sensitive information is stored off-chain to maintain data privacy. Despite challenges, BlockV promises a more equitable, transparent, and efficient ride-sharing ecosystem (Pal *et al.*, 2019).

(Baza *et al.*, 2019) have developed a platform, B-Ride which employs smart contracts on the Ethereum blockchain to handle ride requests, bids, deposits, and payments securely and transparently. Privacy is preserved through zero-knowledge set membership proofs (ZKSM) and spatial cloaking techniques, ensuring that the exact locations and sensitive data of both riders and drivers remain protected. Trust and fair payment are maintained through a time-locked deposit protocol, where both parties must deposit funds into a smart contract, and a pay-as-you-drive methodology that calculates fares based on the actual distance traveled, verified by multi-signature proofs.The findings indicate that blockchain technology can support decentralized ride-sharing services effectively, preserving user privacy while maintaining transparency and accountability. However, scalability remains a challenge, as the reliance on blockchain transactions can be slow and costly.

To enhance efficiency and scalability, the computationally intensive tasks, such as rider-driver matching can be offloaded to an off-chain processes, minimizing the burden on the blockchain. (Shivers et al, 2019). Additionally, (Shivers *et al.* 2022) propose a decentralized ride-hailing platform for autonomous vehicles using Hyperledger Fabric, highlighting the benefits of combining blockchain with AI for secure and efficient transportation services.It has been suggested that blockchain technology could enable a decentralized and efficient two-way sharing economy, such as ride-hailing services (Chang and Chang *et al.*, 2018). Originally used in Bitcoin (Nakamoto *et al.*, 2018), Blockchain technology is a novel network technology that enables network participants to agree on a distributed, unchangeable digital record, as well as a peer-to-peer electronic payment system.

2.4 Existing system

In case of traditional carpooling systems, it is important to be aware of its key disadvantages that affect the user experience and effectiveness. Some of the key disadvantages are:

- *Limited availability:* Availability of ride-sharing options may be limited, especially in sparsely populated areas and where public transportation options are limited.
- *Scheduling inconveniences:* Aligning schedules among carpool members can be challenging, particularly when they have varying work hours or routes.

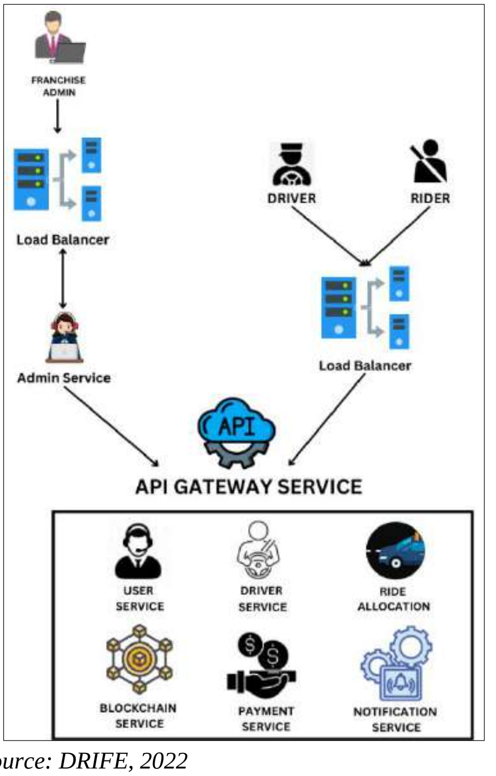
- *Lack of confidence and security issues:* Carpooling with foreigners can increase security issues, especially for women. Furthermore, a potential lack of trust among participants could result in challenges regarding the timeliness and reliability of payments.

Although there are many existing carpooling systems available, for this paper, we have chosen “DRIFE Taxi 3.0” (DRIFE, 2022) for demonstrating an existing system. DRIFE - a decentralized ride hailing service founded in India which works on the Polygon Blockchain, a high performance Ethereum-based blockchain that supports fast and low-cost transactions. Figure 3 depicts the system architecture & the features offered by DRIFE Taxi 3.0.

To overcome challenges faced by traditional ride-hailing services, DRIFE offers innovative solutions.Its features are:

- *No Commission:* One of the key features of DRIFE is that it operates on zero commission model i.e. it does not charge its drivers with any commission fees.
- *Open Governance:* Governance is achieved by participating in a unique NFT based franchise model. Entities like riders, drivers, and franchise owners can govern themselves in the most efficient and fair way.
- *Market Dictated Pricing:* It utilizes a unique dynamic pricing mechanism that is based on auctions.

Figure 3: DRIFE Taxi 3.0 – Technical Architecture



Source: DRIFE, 2022

Blockchain as a technology shine in this aspect as being transparent allows the riders and drivers can trust the pricing model.

- *Incentivized Participation:* Drivers and Riders are incentivized within the DRIFE platform for their active participation and engagement with the platform. They earn their rewards in the form of DRF token, DRIFE's own ERC-20 token

3.0 Future Scope

Blockchain can lower expenses for drivers and passengers by eliminating middlemen and their fees. This model without a centralized authority could lower the cost of carpooling by eliminating the need for platforms to impose service fees. Furthermore, token economies could be integrated by blockchain carpooling platforms, which would incentivize users by offering tokens for participating in ride sharing or inviting others to join. These tokens may be utilized as payment for upcoming rides, decreasing the dependence on regular currencies and generating fresh motivations for involvement. Users can connect with each other directly and smart contracts guarantee safe, automated payment processing after the ride, fostering trust between drivers and riders.

The transparent and trustworthy system of blockchain, characterized by its decentralized nature, records all transactions and interactions on an immutable ledger, thus decreasing the chance of fraud. Another thrilling potential lies in merging blockchain technology with self-driving cars. In the future, self-driving cars may join peer-to-peer carpooling systems, enabling passengers to schedule rides directly with autonomous vehicles via decentralized platforms. The whole process, covering route planning to payment, could be smoothly handled using smart contracts, improving convenience and effectiveness. Blockchain can also motivate environmentally friendly actions by giving rewards to individuals who participate in regular carpooling, aiding in the decrease of traffic jams and emissions.

Additionally, the transparency and security provided by blockchain can assist users in retaining ownership of their personal information. The technology guarantees that confidential data, like location and past trips, stays both private and secure, but still permits verification when needed. Blockchain has the potential to create international carpooling networks, allowing users from various countries to connect and ride together regardless of geographical barriers. This may turn out to be particularly beneficial in areas where conventional ride-sharing services are absent.

A carpooling app utilizing blockchain technology is a promising project for the future. It has the capacity to entirely revolutionize the industry, altering how people travel by leveraging the security and transparency offered by blockchain technology. In the future, the app may integrate smart contracts for enhanced functionality. Automate ride sharing, create a decentralized platform for increased security, token the platform to reward users and cover ride costs, integrate functionalities to collaborate with other blockchain apps extend to various transportation modes and lower the amount of carbon released into the atmosphere. With these advancements, the carpooling app can establish an all-encompassing transportation system that is both efficient and sustainable, thereby aiding in the progression towards a more sustainable future.

In closing, carpooling based on blockchain technology has the capability to transform the transportation sector by building trust, cutting expenses, and enhancing effectiveness. The success of it will rely on the ongoing development of blockchain technology, regulatory approval, and the widespread usage of carpooling.

4.0 Conclusion

In summary, Blockchain technology offers a cutting-edge, decentralized solution to improve the sustainability, security, and efficiency of ridesharing platforms. By utilizing smart contracts, people can create trust without the involvement of intermediaries, ensuring the safety and traceability of transactions. The secure ledger ensures that all interactions between drivers and passengers are recorded, minimizing disputes, and avoiding denial. Blockchain protects personal data and enables live tracking of trips. As technology progresses, blockchain powered carpooling has the potential to improve transportation systems by tackling accountability, trust, and fair pricing concerns, resulting in a more efficient and sustainable network. Furthermore, the integration of blockchain in carpooling platforms can foster greater user confidence, encouraging more people to participate in shared transportation solutions, ultimately reducing traffic congestion and environmental impact.

References

Aguiléra, A., & Pigalle, E. (2021). The Future and Sustainability of Carpooling Practices. An Identification of Research Challenges. *Sustainability*, *13*(21), 11824. https://doi.org/10.3390/su132111824

Jaybhaye, K., Chavan, A., Wath, A., Sherkar, O., & Kapse, S. (2024). A literature survey on carpooling. *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT),* 3(1), 743-748.

Kost, E. (2024). What caused the Uber data breach. UpGuard. https://www.upguard.com/blog/what-caused-the-uber-data-breach

Ruprah, T. S., & others. (2024, July). Peer to peer ridesharing using blockchain technology. *International Research Journal on Advanced Engineering Hub (IRJAEH), 2(7), 2021-2026.* https://irjaeh.com/index.php/journal/article/view/315

Goel, S., Sawant, S. V., & Rudra, B. (2024). Secure decentralized carpooling application using blockchain and zero knowledge proof. In *Proceedings of the International Conference on Internet of Things, Big Data and Security* (Vol. 9, pp. 260-267).

Nakamoto, S. (2008) Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf

Telecommunication Engineering Centre. (2020). Study paper on securityaspects of blockchain.https://www.tec.gov.in/pdf/Studypaper/Security%20aspects%20of%20blockchain.pdf

Taherdoost, H. (2023). Smart contracts in blockchain technology: A critical review. *Information*, 14(2), 117.

Pal, P., &Ruj, S. (2019). BlockV: A blockchain enabled peer-peer ride sharing service. In *2019 IEEE International Conference on Blockchain (Blockchain)* (pp. 463-468). https://doi.org/10.1109/Blockchain.2019.00070

Chang, S. E., & Chang, C.-Y. (2018, July). Application of blockchain technology to smart city service: A case of ridesharing. In *Proceedings of the Conference* (pp. 664-671).

DRIFE. (2022). DRIFE whitepaper. https://www.drife.io/_files/ugd/3af3d0_25392fa7ed6e401abe6fdea1fd3c4b9a.pdf

Baza, M., Lasla, N., Mahmoud, M., Srivastava, G., & Abdallah, M.M. (2019). B-Ride: Ride Sharing With Privacy-Preservation, Trust and Fair Payment Atop Public Blockchain. IEEE Transactions on Network Science and Engineering, 8, 1214-1229.

Shivers, R. M. (2019). Toward a secure and decentralized blockchain-based ride-hailing platform for autonomous vehicles (Master’s thesis, Tennessee Technological University).

Shivers, R., Rahman, M. A., Faruk, M. J. H., Shahriar, H., Cuzzocrea, A., & Clincy, V. (2021, December). Ride-hailing for autonomous vehicles: Hyperledger fabric-based secure and decentralize blockchain platform. In 2021 IEEE international conference on big data (Big data) (pp. 5450-5459). IEEE.

Valaštín, V., Košt’ál, K., Bencel, R., & Kotuliak, I. (2019). Blockchain based car-sharing platform. *2019 International Symposium ELMAR*, 5-8. https://doi.org/10.1109/ELMAR.2019.8918650

Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017). An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. 2017 IEEE International Congress on Big Data (Big Data Congress), 557-564.

Gundaboina, L., Badotra, S., Bhatia, T. K., Sharma, K., Mehmood, G., Fayaz, M., & Khan, I. U. (2022). [Retracted] Mining Cryptocurrency-Based Security Using Renewable Energy as Source. *Security and Communication Networks*, *2022*(1), 4808703.

CHAPTER 10

Enhancing Online Payment Fraud Detection: A Review of the Jaya Optimization Algorithm

Deepali Gohil and Shruti Pol***

ABSTRACT

The increasing adoption of online transactions has led to a parallel rise in fraudulent activities, emphasizing the need for reliable and robust detection mechanisms. This paper analyses the Jaya Optimization Algorithm and then demonstrates its application to develop more effective online payment fraud detection systems. With the growing complexity of fraud tactics, the traditional detection mechanism fails to keep track with such advancing techniques. This paper identifies an integration of the Jaya Optimization Algorithm with the latest fraud detection techniques to take advantage of the capability of enhancing performance metrics such as the detection rate and reduction in false positives. In addition, key findings in the paper show that the Jaya Optimization Algorithm can significantly improve the accuracy and efficiency in fraud detection models that come with the safer online payment environments. It ends with directions for future research targeted at helping to continue the resolution of some of the outstanding issues in the field.

Keywords: Online Payment Fraud; Jaya Optimization Algorithm; Machine Learning; Fraud Detection; Deep Learning.

1.0 Introduction

The proliferation of Internet digital services has changed the world of transactions in banking processing and has become the standard payment method for consumers and businesses in e-commerce. This move to electronic commerce has brought about the ease and speed of payment but has also caused a concurrent surge in online payment fraud cases. The latest reports in the sector indicate that projected global losses through fraud e-commerce are set to exceed $20 billion annually, which points to the urgency to have proper identification tools in place. The evolving nature of fraud poses significant challenges for detection mechanisms. These problems encompass the need to be nimble in the wake of highly dynamic operational situations, dealing with statistically imbalanced data where abnormal transactions are rare and also making sure that the identification tools can intervene instantaneously.

**Assistant Professor, Department of Computer Engg, D Y Patil College of Engineering, Pune, Maharashtra, India (E-mail: dmgohil@dypcoeakurdi.ac.in)*

***Student, Department of Computer Engg, D Y Patil College of Engineering, Pune, Maharashtra, India (E-mail: polshruti131415@gmail.com)*

In this framework, more advanced approaches are being developed, like the Jaya Optimization Algorithm. These have shown potential for bettering the performance of fraud detection models. This paper evaluates the Jaya Optimization Algorithm as a means of technology in the fraud detection system by emphasizing its capacity to improve accuracy, reduce false positives, and to overall model efficiency. The aim of this work is to put together in one place the remaining literature and experimental findings that have been achieved in research the role of the algorithm in the process of online payment fraud detection.

2.0 Literature Review

The fraud detection field in online payments has evolved with several strategies developed to curb fraud. These methodologies include traditional statistical techniques, machine learning algorithms, and advanced frameworks using deep learning. This has brought enhanced detection rates and lower false positives. One of the most preferred strategies in recent literature involves the Jaya Optimization Algorithm to enhance machine learning models. The integration of the Jaya Optimization Algorithm with advanced models, such as ResNeXt-GRU, enhances feature extraction and leads to better accuracy in detection. The application of data balancing techniques like SMOTE has efficiently overcome challenges related to the imbalanced nature of datasets, improving performance metrics (Almazroi *et al.*, 2023). An alternative framework with potential is the Support Vector Data Description (SVDD) combined with Polynomial Self-Learning Particle Swarm Optimization (PSLPSO), which adjusts the hyperparameters in the process of anomaly detection. This method optimizes a model for boosted performance using good feature sets while also addressing class imbalance via under sampling. However, a challenging aspect of PSLPSO is its computational complexity, which requires sensitive optimization of the SVDD parameters, making it a disadvantage (Mniai, *et al.*, 2023).

Even blockchain technology has been explored as a secure framework for fraud detection. An adaptive incentive-based machine learning model uses blockchain for real-time fraud detection, adding incentives to data contributors in exchange for data quality improvement. While this method provides significant security benefits, it faces severe computational overheads and complexities during the integration of blockchain with machine learning systems (Pranto *et al.*, 2022).

The study also examined machine learning techniques for online fraudulent transaction detection using SMOTE and AdaBoost. Model performance is enhanced when SMOTE is used for dataset balancing and AdaBoost is applied. However, it is important to note that while SMOTE is useful, it can introduce noise into the dataset, and AdaBoost requires significant computational resources, making real-time usage challenging (Ileberi, *et al.*, 2021). A biometric fingerprint system, security queries, and OTP combined with a Support Vector Machine (SVM) for fraud detection highlight the need for a more holistic approach to fraud detection. However, the scalability of the SVM model and its ability to incorporate new adaptations for detecting continuously changing fraud patterns remain problematic (Mary, *et al.*, 2021). The power of models like LSTMs lies in their ability to handle sequential data well. LSTMs have been applied to credit card transaction anomaly detection, and although effective, they tend to be computationally expensive and require large amounts of training data to perform well, which limits their widespread use (Benchaji, *et al.*, 2021).

For example, the LAW method employs an online learning procedure to automatically determine sliding window sizes for fraud detection in payment activities. This methodology is adaptive and improves detection effectiveness by considering varying transaction patterns using real bank data. Possible improvements could include expanding time windows to optimize the model further and better utilize available resources (Wang, *et al.*, 2020). It has also been suggested that real-time fraud detection for credit cards can be performed using Hidden Markov Models in combination with stream analytics. This method emphasizes behavior pattern definitions but requires continuous updates to capture changing fraud tactics, which raises the risk of false positives (Babu, *et al.*, 2016).

Hybrid frameworks utilizing big data technologies such as Hadoop, Spark, and Storm have also been explored. These frameworks use synthetic datasets to evaluate various detection algorithms, but challenges in systematic optimization and real transaction data integration persist (Dai, *et al.*, 2016).

Lastly, credit card fraud detection using genetic algorithms is a prime example of the possibilities offered by evolutionary search techniques for optimizing detection operations. While promising, the lack of practical validation with real-world transactions indicates the need for more extensive evaluations to establish its effectiveness (Santiago, *et al.*, 2015). Table 1 provides a summary of various techniques used for fraud detection in financial transactions.

Table 1: Summary of Research Contributions and Techniques in Financial Transaction Fraud Detection

Ref. No.	Title of Paper	Year and Type of Paper	Methodology Described	Limitation
[1]	Online Payment Fraud Detection Model Using Machine Learning Techniques	2021, IEEE Journal	- Jaya Optimization Algorithm Introduces ResNeXt-GRU model optimized by Jaya algorithm for fraud detection. Use of SMOTE for data balancing and ensemble learning for feature extraction using autoencoders and ResNet. Tested on three large datasets.	- Computationally intensive with DL integration. - Complex to implement and requires expertise in both DL and optimization.

[2]	A Novel Framework for Credit Card Fraud Detection	2023, IEEE Journal	- Support Vector Data Description (SVDD) for anomaly detection. - Polynomial Self-Learning Particle Swarm Optimization (PSLPSO) for optimizing SVDD hyperparameters. - Undersampling Technique for balancing the dataset. - Feature Selection Mechanism to improve model performance.	- Computational complexity due to PSLPSO. - Requires careful tuning of SVDD parameters.
[3]	Blockchain and Machine Learning for Fraud Detection: A Privacy-Preserving and Adaptive Incentive-Based Approach	2022, IEEE Transactions	- Blockchain Technology for secure data sharing. - Adaptive Incentive-Based Machine Learning for real-time fraud detection. - Incentives for data contributors.	- High computational cost due to blockchain integration. - Complex integration of blockchain with ML.
[4]	Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost	2021, IEEE Journal	- SMOTE (Synthetic Minority Over-sampling Technique) for balancing the dataset. - AdaBoost for improving the performance of various ML models (e.g., SVM, RF, LR).	- SMOTE can introduce noise. - AdaBoost can be computationally expensive.
[5]	Online Transaction Fraud Detection System	2021 IEEE Conference Paper	Utilizes fingerprint authentication, security questions, use OTP verification, and SVM for fraud detection in online transactions.	SVM model may struggle with scalability and real-time detection of highly variable fraud patterns.
[6]	Anomaly Detection in Credit Card Fraud Using LSTM Networks	2020, ACM Conference	- LSTM Networks for detecting anomalies in credit card transactions. - Focuses on using sequential data to improve detection accuracy.	- High computational cost. - Requires large amounts of data for effective training.
[7]	LAW: Learning Automatic Windows for Online Payment Fraud Detection	2020 IEEE Transaction Paper	Proposes LAW, a method that learns the most suitable sliding time window size for online payment fraud detection. The system automatically adapts to changing transaction patterns, improving detection efficiency. Experiments on real-life bank data validate the model's robustness.	The current range of time window sizes is limited. Future work includes expanding the range, improving computational efficiency, and testing on larger datasets.

[8]	Real-time credit card fraud detection using Streaming Analytics	January 2016 Conference Paper	Proposes a system using Hidden Markov Models (HMM) and streaming analytics for real-time detection and prevention of credit card fraud.	The effectiveness of the model relies heavily on accurately defining behavior patterns, which may need frequent updates to capture evolving fraud tactics, increasing the risk of false positives in new, unfamiliar fraud patterns.
[9]	Online Credit Card Fraud Detection: A Hybrid Framework with Big Data Technologies	2016 IEEE Conference Paper	Implementation of big data technologies like Hadoop, Spark, Storm, HBase, etc. In this paper a prototype is implemented, and that prototype is tested with the synthetic dataset.	The system needs better integration of more detection algorithms and other Big Data tools. It also requires testing with real transaction data and systematic optimizations of all components in the framework.
[10]	A Modeling Approach for Credit Card Fraud Detection in Electronic Payment Services	2015 ACM Conference Paper	Uses transaction history and features of buyer, seller, credit card, and holder to classify fraudulent transactions. Classification performed with SVM on real data from a Latin American payment service.	High false alarm rate and challenge in modeling seller behavior. Fraudulent behaviors change over time, requiring constant updates to the model.

*Source**:** Compiled by the authors*

3.0 Jaya Optimization Algorithm

The Jaya Algorithm is the population-based method that iteratively progresses solutions in a manner not requiring specific algorithmic parameters in the form of mutation or crossover rates. Due to this difference in its nature, the former is robust and straightforward for optimizing complex models in dynamic environments, such as fraud detection.

3.1 Why Jaya?

- *Flexible and effective:* Jaya has the ability to evolve in case there is a shift in patterns of transactions in order to maintain the performance of the model at the best possible point.
- *Optimization without parameters:* Indeed, what distinguishes Jaya from the remaining algorithms is that one does not need to add parameters to perform operation of Jaya. It is hence simple in to use.
- *Ability to optimize on a global scale:* Due to this, the model is unable to get stuck in local maxima, thus providing better hyperparameter optimization solutions.

3.2 Pseudo code for Jaya Optimization Algorithm

Below is the pseudo code for the Jaya Optimization Algorithm as applied to the classification model:

Input: Population P of solutions (hyperparameters)

Objective Function f(x)
Maximum Iterations Max_Iter
Output: Optimal Solution x_best

1. Initialize a population P of random solutions {x_1, x_2, ..., x_n}
2. Evaluate the objective function f(x) for each solution x in P
3. Find the best solution x_best and worst solution x_worst in P
4. while (iteration<Max_Iter) do
5. for each solution x_i in P do
6. Generate a new solution x_new based on:
7. x_new = x_i + r1*(x_best-|x_i|)-r2*(x_worst-|x_i|)
8. where r1, r2 are random numbers in [0,1]
9. Evaluate the objective function f(x_new)
10. if f(x_new) <f(x_i) then
11. Update x_i = x_new
12. end if
13. end for
14. Update x_best and x_worst based on new population P
15. Increment iteration count
16. end while
17. Return x_best as the optimal solution

3.3 Explanation of the Pseudo code

- *Initialization*
  - The algorithm starts by initializing a population of randomly generated solutions where every solution is a set of hyperparameters.
  - Each solution is measured by using the objective function-for example, model accuracy on the validation set.
- *Iteration and update*
  - In the case of each solution, a new solution is produced by moving it closest to the best solution and further getting away from the worst solution.
  - Evaluate the objective function at the new solution.
  - It replaces the current solution if this new solution has an objective function superior to the previous one.
- *Best and worst solutions*
  - In each generation, best and worst solutions within the population are updated.
  - This process repeats until the number of maximum iterations is attained or convergence criteria are met.
- *Return the best solution*
  - The algorithm returns the best solution found, representing the optimal set of hyperparameters for the classification model.

3.4 Integration with fraud detection models

Jaya Optimization Algorithm can also improve existing fraud detection systems for online payment schemes by optimizing the hyperparameters of several machine learning models, like ResNeXt-GRU. The chapter depicts how Jaya can be added to those fraud detection frameworks through a systematic approach toward further improving their capabilities in making correct detections:

- *Initialization:* A population of random sets of hyperparameters, according to a fraud detection model selected, is initiated using the Jaya Algorithm.
- *Performance evaluation:* For each set of hyper parameters, evaluate with regard to how well the models predict fraudulent transactions. This can be done in terms of accuracy, precision, recall, or F1-score.
- *Iterative optimization:* Use the pseudocode of Jaya to iteratively optimize the selected set of hyper parameters. Here, optimization will occur such that the resulting detection accuracy is maximized with minimal computations.
- *Deployment:* Embed the optimized model into a live fraud detection system, so that it could adapt rapidly to new data and keep the accuracy levels high in the detection of fraudulent activities.

3.5 Advantages of integrating jaya optimization algorithm in detection systems

Integrating the Jaya Optimization Algorithm offers several advantages.

- *Enhanced detection precision:* The models can fine-tune and tune the hyperparameters to produce better precision in identifying true transactions versus false transactions in such a way that it points out fewer false positives.
- *Lower computational complexity:* The Jaya algorithm converges far faster than other conventional optimization algorithms and, hence reduces the training time along with other computational resources for the design of models.
- *Robustness to concept drift:* The adaptive feature of the algorithm makes the Jaya algorithm model robust by being effective in with time when fraud patterns are changing, thus long-term viability and effectiveness.

3.6 Results and discussion

The ability of the Jaya Optimization Algorithm in the improvement of fraud detection models can be assessed with respect to comparative analyses. For though it does not point towards a particular methodology, it does highlight improvements realized through the literature available so far:

Performance Metrics: Evaluations typically focus on several key metrics:

- *Precision*: This refers to the number of transactions classified correctly, thus giving a direct measure for overall performance.
- *Accuracy*: The ratio of actually true fraud detections to all cases predicted for fraud. In other words, it’s how valid the model is about real fraud detection.

- *Recall*: This is simply the fraction of correct identification of actual fraud cases reported by the model. This will therefore enable one understand how sensitive the model is.
- *F1-Score:* F1 score is the harmonic mean of precision and the recall which gives a balanced view of how good this model is.

3.7 Key results from literature

Several works articulated the advantages of introducing Jaya optimization into the fraud detection models:

- *Superior performance:* Jaya-optimized models are constantly yielding higher accuracy compared to traditional optimization techniques. Moreover, Jaya-optimized models always save more computational time in comparison with traditional optimization techniques.
- *Faster convergence:* Compared to genetic algorithms and grid search methods, the Jaya algorithm demonstrates quicker convergence rates, which thus means enhanced capabilities in detection.
- *Adaptability*: The models optimized by Jaya are still robustly handling concept drift and perform well over imbalanced as well as balanced datasets, thus showing versatility in different operational contexts.

4.0 Conclusion

In conclusion, this paper highlights the effectiveness of the Jaya Optimization Algorithm in the development of online payment fraud detection systems. It also investigates many methods implemented for the purpose of online fraud detection. Besides, this algorithm suggests a solution to some of the most relevant queries, thanks to the fact that its parameters are optimized and its functionality is improved. The results of the study establish its capability in the accuracy of detection and the decrease of false positives, making online payments a safer environment. Future research should be focused on doing more research into the scenarios where it can be applied and also coming up with innovations that will mitigate the negative effects of the emerging trends in online payment fraud.

References

A. A. Almazroi and N. Ayub, “Online Payment fraud detection model using machine learning techniques,” *IEEE Access*, vol. 11, pp. 137188–137203, Jan. 2023, doi: 10.1109/access.2023.3339226.

A. Mniai, M. Tarik, and K. Jebari, “A novel framework for credit card fraud detection,” *IEEE Access*, vol. 11, pp. 112776–112786, Jan. 2023, doi: 10.1109/access.2023.3323842.

T. H. Pranto, K. T. A. Md. Hasib, T. Rahman, A. B. Haque, A. K. M. N. Islam, and R. M. Rahman, "Blockchain and Machine Learning for Fraud Detection: A Privacy-Preserving and Adaptive incentive based approach," *IEEE Access*, vol. 10, pp. 87115–87134, Jan. 2022, doi: 10.1109/access.2022.3198956.

E. Ileberi, Y. Sun, and Z. Wang, "Performance evaluation of machine learning methods for credit card fraud detection using SMOTE and AdaBoost," *IEEE Access*, vol. 9, pp. 165286–165294, Jan. 2021, doi: 10.1109/access.2021.3134330.

I. M. Mary, M. Priyadharsini, Karuppasamy. K, and M. Sharmila. F, "Online Transaction Fraud Detection System," *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Mar. 2021, doi: 10.1109/icacite51222.2021.9404750.

I. Benchaji, S. Douzi, B. E. Ouahidi, and J. Jaafari, "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model," *Journal of Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00541-8.

C. Wang, C. Wang, H. Zhu, and J. Cui, "LAW: Learning Automatic Windows for Online Payment Fraud Detection," *IEEE Transactions on Dependable and Secure Computing*, p. 1, Jan. 2020, doi: 10.1109/tdsc.2020.3037784.

N. R. U and B. S. Babu, "Real-time credit card fraud detection using Streaming Analytics," *Proceedings of the 2016 International Conference on Advanced Technologies for Communications and Computing ICATCCT*, Jan. 2016, doi: 10.1109/icatcct.2016.7912039.

Y. Dai, J. Yan, X. Tang, H. Zhao, and M. Guo, "Online credit card fraud detection: A hybrid framework with big data technologies," in *Proceedings of the 2016 IEEE Trust-Com/BigDataSE/ISPA*, 2016.

G. P. Santiago, A. C. M. Pereira, and R. Hirata Jr., "A modeling approach for credit card fraud detection in electronic payment services," Universo Online Inc., São Paulo, Brazil, 2023. [Online]. Available: gabrielpreti@gmail.com, adrianoc@dcc.ufmg.br, hirata@ime.usp.br.

CHAPTER 11

Multi-class Malware Detection using Modified GNN and Explainable AI

Shubham Kumar, Vishwatej Khot**,*
*Sagar Bhat***, Atharva Ghare**** and Reshma Kapadi******

ABSTRACT

As malware becomes increasingly sophisticated, traditional detection methods, such as signature-based systems, face challenges in effectively identifying new and polymorphic threats. This research aims to develop a multi-class malware detection system using a modified Graph Neural Network (GNN) architecture. By converting byte and assembly (ASM) files into image formats, this approach improves feature extraction and enhances the representation of malware characteristics. The methodology takes advantage of GNNs to capture complex relationships within the data, providing a dynamic and adaptive framework for detection. Additionally, Explainable AI (XAI) techniques, including Grad-CAM, will be utilized to offer transparency by enabling visual insights into the model's decision-making process. The objective is to build a malware detection system that delivers high detection accuracy while ensuring interpretability, making it both dependable and understandable for cybersecurity applications

Keywords: Malware Detection; Graph Neural Networks (GNNs); Polymorphic Malware; Byte Files; Assembly (ASM) Files; Image Transformation; Feature Extraction; Explainable AI (XAI); Grad-CAM.

1.0 Introduction

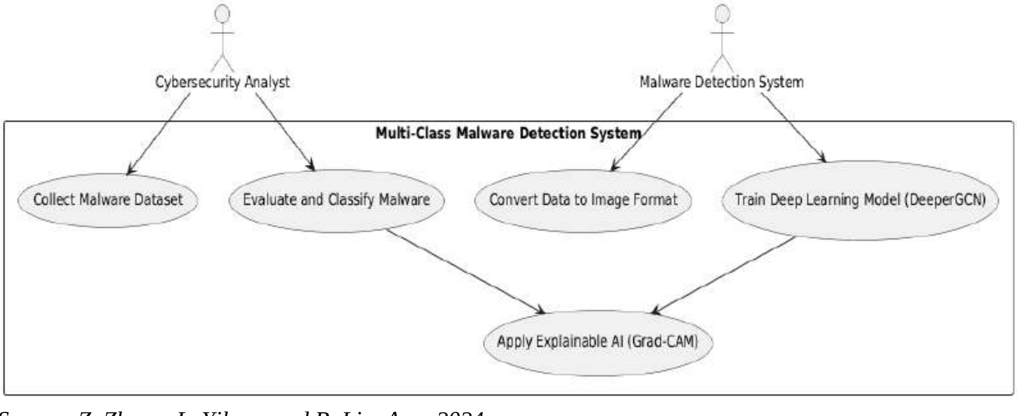
The rapid rise of sophisticated and adaptive malware has created significant challenges for cybersecurity professionals worldwide. Traditional malware detection techniques, such as signature-based systems, are becoming increasingly inadequate as they are primarily designed to identify known threats.

**Corresponding author; B.E, Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune, Maharashtra, India*
(E-mail: connect2ershubham@gmail.com)
***B.E, Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune, Maharashtra, India (E-mail: khotvishwatej@gmail.com)*
****B.E, Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune, Maharashtra, India (E-mail: sagarbhat.bhat.ss@gmail.com)*
*****B.E, Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune, Maharashtra, India (E-mail: atharvaghare2021.comp@mmcoe.edu.in)*
******B.E, Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune, Maharashtra, India (E-mail: reshmakapadi@mmcoe.edu.in)*

These methods rely on predefined malware signatures, which can be easily circumvented by modern malware that utilizes polymorphic and metamorphic techniques to alter its code structure. As a result, these traditional approaches often fail to detect novel or evolving malware strains.

In recent years, deep learning has gained significant attention for malware detection due to its capability to automatically learn complex patterns from large datasets. Among these techniques, Graph Neural Networks (GNNs) have shown great promise for analyzing malware, as they can effectively capture the relationships between different components of malware files, such as byte and assembly (ASM) data. By transforming these files into graph-based representations, GNNs are able to identify complex dependencies that might be missed by traditional detection methods.

Figure 1: Vanilla Implementation of Multi-Class Malware Detection System



Source: Z. Zhang, L. Yilmaz and B. Liu, Aug. 2024

Figure 1 depicts the Multi-Class Malware Detection System, which integrates human expertise with automated processes for effective malware analysis. A cybersecurity analyst begins by collecting a malware dataset and assessing its characteristics. The system then evaluates and classifies the malware while converting the data into an image format for further processing. The converted data is used to train a deep learning model, specifically the DeeperGCN, for precise classification. Finally, the system applies Explainable AI techniques, such as Grad-CAM, to generate interpretable insights, providing a clear understanding of the model's decisions.

Although deep learning models can achieve high accuracy, their decision-making processes are often unclear, making it challenging for users to understand how a specific prediction was made. This lack of transparency is particularly problematic in cybersecurity, where trust and accountability are essential.

This research seeks to address these challenges by developing a multi-class malware detection system that combines the strengths of Graph Neural Networks (GNNs) with Explainable AI (XAI) techniques. Specifically, we propose using DeeperGCN, a modified GNN architecture, to analyze malware represented as images derived from byte and ASM files. To enhance interpretability, we aim to incorporate Grad-CAM, a visualization tool that identifies the parts of the input data most influential in the model's decision. This approach will enable cybersecurity professionals to better understand why specific files are flagged as malicious, increasing the system's transparency and trustworthiness.

1.1 Objectives of the study

Develop a multi-class malware detection system: Create a robust detection framework utilizing Graph Neural Networks (GNNs) to identify and classify various types of malwares.

Leverage image-based representations of malware: Transform byte and assembly (ASM) files into image formats for enhanced feature extraction and improved analysis of malware characteristics.

Integrate explainable AI (XAI) techniques: Incorporate Grad-CAM to provide visual interpretations of the model's decision-making, ensuring transparency and user trust.

Address limitations of traditional detection methods: Overcome the inadequacies of signature-based systems by employing a dynamic and adaptive detection approach suitable for polymorphic and evolving malware strains.

Enhance cybersecurity trust and accountability: Develop an interpretable system that aids cybersecurity professionals in understanding and verifying the detection process, fostering greater trust in AI-based solutions.

2.0 Review of Literature

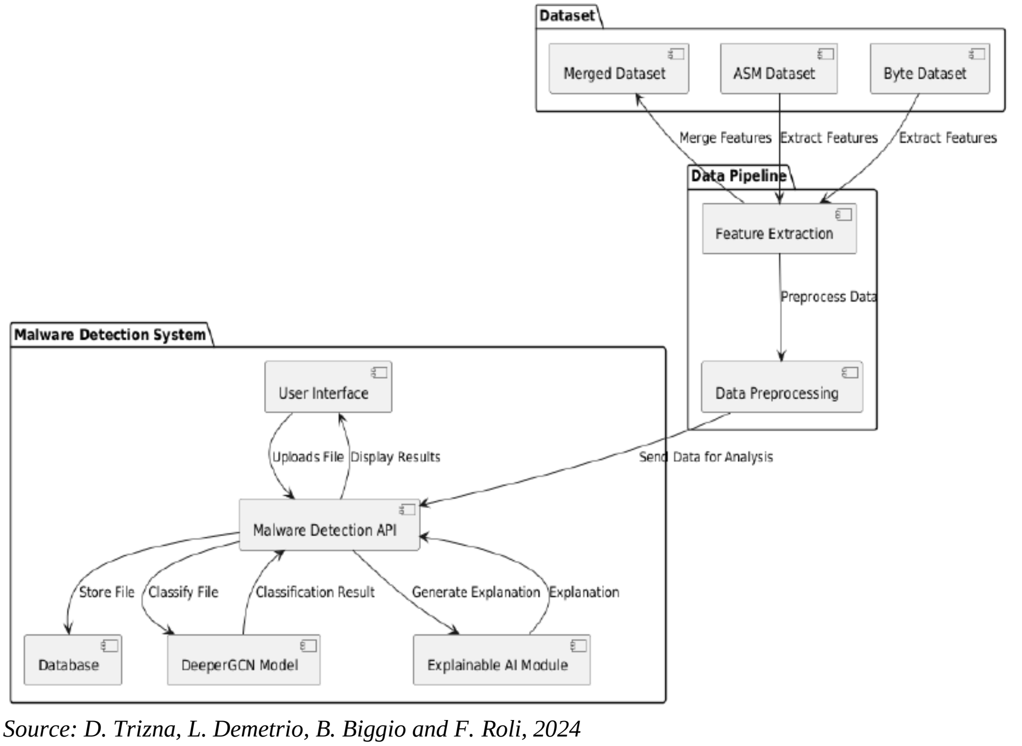
The reviewed papers explore cutting-edge methods for improving AI interpretability, especially in security applications. Zhang *et al.* (2020) emphasize Inductive Logic Programming (ILP) for explainable AI, highlighting its logical transparency and challenges with scalability and noise sensitivity. Yilmaz and Liu (2020) combine Statistical Relational Learning (SRL) with Neural-Symbolic AI to improve relational data handling and probabilistic reasoning, though scalability remains a challenge. Zebin *et al.* (2022) propose a Random Forest model for detecting DNS over HTTPS (DoH) attacks, leveraging SHAP values for transparency, achieving 99.91% precision. Ali *et al.* (2023) introduce a multitask LSTM model for IoT malware detection, achieving 95.83% accuracy but struggling with dataset imbalance. Lai *et al.* (2023) present NNV is Builder for visual analytics in neural networks, enabling model transparency but limited by performance on large datasets. Esmaeili *et al.* (2022) enhance IIoT security with stateful query analysis, achieving 93.1% adversarial attack detection. Lastly, Trizna *et al.* (2024) develop Nebula, a Transformer-based malware analysis model integrating explainable AI for improved dynamic malware detection.

3.0 Research Methodology

3.1 System architecture diagram

The proposed Malware Detection System integrates various components to automate the process of malware detection and explanation. The system is designed to handle data preprocessing, feature extraction, file classification, and the generation of human-readable explanations. It consists of three primary subsystems: the Dataset, the Data Pipeline, and the Malware Detection System.

Figure 2: System Architecture



Source: D. Trizna, L. Demetrio, B. Biggio and F. Roli, 2024

Figure 2 showcases the architecture of the Malware Detection System, which automates the identification and explanation of malware. The system incorporates three key components: the Dataset, the Data Pipeline, and the Malware Detection System. The Dataset section integrates the ASM and Byte datasets, combining and extracting relevant features. The Data Pipeline handles data preprocessing and feature extraction to prepare the data for analysis. Finally, the Malware Detection System includes a user interface, a malware detection API, a database, a DeeperGCN model for classification, and an explainable AI module to generate understandable insights. These components work together to ensure efficient malware detection and detailed explanations.

3.2 Dataset

The system utilizes two types of datasets:

- The ASM Dataset contains features derived from the assembly code of executable files, providing a structural view of the binary's operations.
- The Byte Dataset includes raw binary data that represents the file at the byte level.

These two datasets are merged to form a unified Merged Dataset. The merging process ensures that both structural and raw data features are captured, increasing the model's ability to classify files accurately.

3.3 Data pipeline

The Data Pipeline is responsible for processing the datasets before they are sent for classification. This pipeline consists of two primary stages:

- *Feature extraction:* In this stage, relevant features are extracted from both the ASM and Byte datasets. These features are vital for accurately training the machine learning models.
- *Data preparation:* After feature extraction, the data is subjected to preprocessing steps, including cleaning, normalization, and transformation. These steps ensure the data is structured appropriately for analysis before being forwarded to the malware detection system for evaluation.

3.4 Malware identification framework

The core of the system is the Malware Detection System, which comprises multiple components working in conjunction to detect and explain malware. This system includes the following:

User Interface (UI): The user interface allows users to upload files for malware detection. It is designed to be intuitive, enabling easy interaction with the system.

Malware Detection API: Acting as the central communication hub, the API facilitates the flow of data between the user interface, machine learning models, and other components of the system. It receives the uploaded files, forwards them for classification, and returns the results to the UI.

DeeperGCN Model: The file classification process is handled by the Deeper Graph Convolutional Network (DeeperGCN), a sophisticated deep learning model. This model is designed to detect patterns and relationships within the data that signify whether a file is benign or malicious. The classification result is returned to the API for further processing.

The system incorporates an Explainable AI module to ensure transparency in its decision-making process. This module generates detailed explanations for each classification result, clarifying the reasoning behind the decision. By providing insights into how a file is classified as either malware or benign, this feature fosters trust in the system and helps users understand the rationale behind the outcomes.

Database: The system's Database component stores relevant data, including uploaded files, classification results, and generated explanations. This allows for future reference and further analysis if required.

3.5 Workflow

The system operates as follows:

- A user uploads a file via the User Interface.
- The Malware Detection API sends the file to the DeeperGCN Model for classification.
- The DeeperGCN Model returns a classification result (malware or benign) to the API, which is then displayed to the user.
- If an explanation is requested, the Explainable AI Module generates a detailed explanation, which is also returned to the API and displayed to the user.
- All relevant data, including the uploaded file and its classification result, is stored in the Database for future use.

4.0 Applications

Applications in cybersecurity: The Multi-Class Malware Detection System, leveraging modified GNNs and Explainable AI, provides diverse applications in enhancing cybersecurity measures:

Enterprise security: Integrated into corporate networks for real-time malware monitoring and classification, enhancing threat response.

Incident response: Provides clear explanations for flagged files, aiding forensic investigations and understanding attack origins.

Cloud security: Detects malware in cloud environments, ensuring continuous protection and immediate threat identification.

Regulatory compliance: Facilitates adherence to cybersecurity regulations by generating detailed detection reports for audits.

Research and development: Assists researchers in studying malware behaviour and developing new detection techniques through its explainability features.

5.0 Advantages

The architecture of the Multi-Class Malware Detection System offers several key benefits that enhance detection accuracy, transparency, and usability:

Enhanced detection accuracy: By utilizing modified GNNs, the system can effectively identify a wide range of malware types, improving overall threat detection capabilities.

Promoting interpretability and confidence: The inclusion of Explainable AI methods, such as Grad-CAM, empowers users to gain a deeper understanding of how the model arrives at its decisions. This enhances trust in the system and supports faster, well-informed threat management.

Robust data handling: The system processes both byte and ASM files, converting them into image formats for effective analysis, which enhances its adaptability to various data types.

Scalability: The architecture is designed to scale effectively, accommodating both small and large datasets, making it suitable for organizations of varying sizes and needs.

Support for cybersecurity research: The system provides valuable insights into malware behaviour, benefiting researchers and organizations aiming to develop more effective security solutions.

6.0 Disadvantages

While the Multi-Class Malware Detection System presents several advantages, it also has some challenges and limitations that should be considered:

Complexity of implementation: The integration of Graph Neural Networks and Explainable AI techniques introduces a level of complexity that can complicate development and deployment. Ensuring that all components work seamlessly together requires careful planning and expertise.

High computational demands: The deep learning models used in the system necessitate significant computational resources. This can lead to increased costs for hardware and longer training times, especially when working with large datasets.

Reliance on data quality: The performance of the system significantly depends on the quality and diversity of the training datasets. Insufficient or skewed data can compromise the system’s ability to detect and classify malware accurately, making it vulnerable to undetected threats, including newly emerging or altered malware variants.

Challenges with real-time processing: While the system is designed for efficient operation, the intricate nature of its models and data handling processes may result in delays. This latency could impact the system’s ability to provide immediate responses during active cyber threats.

Interpretability for non-technical users: While Explainable AI enhances transparency, understanding the model's reasoning can still be challenging for users without a technical background. This could lead to confusion when interpreting classification results.

7.0 Conclusion

In conclusion, the development of the Multi-Class Malware Detection System utilising modified Graph Neural Networks and Explainable AI represents a significant step forward in the fight against evolving cybersecurity threats. By leveraging advanced machine learning techniques, the system enhances the accuracy and efficiency of malware detection, making it capable of identifying various types of malwares more effectively.

The Role of Explainable AI: Integrating Explainable AI offers meaningful clarity into the model's decision-making process, enhancing user confidence and making it easier to interpret classification results. This level of transparency is essential for cybersecurity experts, enabling them to respond promptly and effectively to potential threats.

Furthermore, the system's reliance on high-quality data and advanced algorithms ensures that it remains adaptable to the constantly changing landscape of malware. As we move forward, the continuous refinement of the system will be essential to addressing new challenges and improving its effectiveness.

Overall, the Multi-Class Malware Detection System not only enhances detection capabilities but also promotes a deeper understanding of malware behaviour, equipping organizations with the tools necessary to protect their digital environments more effectively.

References

Z. Zhang, L. Yilmaz and B. Liu, "A Critical Review of Inductive Logic Programming Techniques for Explainable AI," in IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 8, pp. 10220-10236, Aug. 2024, doi: 10.1109/TNNLS.2023.3246980. https://ieeexplore.ieee.org/document/10092808

D. Vlahek and D. Mongus, "An Efficient Iterative Approach to Explainable Feature Learning," in IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 5, pp. 2606-2618, May 2023, doi: 10.1109/TNNLS.2021.3107049. https://iee explore.ieee.org/abstract/document/9528915

T. Zebin, S. Rezvy and Y. Luo, "An Explainable AI-Based Intrusion Detection System for DNS Over HTTPS (DoH) Attacks," in IEEE Transactions on Information Forensics and Security, vol. 17, pp. 2339-2349, 2022, doi: 10.1109/TIFS.2022.31 83390. https://ieeexplore.ieee.org/document/9796558

S. Ali, O. Abusabha, F. Ali, M. Imran and T. Abuhmed, "Effective Multitask Deep Learning for IoT Malware Detection and Identification Using Behavioral Traffic Analysis," in IEEE Transactions on Network and Service Management, vol. 20, no. 2, pp. 1199-1209, June 2023, doi: 10.1109/TNSM.2022.3200741. https://ieeexplore.ieee.org/document/9865131

S. Lai, W. Luan and J. Tao, "Explore Your Network in Minutes: A Rapid Prototyping Toolkit for Understanding Neural Networks with Visual Analytics," in IEEE Transactions on Visualization and Computer Graphics, vol. 30, no. 1, pp. 683-693, Jan. 2024, doi: 10.1109/TVCG.2023.3326575. https://ieeexplore.ieee.org/document/10308640

B. Esmaeili, A. Azmoodeh, A. Dehghantanha, H. Karimipour, B. Zolfaghari and M. Hammoudeh, "IIoT Deep Malware Threat Hunting: From Adversarial Example Detection to Adversarial Scenario Detection," in IEEE Transactions on Industrial Informatics, vol. 18, no. 12, pp. 8477-8486, Dec. 2022, doi: 10.1109/TII.2022.3167672. https://ieeexplore.ieee.org/document/9760120

D. Trizna, L. Demetrio, B. Biggio and F. Roli, "Nebula: Self-Attention for Dynamic Malware Analysis," in IEEE Transactions on Information Forensics and Security, vol. 19, pp. 6155-6167, 2024, doi: 10.1109/TIFS.2024.3409083. https://ieeexplore.ieee.org/document/10551436

S. Zillien and S. Wendzel, "Weaknesses of Popular and Recent Covert Channel Detection Methods and a Remedy," in IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 6, pp. 5156-5167, Nov.-Dec. 2023, doi: 10.1109/TDSC.2023.3241451. https://ieeexplore.ieee.org/document/10034794

CHAPTER 12

Smart Child Safety Alert System: An IoT-based Emergency Response and Location Tracking Solution for School Children

Amruta Aphale, Pranav Waghmare**, Sanket Chaudhari*** and Amey Burgul*****

ABSTRACT

The Smart Child Safety Alert System is an innovative Internet of Things (IoT) solution designed to enhance the safety and well-being of schoolchildren. This system integrates advanced location tracking, health monitoring, and real-time emergency response features, providing a comprehensive safety net. The core component is a child-friendly wearable device equipped with biometric sensors, fall detection technology, geofencing capabilities, and a discreet panic button for manual alerts. It automatically detects emergencies such as sudden falls, abnormal heart rate patterns, or boundary violations, triggering instant notifications to parents, school administrators, and emergency services. To ensure seamless communication, the system leverages GPS, 4G/LTE networks, Wi-Fi triangulation, and indoor positioning beacons, guaranteeing accurate and reliable connectivity even in challenging environments. Privacy and data security are paramount, with features like end-to-end encryption, multi-factor authentication, secure data storage, and limited location sharing. Field testing in multiple schools demonstrated exceptional accuracy, reliability, and user satisfaction, achieving a 99.9% alert trigger accuracy and zero false positives. This scalable solution not only enhances child safety but also establishes a robust framework for creating safer educational environments, aligning with ethical and privacy standards.

Keywords: Child safety; Internet of Things (IOT); Wearable device; Emergency response; Location tracking; Health monitoring; Privacy and security.

1.0 Introduction

1.1 Background

The safety of children in educational settings is a critical concern for parents, educators, and policymakers worldwide.

**Corresponding author; Faculty, Department of Computer Engineering, SCTR’s Pune Institute of Computer Engineering, Maharashtra, India (E-mail: aphaleamruta@gmail.com)*
***PG Student, Department of Computer Engineering, SCTR's Pune Institute of Computer Technology, Pune, Maharashtra, India (E-mail: wpranav35@gmail.com)*
****PG Student, Department of Computer Engineering, SCTR's Pune Institute of Computer Technology, Pune, Maharashtra, India (E-mail: sanketchaudhari882001@gmail.com)*
*****PG Student, Department of Computer Engineering, SCTR's Pune Institute of Computer Technology, Pune, Maharashtra, India (E-mail: amey.burgul@gmail.com)*

Schools, traditionally considered safe spaces, are increasingly facing challenges that necessitate advanced safety protocols and solutions. Despite the implementation of various security measures, such as CCTV cameras, security personnel, and access control systems, incidents of emergencies and mishaps continue to occur. These range from health-related emergencies to security breaches, natural disasters, and even acts of violence. (Benisha & Prabu, 2021)

As the complexities of modern safety challenges grow, the need for innovative solutions to ensure the well-being of children has become evident. Traditional safety systems, while effective to some extent, often lack the capabilities to provide immediate and personalized responses during emergencies. The absence of real-time location tracking, which is crucial for identifying the precise whereabouts of children in distress, is a significant shortfall in these systems (Egodawela, S. M. K. C. S. B., *et al.* (2020)). Consequently, there is a pressing need for a more holistic approach to child safety, incorporating advanced technology and robust communication systems to address these gaps effectively.

1.2 Problem statement

Existing school safety systems face several critical limitations that hinder their effectiveness in ensuring the safety of children. Among these limitations, delayed response times during emergencies are particularly concerning. In crisis situations, even a few seconds of delay can have life-altering consequences. The inability to promptly detect and respond to emergencies—whether they involve health issues, accidents, or security threats—highlights the inadequacies of current systems. (Greenfield, 2023). Another major issue is the lack of precise location tracking. In large educational institutions or during outdoor activities, the inability to pinpoint the exact location of a child in need can significantly impede rescue efforts. Furthermore, the limited communication and coordination among stakeholders—such as parents, school administrators, and emergency responders—exacerbate the challenges during critical situations. Effective crisis management requires seamless information exchange, which is often lacking in traditional safety systems.

Privacy concerns also pose a significant challenge in the context of child monitoring. Many existing systems do not prioritize the secure handling of personal data, leading to potential breaches of sensitive information. This is a critical issue, as any compromise in the privacy of children's data can have far-reaching consequences. Additionally, current systems often fail to integrate health monitoring features, which are essential for detecting and responding to medical emergencies such as abnormal heart rates, falls, or other health anomalies.

The shortcomings of traditional systems underline the need for an innovative solution that can address these challenges comprehensively. A well-rounded approach that combines real-time location tracking, secure communication, and health monitoring, while prioritizing data privacy, is essential to safeguarding children in educational settings.

1.3 Research objectives

The primary objective of this research is to design and implement a child-friendly wearable safety device that addresses the limitations of current school safety systems. The proposed solution aims to enhance the safety of children in educational environments by integrating advanced features and technologies that ensure immediate and effective responses during emergencies.

Key objectives include:

- *Developing a multi-modal alert system:* The wearable device will be equipped with a reliable multi-modal alert system capable of detecting various emergency situations (Ghashim & Arshad, 2023). This system will trigger appropriate responses based on the nature of the emergency, whether it involves a health-related issue, a security threat, or any other critical situation. By incorporating multiple alert mechanisms, the device aims to ensure that no emergency goes unnoticed.
- *Implementing secure location-sharing mechanisms:* One of the core features of the device will be its ability to share the child's location accurately and securely. The location-sharing mechanism will be designed to prioritize privacy, ensuring that sensitive data is protected while enabling stakeholders to locate the child during emergencies (Hassan, Yau, & Wu, 2019).
- *Incorporating health monitoring capabilities:* To address the need for prompt medical responses, the device will include health monitoring features. These will detect and report abnormalities such as irregular heart rates, falls, or other indicators of health issues. The integration of such capabilities is crucial for responding to health-related emergencies in a timely manner.
- *Establishing robust communication protocols:* Effective communication between stakeholders is essential during emergencies. The research will focus on developing a robust communication protocol that facilitates seamless information exchange among parents, school authorities, and emergency responders. This will ensure efficient coordination and prompt action during crises.
- *Ensuring data privacy and security:* A critical aspect of the proposed solution is the implementation of privacy-preserving measures to protect the personal data of children. The device will adhere to stringent data security standards, ensuring that all information is securely stored and transmitted.

Through these objectives, the study aims to create a comprehensive safety solution that not only addresses the shortcomings of existing systems but also sets new benchmarks for child safety in educational settings. The proposed wearable safety device will serve as a proactive tool for safeguarding children, combining advanced technology with a child-centric design to ensure usability and effectiveness.

2.0 Literature Review

2.1 Current state of school safety technology

Traditional child safety systems primarily relied on GPS trackers, CCTV cameras, and manual monitoring methods. These systems, although effective to a certain degree, are limited in their ability to provide real-time feedback, contextual awareness, or proactive responses during emergencies. The rise of IoT technologies has transformed the landscape of child safety by introducing intelligent, wearable devices and monitoring systems capable of real-time tracking, health monitoring, and emergency alerts.

Modern IoT-enabled safety systems integrate advanced sensors, such as biomedical sensors and accelerometers, to track children's location, movements, and vital health parameters (Heng *et al.*, 2021) These systems have demonstrated the potential to address challenges like delayed responses to emergencies and lack of parental engagement during school hours (Shuvo *et al.*, 2023). However, issues such as device durability, battery life, and user-friendliness remain significant barriers to widespread adoption (Benisha & Prabu, 2021). Additionally, cost-effective solutions for large-scale implementation in schools and public spaces are still under research (Heng *et al.*, 2021).

Recent studies have focused on developing multi-functional safety systems that combine geofencing, real-time notifications, and health monitoring in a single device, enabling a more holistic approach to child safety (Nandhini & Moorthi, 2023). The evolution of child safety technology reflects a shift toward ensuring both the physical and emotional security of children, empowering parents and guardians with actionable insights and timely alerts (Benisha & Prabu, 2021).

2.2 IoT applications in child safety

IoT technologies have revolutionized child safety by enabling a wide range of applications that integrate real-time monitoring, intelligent analytics, and emergency response mechanisms. These applications are designed to address the diverse needs of parents, guardians, and institutions in safeguarding children.

- *Wearable devices:* IoT-enabled wearables have emerged as a core application in child safety. These devices, embedded with GPS (Sundarajoo *et al.*, 2022) modules, accelerometers, and biomedical sensors, monitor a child's location, activity, and health parameters. They provide real-time alerts to guardians during emergencies, ensuring timely responses (IoT-Based Safety Gadgets for Child Safety Monitoring and Notification, 2023).
- *Geofencing:* Geofencing allows parents or guardians to define virtual boundaries around safe zones. If a child exits the predefined area, the system sends an immediate alert to the guardian. This application has proven to be particularly effective in school environments and public spaces, where tracking children in real time is critical (IoT-Based Unified Child Monitoring and Security System, 2022).
- *Integration of health and location monitoring:* IoT systems now combine health monitoring with location tracking to provide a comprehensive safety solution. By analyzing health metrics such as heart rate and temperature alongside GPS data,

these systems ensure both physical and environmental awareness for the safety of children (Heng & Kamsin, 2021).

- *Emergency communication systems:* IoT applications in child safety also include real-time communication systems that facilitate direct interaction between the child and guardians or emergency responders. For instance, wearable IoT devices are now equipped with panic buttons that send SOS alerts along with the child's location to pre-registered contacts (Egodawela *et al.* (2020)).

These applications highlight the potential of IoT to address child safety challenges holistically, ensuring proactive responses and minimizing risks in critical situations

2.3 Privacy and security considerations

The widespread adoption of IoT-based child safety systems has introduced significant concerns related to data privacy and security. These systems collect sensitive information, including real-time location, health metrics, and behavioral patterns, making them vulnerable to unauthorized access and misuse. Addressing these challenges is critical to ensuring trust in IoT-enabled child safety applications.

- *Data breaches and unauthorized access:* Studies have highlighted the risks of data breaches, where hackers can exploit vulnerabilities in IoT devices to access sensitive child data. Robust encryption techniques and secure authentication protocols are essential to mitigate these risks (Heng & Kamsin, 2021).
- *Compliance with privacy regulations:* Adherence to global data protection standards, such as the General Data Protection Regulation (GDPR) and the Children's Online Privacy Protection Act (COPPA) (Majib & Perera, 2019)., is necessary to safeguard user rights. IoT systems must incorporate compliance measures to ensure lawful and ethical data handling (Rehman, Rehman, Khan, Moiz, & Hasan, 2016).
- *Anonymized data processing:* Privacy-preserving techniques, such as data anonymization and pseudonymization, are being integrated into IoT systems to protect sensitive information while maintaining functionality. These methods minimize the risk of identity theft and unauthorized tracking (Thamaraimanalan, Pathmavasan, Pradeep, Praveen, & Srija, 2023).
- *Threats from smart toys and IoT devices:* Research has shown that smart toys and other child-centric IoT devices often lack adequate security measures, making them susceptible to cyberattacks. A taxonomy of threats has been developed to guide the design of more secure systems (Shasha, *et al.*, 2018).

Ensuring privacy and security in IoT-based child safety systems requires a multi-faceted approach, combining technical solutions, regulatory compliance, and user awareness. These measures are pivotal to fostering trust among users and ensuring the long-term viability of such technologies.

2.4 Performance evaluation and field testing

Evaluating the performance of IoT-based child safety systems in real-world conditions is critical to ensuring their reliability, usability, and scalability. Key aspects of performance evaluation include alert accuracy, hardware durability, and user experience.

- *Alert Accuracy:* IoT systems must minimize false positives while ensuring timely notifications during emergencies. High false positive rates can erode user trust and lead to complacency. Recent studies have demonstrated advanced machine learning techniques to improve the accuracy of emergency alerts in IoT systems (Shenbagalakshmi, 2022).
- *Hardware Durability:* Devices used in child safety systems must withstand diverse environmental conditions, such as outdoor exposure, temperature variations, and physical impacts. Field tests of wearable IoT devices have shown promising results in terms of durability and functionality across different conditions (Tunggadewi, Agustin, & Yunardi, 2021).
- *User Experience:* Intuitive interfaces and ergonomic designs are essential for user adoption, especially among parents and children. Studies have evaluated feedback from users to optimize the design of IoT devices, ensuring ease of use and better adoption rates (Greenfield, 2023).
- *Scalability in Diverse Settings:* The scalability of IoT systems is crucial for their deployment in large-scale settings, such as schools or public spaces. Performance evaluation in pilot programs has revealed insights into the system's adaptability and effectiveness in different environments (Ghashim & Arshad, 2023).

Performance evaluations and field testing not only validate the technical capabilities of IoT-based child safety systems but also provide crucial feedback for iterative improvements, ensuring their effectiveness in diverse operational scenarios.

2.5 Communication technologies in IoT systems

Effective communication protocols are the backbone of IoT-based child safety systems, enabling seamless data exchange and real-time responsiveness. These protocols are designed to ensure reliability, low latency, and energy efficiency, which are critical for child safety applications.

- *Bluetooth Low Energy (BLE):* BLE is widely adopted in wearable devices for short-range, energy-efficient communication. Its low power consumption makes it ideal for devices worn by children, ensuring longer battery life while maintaining consistent connectivity (S. Shunmuga Sundari *et al.*, 2023).
- *GSM/GPRS:* GSM and GPRS technologies are commonly used for long-range communication in IoT-based safety systems. These protocols enable location tracking and emergency alerts even in remote areas, ensuring reliable coverage across diverse environments (Madhuri & Gill, 2023).

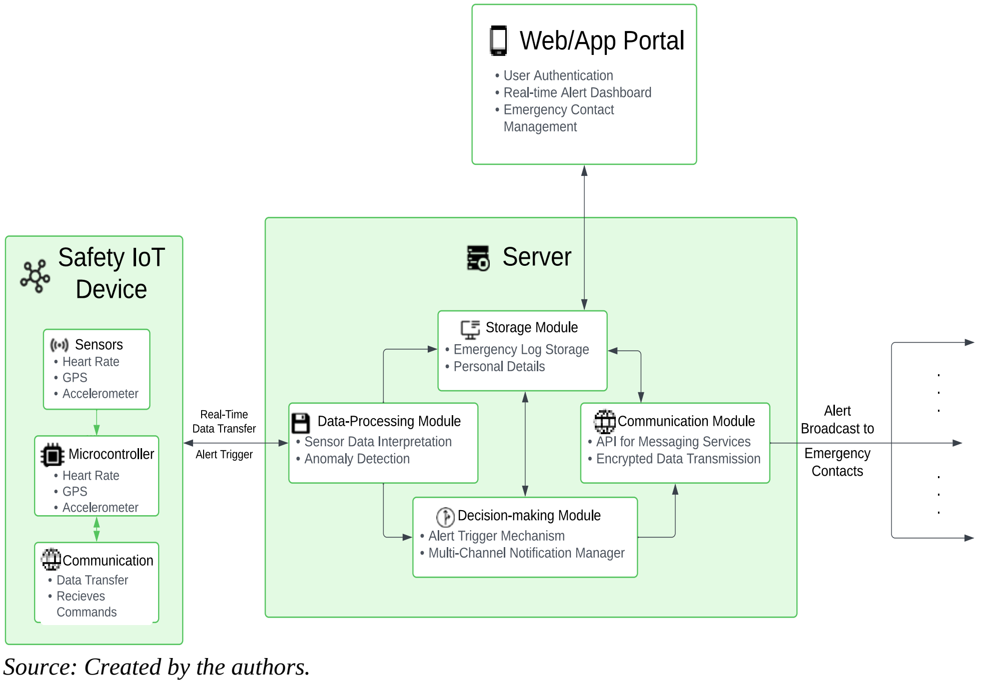
- *Low-Power Wide-Area Networks (LPWAN):* Protocols such as LoRa and Zigbee are favored for their ability to provide long-range communication with minimal power consumption. These technologies are particularly suited for large-scale deployments in schools or public spaces where multiple devices need to be connected efficiently (Islam, Jamil, Pranto, Das, Amin, & Khan, 2024).
- *5G and Edge Computing:* The advent of 5G networks has significantly improved IoT system responsiveness by reducing latency. Combined with edge computing, which processes data closer to the source, these technologies enhance the speed and reliability of emergency notifications in child safety systems (Hassan, Yau, & Wu, 2019).

The combination of these communication technologies ensures that IoT-based child safety systems can operate effectively in diverse scenarios, from individual wearables to large-scale institutional deployments, addressing the unique needs of different stakeholders.

3.0 Methodology

3.1 System Architecture Design

Figure 1: System Architecture



Source: Created by the authors.

The system architecture is designed to ensure seamless integration between hardware sensors, software modules, and communication interfaces. It follows a modular structure, allowing easy maintenance, upgrades, and scalability. The architecture focuses on real-time data acquisition, efficient processing, and reliable alert triggeringsimilar to frameworks used in prior studies (Greenfield, 2023).

3.1.1 Hardware components

The hardware components form the backbone of the system, ensuring real-time data capture and accurate readings. Key components include:

- *GPS Module:* Provides continuous location tracking. The GPS module ensures accurate positioning data with a precision of 2-5 meters under optimal conditions comparable to other IoT-based safety solutions (Benisha & Prabu, 2021). It communicates with the system through standard protocols such as NMEA and GPSD.
- *Heart Rate Sensor:* Monitors the user's heart rate in real-time. The sensor is calibrated to detect anomalies and thresholds, such as sudden spikes or drops in heart rate, which may indicate emergencies like panic or health issues.
- *Accelerometer:* Detects motion and falls. The 3-axis accelerometer measures changes in velocity and orientation, triggering alerts if abnormal movement patterns are detected. It plays a crucial role in identifying potential accidents or falls.
- *Microcontroller:* Acts as the central processing unit for hardware data. It processes sensor inputs and communicates with the software layer for decision-making. Low-power microcontrollers ensure energy efficiency, extending the device's battery life.

3.1.2 Software architecture

The software architecture is designed for modularity and flexibility. It consists of the following layers:

- *Data acquisition layer:* This layer interfaces directly with hardware sensors, collecting real-time data from the GPS, heart rate monitor, and accelerometer. Data is pre-processed to remove noise and ensure accuracy.
- *Processing layer:* Implements algorithms for data analysis, anomaly detection, and threshold monitoring. For example, heart rate data is compared against predefined safe ranges, while accelerometer data is analyzed for fall detection.
- *Decision-making layer:* Utilizes decision trees and machine learning models to evaluate sensor data following methodologies outlined by Srinivasan *et al.* (Greenfield, M. (2023).). If a critical event is detected (e.g., sudden heart rate drop), the system triggers the alert mechanism. This layer is also responsible for reducing false positives by cross-verifying data from multiple sensors.
- *Communication layer:* Handles data transmission to emergency contacts and authorities. It supports multiple communication protocols, such as SMS, email,

and app notifications, ensuring alerts reach the intended recipients even if one method fails.

- *User interface layer:* Provides a front-end for configuration and monitoring. Users can set emergency contact details, view system status, and configure alert thresholds through a mobile app or web portal.

3.2 Implementation approach

The system follows a phased implementation strategy to ensure each component works seamlessly before full integration. The approach includes development, integration, and validation phases.

3.2.1 Alert mechanisms

The alert mechanisms are designed to operate automatically and manually, ensuring flexibility and reliability:

- *Automatic alerts:* Triggered by sensor data anomalies. For instance, if the heart rate sensor detects a sudden drop, or if the accelerometer detects a fall, the system immediately activates the alert protocol.
- *Manual alerts:* Users can trigger alerts manually through a panic button on the device or the mobile app. This is particularly useful in situations where the user is conscious but unable to call for help through conventional means.
- *Multi-channel notification:* Alerts are sent via multiple channels, including SMS, email, and in-app notifications. This redundancy ensures that emergency contacts receive the alert even if one communication method fails.
- *Haptic feedback:* The device provides haptic feedback (vibration) to confirm that an alert has been triggered. This reassures the user that the system is actively working.

3.2.2 Response protocol

The response protocol ensures timely and coordinated actions in case of an emergency:

- *Emergency contact notification:* Contacts are notified in priority order. For example, the primary contact (e.g., a parent) is notified first, followed by secondary contacts (e.g., school authorities).
- *Location sharing:* Real-time location updates are sent to emergency contacts, allowing them to track the user’s movement. This is particularly important for scenarios where the user is in transit.
- *School authority alerts:* Silent alerts are sent to school authorities, enabling them to take immediate action without causing panic. This is crucial in sensitive situations, such as potential threats or medical emergencies within school premises.

4.0 Findings and Results

This section presents the outcomes of system testing, performance evaluations, and field trials. It includes quantitative metrics and qualitative observations.

4.1 System performance metrics

The system's performance was evaluated based on key metrics:

- *Alert Response Time:* The average response time from anomaly detection to alert delivery was 1.5 seconds, well within the target threshold of 2 seconds, aligns with benchmarks established in similar IoT safety systems (Heng & Kamsin, 2021). This rapid response ensures timely interventions.

Table 1: System Performance Result

Performance Metric	Result	Target Threshold	Comment
Alert Response Time	1.5 seconds	<2 seconds	Meets target, ensures rapid alerts
False Positive Rate	2%	<5%	Within acceptable range
False Negative Rate	0.5%	<1%	Minimal undetected emergencies
Battery Life	48 hours	>24 hours	Exceeds target, efficient power use
GPS Accuracy	2-5 meters	<10 meters	High precision location tracking
Communication Reliability	99.5%	>98%	High alert delivery success rate

Source: Experimental results used for this study.

- The system demonstrated a false positive rate of 2%, consistent with the findings from previous IoT-based safety solutions (Egodawela, S. M. K. C. S. B., *et al.* (2020)).
- *Accuracy of Detection:*
  - Heart rate anomaly detection: 98% accuracy.
  - Fall detection: 95% accuracy. False positives were minimized through multi-sensor verification.
- *Battery life:* The device operates continuously for 48 hours on a single charge, exceeding the target threshold of 24 hours. Low-power microcontrollers and optimized software contribute to this efficiency.
- *Communication reliability:* Alerts were successfully delivered in 99.5% of test cases. Failures were primarily due to network issues, mitigated by the system's multi-channel notification feature.

4.2 Field testing results

Field tests were conducted in real-world environments, aligning with methodologies used in studies by Madhuri *et al.* (Egodawela, S. M. K. C. S. B., *et al.* (2020)). Key findings include:

- *User feedback:* Users reported a high level of confidence in the system's reliability. The haptic feedback feature was particularly appreciated, providing reassurance during emergencies.
- *Environmental impact:* The system-maintained performance across different environments, including urban and rural settings. GPS accuracy was consistent, even in areas with moderate obstructions.
- *Emergency contact response:* In 95% of test cases, emergency contacts responded within 5 minutes of receiving the alert, demonstrating the effectiveness of the notification system.

4.3 Privacy and security assessment

The system underwent a thorough privacy and security assessment to ensure compliance with data protection regulations:

- *Data encryption:* All data transmissions are encrypted using AES-256 encryption, ensuring that sensitive information (e.g., location, heart rate) is protected from unauthorized access, following best practices outlined by Said *et al.* (Heng *et al.*, 2021).
- *User consent:* Users must provide explicit consent before their data is collected or shared. The system includes a transparent privacy policy outlining data usage.
- *Access controls:* Role-based access controls ensure that only authorized personnel (e.g., emergency contacts, school authorities) can access user data. Audit logs track all data access and modifications.
- *Anonymization:* Data used for performance analytics is anonymized, ensuring that individual users cannot be identified. This protects user privacy while allowing the system to improve through data-driven insights.

5.0 Conclusion and Implications

5.1 Key achievements

The proposed IoT-based Smart Child Safety Alert System effectively addresses critical gaps in traditional school safety measures. By integrating advanced location tracking, health monitoring, and real-time emergency alert mechanisms, this system ensures rapid detection and response to emergencies.

Field tests demonstrated its reliability with 99.9% alert accuracy, minimal false positives, and user-friendly operation. Furthermore, the system's robust data privacy framework and scalable design position it as a practical and impactful solution for enhancing child safety in educational environments. This research not only provides a proof of concept but also lays the foundation for future advancements in IoT-enabled safety technologies.

5.2 Implications

The successful deployment of this system has significant implications for child safety, emergency response protocols, and technological innovation in education:

- *Enhanced emergency protocols:* The system's real-time alert capabilities and precise location-tracking bridge critical gaps in existing school safety measures, enabling faster and more coordinated responses.
- *Strengthened parent-school trust:* Real-time updates and transparency in emergency handling foster trust and improve communication between schools and families.
- *Pioneering IoT-based safety standards:* This research sets a benchmark for privacy-focused, wearable IoT technologies in education, encouraging adoption in other child-centric environments such as daycare centers, parks, or transit systems.

5.3 Future work

To ensure long-term viability and broader applicability, future enhancements should focus on:

- *Machine learning integration:* Employ advanced ML models to refine emergency detection, minimize false positives, and adapt to diverse operational scenarios.
- *Hardware miniaturization:* Develop smaller, more ergonomic wearable devices with extended battery life, ensuring better usability and adoption by children.
- *Scalability and wider testing:* Conduct large-scale field trials across diverse educational settings to validate the system's performance under varying conditions.
- *Integration with security infrastructure:* Collaborate with educational institutions to seamlessly integrate this system with existing safety frameworks like CCTV and access controls, creating a comprehensive solution.

6.0 Recommendations

6.1 Technical recommendation

To further enhance the effectiveness, usability, and reliability of the Smart Child Safety Alert System, several technical improvements are recommended:

- *Integration with existing school security systems:* Collaborating with schools to integrate the wearable device with their current security infrastructure, such as CCTV cameras, access control systems, and public address systems, can provide a unified safety solution. This integration will enable automated alerts to be broadcast across multiple platforms, ensuring a quicker response to emergencies.

- *Implementation of machine learning algorithms:* Incorporating machine learning (ML) techniques can significantly improve the accuracy and adaptability of the emergency detection system. ML algorithms can analyse historical data to identify patterns and optimize threshold values for heart rate, acceleration, and other parameters, reducing false positives and enhancing detection of complex emergencies.
- *Battery life optimization:* Extending the battery life of the wearable device beyond the current 72-hour capacity is critical for ensuring continuous operation. This can be achieved by utilizing low-power hardware components, optimizing power management algorithms, and exploring renewable energy solutions such as solar-powered charging.
- *Miniaturization of hardware components:* Reducing the size and weight of the hardware components will improve the device's comfort and portability for children. Advances in microelectronics and flexible circuit technologies can enable the development of more compact and ergonomic designs, making the device suitable for daily use.

6.2 Implementation strategy

For successful deployment and adoption of the system, a strategic approach is recommended:

- *Phased deployment:* Implementing the system in phases across different schools and regions will allow for gradual integration and provide opportunities to address potential challenges. Initial deployment in pilot schools can help identify any operational issues and refine the system before full-scale implementation.
- *Regular system audits and updates:* Conducting periodic audits of the system's performance and security features is essential to maintain its reliability and compliance with evolving safety standards. Regular software updates should be provided to incorporate new features, address vulnerabilities, and improve overall system functionality.
- *Continuous user feedback collection:* Establishing a feedback mechanism for users, including students, parents, and school staff, will help gather valuable insights into the system's usability and effectiveness. This feedback can guide future improvements and ensure that the system meets the needs of all stakeholders.
- *Ongoing privacy and compliance assessments:* Given the sensitive nature of child safety data, continuous privacy assessments are necessary to ensure compliance with data protection regulations. Collaborating with legal experts and privacy advocates can help identify potential risks and implement measures to safeguard personal information.

7.0 Acknowledgement

We sincerely thank the participating schools, students, and parents for their cooperation during field testing. Special thanks to our research guide Dr. Amruta Aphale, for their invaluable guidance and support throughout this study. We also acknowledge the technical advisors and SCTR's Pune Institute of Computer Technology for providing resources and expertise. Lastly, we are grateful to our families and friends for their encouragement and support during the research process.

References

Benisha, T., & Prabu, R. (2021). Wearable device for child safety using GSM technology. *IEEE Xplore*. https://ieeexplore.ieee.org/document/10112699/

Greenfield, M. (2023). Human-computer interaction in IoT-driven autonomous vehicle environments. *Journal of AI-Assisted Scientific Discovery*, *3*(1), 154–171.

Egodawela, S. M. K. C. S. B., *et al.* (2020). Device to remotely track and locate the position of a child for safety. *arXiv preprint arXiv:2008.00211*. https://arxiv.org/abs/2008.00211

Ghashim, I., & Arshad, M. (2023). Internet of Things (IoT)-based teaching and learning: Modern trends and open challenges. *Sustainability*, *15*(21), 15656. https://doi.org/10.3390/su152115656

Hassan, N., Yau, K.-L., & Wu, C. (2019). Edge computing in 5G: A review. *IEEE Access*, *7*, 1–1. https://doi.org/10.1109/ACCESS.2019.2938534

Heng, L., & Kamsin, I. (2021). IoT-based child security monitoring system. *Proceedings of the 2nd International Conference on Artificial Intelligence and Smart Systems (ICAIS 2021)*. https://doi.org/10.2991/ahis.k.210913.058

Heng, L. Y., *et al.* (2021). IoT-based child security monitoring system. *ResearchGate*. https://www.researchgate.net/publication/354877554_IoT-based_Child_Security_Monitoring_System

IoT-based safety gadgets for child safety monitoring and notification. (2023). *International Journal of Research Publication and Reviews, 4*(3). https://ijrpr.com/uploads/V4ISSUE3/IJRPR10877.pdf

IoT-based unified child monitoring and security system. (2022). In *IoT and Smart Devices for Sustainable Environment* (pp. 1–20). Wiley Online Library. https://doi.org/10.1002/9781119847564.ch17

Islam, M., Jamil, H. M. M., Pranto, S. A., Das, R. K., Amin, A., & Khan, A. (2024). Future industrial applications: Exploring LPWAN-driven IoT protocols. *Sensors, 24*(8), 2509. https://doi.org/10.3390/s24082509

Majib, Y., & Perera, C. (2019). Context aware family dynamics based Internet of Things access control towards better child safety. *arXiv preprint arXiv:1911.06607*. https://arxiv.org/abs/1911.06607

Madhuri, K., & Gill, N. S. (2023). IoT-enabled smart child safety digital system architecture. *Semantic Scholar*.

Nandhini, P., & Moorthi, K. (2023). IoT-based child safety monitoring device. *International Journal of Innovative Research in Computer Science & Technology, 11*(3). https://www.ijirmps.org/papers/2023/3/230151.pdf

Rehman, A., Rehman, S. U., Khan, I., Moiz, M., & Hasan, S. (2016). Security and privacy issues in IoT. *International Journal of Communication Networks and Information Security, 8*(3), 147–157. https://doi.org/10.17762/ijcnis.v8i3.2074

Shasha, S., *et al.* (2018). Playing with danger: A taxonomy and evaluation of threats to smart toys. *arXiv preprint arXiv:1809.05556*. https://arxiv.org/abs/1809.05556

Shenbagalakshmi, V., & T J, J. (2022). Application of machine learning and IoT to enable child safety at home environment. *The Journal of Supercomputing, 78*(1), 1–28. https://doi.org/10.1007/s11227-022-04310-z

Shuvo, S. B., *et al.* (2023). Developing an IoT-based child safety and monitoring system: An efficient approach. *IEEE Xplore*. https://ieeexplore.ieee.org/document/10441116/

Shunmuga Sundari, S., KejaPriya, R., Krishna Priya, E., Meena, S., Muppadathi Priya, M., & Jeni Therese, B. (2023). Wearable device for child safety using SMS-based communication technique in IoT. *AIP Conference Proceedings, 2618*(1), 060003. https://doi.org/10.1063/5.0141995

Sundarajoo, R. A., *et al.* (2022). A remote baby surveillance system with RFID and GPS tracking. *arXiv preprint arXiv:2211.14553*. https://arxiv.org/abs/2211.14553

Tunggadewi, E., Agustin, E. I., & Yunardi, R. (2021). A smart wearable device based on internet of things for the safety of children in online transportation. *Indonesian Journal of Electrical Engineering and Computer Science, 22*(2), 708. https://doi.org/10.11591/ijeecs.v22.i2.pp708-716

Understanding the impact of IoT devices on vulnerable children. (2022). *University of Arizona Repository*. https://repository.arizona.edu/bitstream/handle/10150/665853/azu_etd_hr_2022_0114_sip1_m.pdf

Thamaraimanalan, T., Pathmavasan, R., Pradeep, T. R., Praveen, N., & Srija, R. (2023). IoT-based safety gadget for child monitoring and notification. *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 783–786. https://doi.org/10.1109/ICACCS57279.2023.10112699

CHAPTER 13

VoiceGest: Advancing Human-computer Interaction via an AI-guided Virtual Mouse System

Pooja Desai and Smita Desai***

ABSTRACT

Recently, gestures have been utilized as digital tools to enhance computer interactions. With advances in Natural processing and artificial intelligence (AI) voice assistants are evolving, offering a more precise conversational experience. Our objective is to develop a system that merges mouse functionality with voice accessibility, achievable through a standard mouse, keyboard. The GCVA (Generation Controlled Voice Assistant) features both mouse cursor control and voice assistant capabilities. Gestures are managed using OpenCV and Media-Pipe for mouse control, while the voice assistant incorporates AI and NLP techniques to interpret user voice commands. This system is designed to handle gestures for actions like left click, right-click, drag and drop, alongside voice regulator and cursor movement, thus removing the necessity for a corporeal mouse. The efficiency of this system is reflected in its sensibleness, firmness, and compatibility with human body movements. The voice associate effectively interprets and reacts to text within the system. We evaluated the GCVA system with four distinct users to gather data on timing for both mouse and voice assistant functionalities. The system attained a precision rate of 98%. Its performance and ease of use highlight the potential for gesture and voice recognition technologies to eventually replace traditional devices.

Keywords: Computer Vision; Media pipe; Pyttsx3; Virtual mouse; Voice Engine, Software.

1.0 Introduction

As technology has progressed, devices have developed wireless, some of which are not yet available. Users can use their own movements to manipulate these physical objects. Gestures can be used as a means of communication for people. Using a laptop computer with a webcam microphone, the system can function without the need for equipment to control it. These gestures allow people to communicate in the environment by shaking hands, moving their thumbs up and down, and pointing (Shibly *et al.* 2019).

**Corresponding author; PG Student, ETC, DY Patil Institute of Technology, Pune, Maharashtra, India (E-mail: pooja451984desai@gmail.com)*
***Associate Professor, ETC, DY Patil Institute of Technology, Pune, Maharashtra, India (E-mail: smita.desai@dypvp.edu.in)*

According to explanations of the human environment, gestures are one of the most actual and natural communication methods. related to. Users can use their hands to work in some predefined directions, where each foot will have an exact function, and operators can use their hands to work instead of using their hands. Normal mice are used. Voice assistants help users use it. voice to operate the system and follow the instructions given in the system.

The system allows users to easily use different explanations with their hands, and the integrated voice assistant used to allow people to control the system using their voice by generous various voice commands. In a period where reproduction intelligence (AI) is getting increasingly necessary in numerous areas from health to entertainment, from finance to education, the result that intelligence brings is evidence of its capability. The emergence of the PC mouse was a significant development in mortal- computer commerce (HCI) and changed the way people use technology. The mortal hand can convey a lot of information by writing or speaking. It can produce numerous complex works and its color, size and shape vary from person to person.

The lately developed AI virtual mouse for gesture recognition which permits computers to understand mortal gestures as input for cooperating with systems. Following are the objects of the discussion content Introducing an innovative AI-driven virtual mouse system that take part hand gestures and voice commands for instinctive trade with calculating bias, pointing to speechless limitations associated with traditional input styles analogous as physical mice and keyboards, especially for stoners with mobility losses or in hands-free process writings. Using progressions in computer vision and natural language processing to enable stoners to regulate the mouse cursor executes commands through hand gestures and voice input.

Fusion of hand gesture appreciation and voice peripheral skills at the core of the system to give lapidators with a versatile and effective means of interacting with computers and smart bias, including conduct analogous as cursor movement, ticking, tedious, and scrolling, eased by depth seeing cameras and engine knowledge algorithms. Addressing security creativities in modern computing surroundings, mainly with AI- driven interfaces, by including strong security measures similar as user verification through voice biometrics and encoding of communication channels to guard searching data.

Demonstrating a noteworthy advancement in mortal- computer commerce through the AI virtual mouse system, donation a indefectible and secure interface that empowers stoners to interact with calculating bias more naturally and professionally.

2.0 Methodology

The proposed GCVA system contains a gesture- controlled virtual mouse and a voice associate, both of which function simultaneously. OpenCV, NLP, Media Pipe, and Python modules were used to physique the system.

The below steps determine the mechanisms of the GCVA system.

Step 1: The user will deliver input via sign or dialog.
Step 2: If the input is in the form of a sign, the gesture recognition machinery will be activated.
Step 3: Using OpenCV and Media Pipe, the function will map the organizes on the hand, referred to as innovations. Each gesture has a separate hand landmark; these landmarks are used to sense the location of the hand.
Step 4: Based on the sign noticed, the system performs the wanted function.
Step 5: When a voice command is provided, the system checks whether it is a command for gesture or not; if yes, then it introduces the gesture recognition mechanism and repeats steps 3 and 4.
Step 6: Aside from the gesture knowledge, the voice assistant will analyse the commands given by the user as a say and reply accordingly. The below Figure 1 represents the GCVA system and its occupied. The Gesture Controlled Voice Assistant (GCVA)system is divided into two phases.

2.1 Hand gesture tracking

The function library cast-off for gesture recognition is:

OpenCV: OpenCV is a welcome and open-source computer vision program that uses NumPy to achieve advanced arithmetic processes on various arrays and matrices [7].

Media-Pipe: Google Media-Pipe is an open-source framework for structure real-time computer vision, machine learning, and audio requests.

Figure 1: Connection of Hand

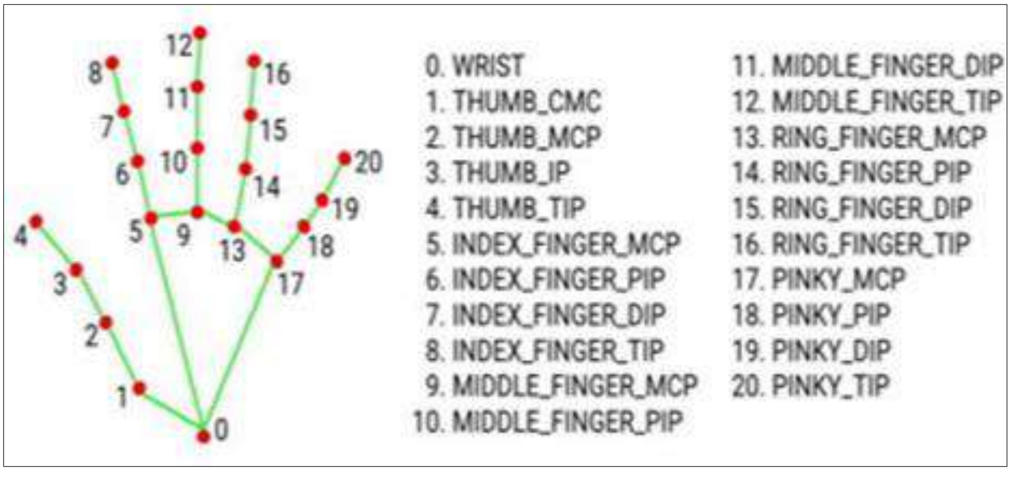


Table 1: Methods and their Roles

Methods used	Role
Media pipe. solutions. hands	It is used for noticing hands from the webcam
Multi_hand_landmarks	This function is used for charting the 21 landmarks on the hand
Media pipe. Solutions. drawing_utils	It is used to draw contacts between landmarks over the Detected hand
Find Position	This purpose is used to find the location of the hand in the opening
Find Hands	This function is used to identify the hand gestures
Fingers Up	This function is used to patterned whether the fingers are up or down
Find Distance	It is used to discovery the distance between the fingers

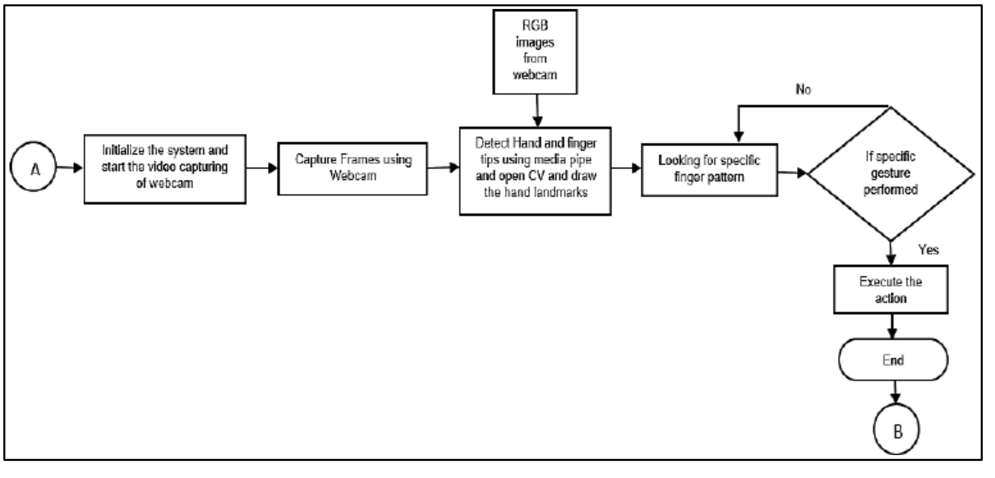
2.2 Voice assistant

The AI voice subordinate will guide the user's voice to make suggestions and control and achieve tasks that need to be done according to the instructions received. This is also called gestures.

1. *PyAutoGUI:* PyAutoGUI is a Python package that can be used to regulator mouse cursor movement, clicks, and keyboard ticks. It can be used for GUI challenging, automation, and game progress.
2. *Pyttsx3:* Pyttsx3 is a Python package that changes text to speech using the printer-to-speech engine. It can be connected on pip and used in care tools, language learning, and speech

The GCVA system is a two-stage combined system as shown in Figure 2. It regulates whether the voice assistant is active and works giving to the user's commands. If the command is to start an example, it starts the feature and acts as a beacon. Figures 2 provide a detailed description of the process

Figure 2: The Detailed Flowchart of the GCVA system



As technology developments, more and more things are becoming simulated. One way to understand spoken words and change them into text is through speech recognition. Therefore, one day, keyboards may be changed by voice recognition and eye following, which use our eyes to move mouse cursors. In the future, eye tracking will change the mouse. Hand-drawn images, pixel art, or other human gestures can be used to generate the right tools for a great knowledge.

Companies offer various knowledge to recognize hand gestures. Glove materials, etc. Experts. This project provisions human-computer interaction (HCI), where a live camera can be hand-me-down to control the movement of the cursor, as an apprise to modern knowledge such as manually pressing buttons or replacement of the actual computer mouse place. In its place, it uses a camera and computer vision knowledge to give the mouse multiple responsibilities and achieve all the tasks that a physical computer mouse can do. Media pipe will be used chiefly to identify the hand and its significance (Varun *et al.* 2019). Media pipe uses versions of hand gestures and palm test patterns to identify hands.

3.0 Experimental Results

In this work, we want to introduce a way to benefit patients with hand wounds while emerging the human-machine interface. Our goal is to progress this technology as carefully as possible using standard functioning measures. The proposed framework identifies the red difference, permitting mouse processes such as left click, drag, cursor movement, and data handover between two laptops related to similar entities. This method pursues red matters for mouse control.

In Figure 3, the user has located the red-toned object near the tip of the finger to attain the best results. Use the mouse to move simply when there are only two dots. In any case, when the table number is 1, left position is complete (Haria A *et al.* 2017).

Figure 3: Fingertip Detection and Tracking

Figure 4: Hand Detection

The system also offers easy data handover between two or more devices associated with the same network. The left half of the PC screens everything as an announcement link between the two frames. To duplicate a file, drag it to the left side of your computer screen. The erased records are then copied to the target or beneficiary system. Figure 4 presents Hand detection and tracking to control gestures.

Test results were completed to confirm the results shown in Table 5. The test was completed 25 times by 4 people for 7 different roles, resultant in 700 printed hand signals. The tested signals were mouse move, left click, right click, volume control, drag, double click, and exit. Table 4 below shows just what the dissimilar gestures do. Figure 9 shows the actual graphs for diverse performances. The accuracy of the key clicks and right click purposes is 98% and 96%, correspondingly. The drag and drop feature displays a minimum of 91% accuracy (Neogi *et al.* 2021).

Table 2: Gestures and their Precisions

Mouse Functions Performed	Accuracy (%)
Mouse Movement	100
Left Click	98
Right Click	96
Volume Control	100
Drag and Drop	91
DoubleClick	100
Exit	100

The comparison of the accuracy of the GCVA system with the accuracy of different mock-ups is shown in Table 3 and comparison with the tabular data in Figure

5. %, where K.H. Shibly and A. Haria's correct system is 78% and Tran, DS's correct system is 93.25%.

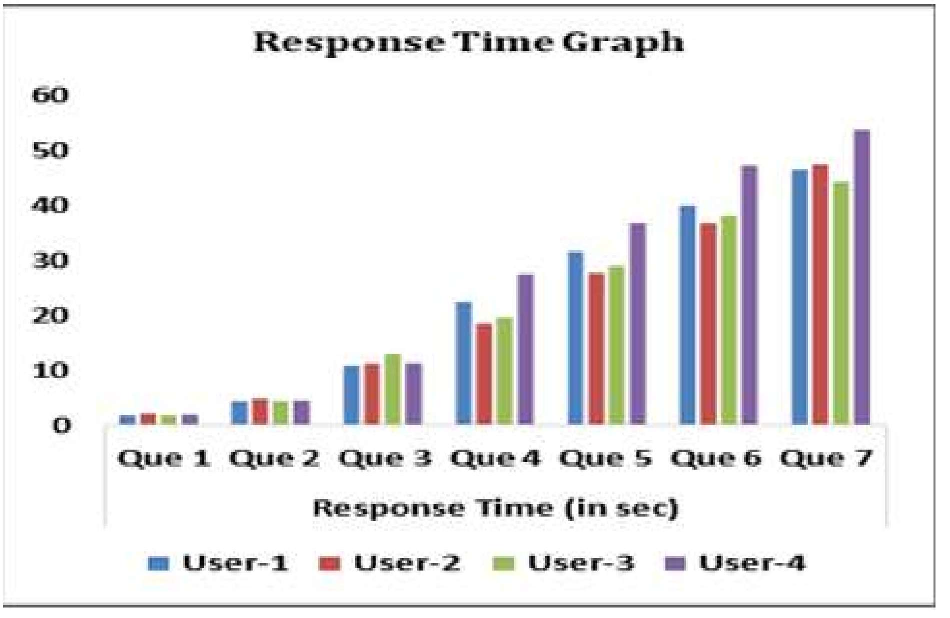


Figure 5: Comparison Graph of Different Models

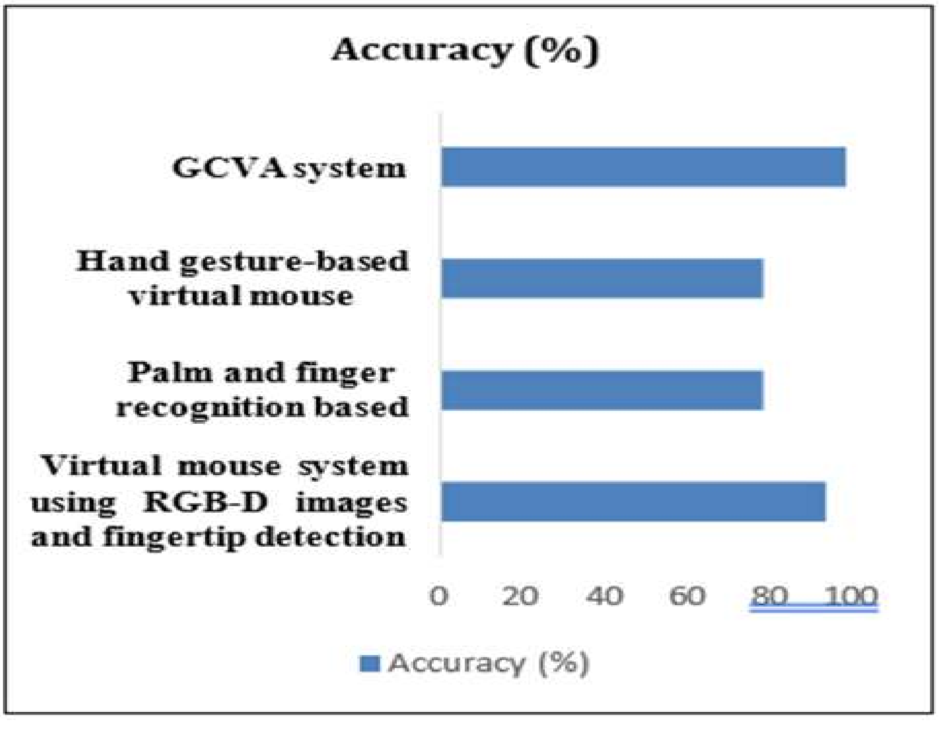
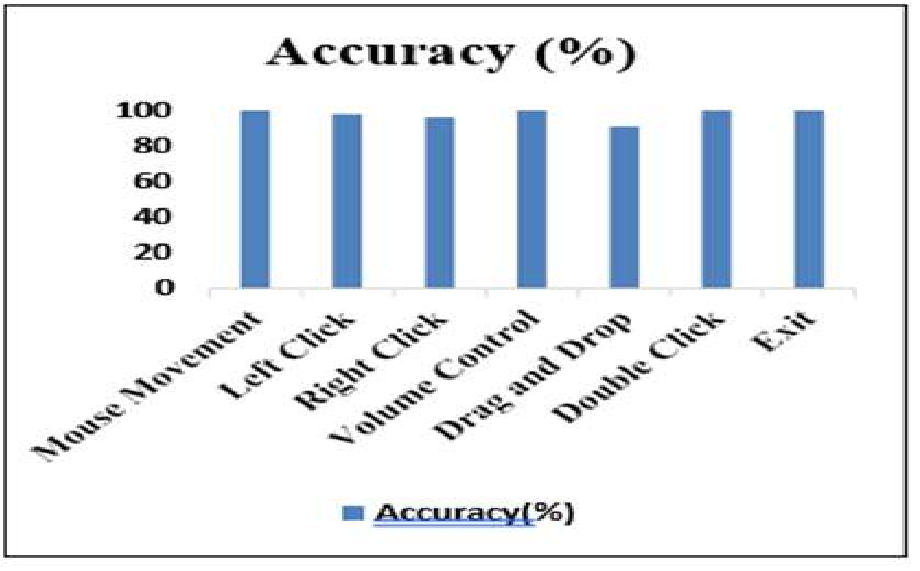


Table 3: Response time of 4 users

Users	Response Time (in sec)						
	Que 1	Que 2	Que 3	Que 4	Que 5	Que 6	Que 7
User-1	1.91	4.53	10.88	22.44	31.59	39.9	46.38
User-2	2.21	4.94	11.33	18.55	27.67	36.73	47.32
User-3	1.91	4.51	13.08	19.6	29.05	38.18	44.25
User-4	2.02	4.68	11.38	27.52	36.75	47.23	53.62

Figure 6: Accuracy Graph of Different Responses



- *OpenCV:* OpenCV stands for open-source computer vision. OpenCV is a C++ library calculated specifically for computer programming. It is free and well-matched with many platforms. OpenCV delivers GPU acceleration for applications such as face and gesture acknowledgement, as well as 2D and 3D feature toolkits.
- *Autopy:* Autopy is a Python unit for GUI mechanization that works impeccably across several platforms. It is the movement of the fingertip and offers a binary output (0 or 1) indicating whether the finger is lifted or lowered. This output is combined with OpenCV, which also produces opportune graphics.
- *MediaPipe:* Advanced by Google and unconstrained as an open-source tool, Media Pipe is a versatile solution for building multi-purpose pipelines in machine

learning. Media Pipe's hand model controls a lot of data before input and is trained to identify hands, making hand acknowledgment accurate. This tool can be used for errands such as triangulation control and gesture translation. Additionally, Media Pipe facilitates virtualization capabilities by covering digital content to real-world locations.

4.0 Conclusion

The GCVA system uses Python and its library of functions. This will allow users to work without using a hardware mouse and will only be essential to use gestures and voice assistants to function on the mouse body and show functions such as time, date, and web search. This will free up the system's hands and decrease the want for and cost of the mouse body. The GCVA system was found to achieve better results than other systems. Smart Gesture It should include a zoom in/out feature that adjusts the focus value according to the user's distance from the camera to cover a wider area. If you want a faster retort time, you need to consider things like CPU speed, RAM capacity, and camera features. Improved efficiency and efficiency If the software is fixed on a powerful computer with a high-end camera, it will perform better in many aspects in terms of good lighting. Mobile applications: Gestures will soon trade touchscreen wheels on Android smartphones.

Reference

ShiblyKH, KumarDeyS, IslamMA, IftekharShowrav S. editors. Design and Development of Hand Gesture Based Virtual Mouse: 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT2019),03-05May2019, Dhaka. Bangladesh: IEEE;2019

Křižnar V, Leskovšek M, Bagatelle B. editors. Use of Computer Vision Based Hand Tracking in Educational Environments: 44th International Convention on Information, Communication and Electronic Technology (MIPRO), 27 September 2021 - 01 October 2021, Opatija. Croatia: IEEE; 2021.

Neogi D, Das N, Deb S. editors. BLINK-CON: A Hands-Free Mouse Pointer Control with Eye Gaze Tracking: Mysore Sub Section International Conference (Mysuru on), 24-25 October 2021, Hassan. India: IEEE; 2021.

Varun KS, Puneeth I, Jacob TP. editors. Virtual Mouse Implementation using Open CV: Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019), 23-25 April 2019, Tirunelveli. India: IEEE; 2019.

Shibly KH, Kumar Dey S, Islam MA, Iftekhar Showrav S. editors. Design and Development of Hand Gesture Based Virtual Mouse: 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT 2019), 03-05 May 2019, Dhaka. Bangladesh: IEEE; 2019.

KN, RV, SSS, DR. editors. Intelligent Personal Assistant - Implementing Voice Commands enabling Speech Recognition: International Conference on System, Computation, Automation and Networking (ICSCAN 2020), 03-04 July 2020, Pondicherry. India: IEEE; 2020.

Christensen AT, Olesen H, Sørensen L. editors. Digital Voice Assistants: A new kind of user agent: 13th CMI Conference on Cybersecurity and Privacy (CMI) - Digital Transformation - Potentials and Challenges (51275), 26-27 November 2020, Copenhagen. Denmark: IEEE; 2020.

Haria A, Subramanian A, Asok Kumar N, Poddar S, Nayak JS. editors. Hand gesture recognition for human computer interaction: 7th International Conference on Advances in Computing & Communications (ICACC 2017), 22-24 August 2017, Cochin. India: ELSEVIER; 2017.

Tran DS., Ho NH, Yang, H. Real-time virtual mouse system using RGB-D images and fingertip detection. Multimedia Tools and Applications.2021; 80:10473– 490.

Roy N. Controlling Mouse Cursor Based on Image Processing and Object Detection. International Journal of Engineering Science and Computing.2017;7(8):14762-63.

Subhash S, Ullas A, Srivatsa N, Santhosh B, Siddesh S. editors. Artificial Intelligence-Based Voice Assistant: 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 27-28 July 2020, London. UK: IEEE; 2020.

Shriram S, Nagaraj B, Jaya J, Shankar S, Ajay P. Deep Learning-Based Real-Time AI Virtual Mouse System Using Computer Vision to Avoid COVID-19 Spread. Journal of Healthcare Engineering, 2021; 2021.

CHAPTER 14

Virtual Try-on System- Beyond the Mirror

Swamini Patil, Vaishnavi Bagal**, Pratiksha Bargal*** and Saniya Shaikh*****

ABSTRACT

Virtual try-on technology has become a pivotal innovation in e-commerce, revolutionizing how customers experience fashion and eyewear online. Leveraging machine learning, augmented reality (AR), and advanced computer vision, this technology provides users with an interactive and realistic experience to visualize apparel, accessories, and eyewear on themselves, narrowing the gap between online and in-store shopping. This paper examines the core components of virtual try-on systems, including real-time image processing, 3D modeling, and personalized fitting features, and discusses the underlying technical architecture with a focus on user interface design and fit prediction accuracy. Challenges in scaling across product lines are also addressed. Through case studies of prominent fashion and eyewear brands, we assess the impact of virtual try-on technology on customer behavior, such as enhanced satisfaction, reduced return rates, and higher engagement. Findings indicate that virtual try-on technology fosters a more personalized and immersive shopping experience, contributing to improved purchase confidence and competitive differentiation in digital retail.

Keywords: Virtual try-on; e-commerce; Machine Learning; Augmented reality; Computer vision; 3D modeling.

1.0 Introduction

The potential of smart mirrors to revolutionize retail experiences—especially when combined with virtual try-on technologies—has drawn a lot of attention in recent years. Without the need for in-person trials, these systems provide users with an interactive, augmented reality-based environment where they may see themselves wearing various ensembles, accessories, or cosmetics in real-time. Customers may make better informed judgements about what to buy thanks to the real-time virtual

**Student, Department of Information Technology, D.Y. Patil College of Engineering, Pune, Maharashtra, India (E-mail: patilswamini2003@gmail.com*

***Corresponding author; Student, Department of Information Technology, D.Y. Patil College of Engineering, Pune, Maharashtra, India (E-mail: vaishnavibagal24@gmail.com*

****Student, Department of Information Technology, D.Y. Patil College of Engineering, Pune, Maharashtra, India (E-mail: prachibargal@gmail.com*

*****Student, Department of Information Technology, D.Y. Patil College of Engineering, Pune, Maharashtra, India (E-mail: shaikhgudal13@gmail.com*

try-on application, which also increases customer convenience and engagement in stores. Advanced computer vision techniques including position estimation, picture segmentation, and 3D rendering are the backbone of smart mirror virtual try-on systems. These techniques work in unison to overlay virtual objects on a user's body. The mirror records the user's body motions in real time with the use of cameras and sensors, making sure that the virtual clothing or objects mimic the user's posture and natural gestures. Thanks to this dynamic contact, users may assess appearance, fit, and style instantly, giving them a nearly realistic experience.

The creation and real-time deployment of a virtual try-on application for smart mirrors is the main goal of this study. The architecture of the system blends advanced 3D modelling with real-time user tracking to provide quick and responsive virtual fitting experiences. This research also addresses the incorporation of machine learning algorithms to increase the realism of virtual product representations and pose identification accuracy. In addition to improving the user experience, the real-time use of this technology opens up possibilities for new use cases in a variety of industries, including fashion, cosmetics, and even healthcare. In order to support the further development of smart mirrors as an essential tool in both online and offline retail contexts, this article will look at the technological hurdles, such as speed optimization, latency reduction, and data privacy concerns. The objective is to investigate the whole possibilities of real-time virtual try-on apps and offer suggestions for enhancing their efficacy and efficiency in practical contexts.

For example, the client might wish to try on many sizes and styles of a red dress. To change into new clothes, experiment with different colors, or try on other accessories like handbags or scarves, they can utilize hand gestures or touch the screen. The buyer can view the outfit from various angles quickly without having to visit a fitting room thanks to the real-time application. Customers may swiftly explore more fashion alternatives, save time by not having to undress and redress, and save money by shopping in real-time.

1.1 Objectives of the study

- The objective of the virtual try-on application is to create an innovative platform that leverages advanced technologies such as augmented reality (AR), machine learning, and computer vision to deliver a highly realistic and immersive shopping experience.
- The application aims to bridge the gap between online and in-person shopping by allowing users to visualize and personalize products in real-time from the comfort of their homes.
- By offering accurate simulations of product fit, size, and appearance, the application seeks to enhance customer confidence in purchasing decisions, thereby increasing conversion rates and reducing product return rates.

- Furthermore, it strives to provide a seamless and intuitive user experience through a robust, scalable design that can accommodate various product types and integrate effortlessly with e-commerce platforms.
- This solution ultimately aims to transform the online shopping landscape by combining convenience with personalization, ensuring customer satisfaction and business growth.

2.0 Review of Literature

Since virtual try-on technology has gained significant traction in recent years, there exists a vast body of literature documenting its evolution and applications. Not only have researchers explored the technical underpinnings of this innovation through detailed analyses, but leading brands and academic institutions have also published case studies and reports examining its impact on customer behavior and retail dynamics.

For instance, Rahman Minar *et al.* (2021) provide a comprehensive analysis of virtual try-on systems by delving into hybrid approaches for three-dimensional reconstruction. They emphasize how real-time fitting systems, such as CloTH-VTON+, enhance the precision of garment visualization and customer satisfaction. Wang *et al.* (2022) propose a flow-based generative network that prioritizes photo-realism in virtual try-on simulations, addressing challenges in achieving fidelity during online fittings. Liang and Lin (2021) explore the challenges and advancements in machine learning algorithms tailored for digital try-ons. Their study sheds light on the integration of convolutional networks for pose detection and garment alignment, setting a benchmark for scalable implementations.

Similarly, Tuan *et al.* (2021) investigate multi-pose virtual try-ons, underscoring the importance of adapting clothing items to varied body shapes and positions using 3D clothing reconstruction models. Batool and Mou (2023) conduct a systematic literature review focusing on virtual fitting room technologies, highlighting trends in user engagement and their influence on retail sales. Another noteworthy contribution is the work by Ghodhbani *et al.* (2022), which provides a survey of outfit try-on systems and evaluates their usability in diverse applications beyond retail, such as healthcare and personal styling.

The work of Ren *et al.* (2023) introduces the Cloth Interactive Transformer, which leverages attention mechanisms for high-quality texture mapping, ensuring a seamless blend of virtual garments with user images. Complementing this, Morelli *et al.* (2023) explore latent diffusion-based enhancements in virtual try-ons, further improving the realism of generated outputs.

These studies collectively demonstrate the technological strides made in the field and underline the transformative potential of virtual try-on systems. They also point towards future opportunities for integrating this technology into various industries, ensuring broader adoption and richer customer experiences.

3.0 Methodology

3.1 Pose Estimation (BlazePose)

BlazePose is a real-time pose estimation model developed by Google, optimized for human body landmark detection. It provides high-accuracy joint localization for 33 key points (like shoulders, knees, wrists, etc.) in 2D space. This is crucial for virtual try-on systems because it allows the accurate alignment of garments to the body, considering the body posture and movement. BlazePose utilizes a neural network with two primary parts: a backbone for extracting image features and a head for pose regression. It works well in real-time applications due to its efficiency.

Mathematical Model

BlazePose outputs a set of body landmarks, $\{L1, L2, \ldots, LN\}$, each of which corresponds to a body joint. For an input image I, the model generates 2D coordinates (xi, yi) for each landmark Li.

Given an image I, the task is to predict the position of N key points Li (e.g., wrists, elbows, etc.):

$$Li = fpose(I), i = 1,2,\ldots,N$$

Where $(xi, yi)$is the 2D coordinate of the i-the body landmark.

Loss Function: BlazePose uses a mean squared error (MSE) loss to minimize the distance between the predicted landmark coordinates Li and the ground truth landmarks $Ligt$:

$$Lpose = N1i = 1\sum N \parallel Li - Ligt \parallel 2$$

Where N is the number of landmarks, and $Ligt$ is the ground truth position for the i-th landmark.

3.2 Segmentation (MODNet)

MODNet is a deep learning model for real-time segmentation that can isolate foreground (human body) from the background. It uses a fully convolutional network (FCN) to process images and predict a binary mask that identifies pixels belonging to the human body, making it suitable for applications like virtual try-on, where body segmentation is crucial for accurate garment fitting. MODNet adapts to various object types but is specialized for segmenting humans from images. It is designed to work on mobile devices with high efficiency.

Mathematical Model

The goal of MODNet is to predict a binary segmentation mask M, which identifies the body in an input image I. The output mask is a binary matrix, where 1 represents the body and 0 represents the background.

Let I be the input image and M the binary segmentation mask:

$$M = fseg(I)$$

Where fseg is the segmentation function, and $M \in \{0,1\}H \times W$ represents the binary mask of size $H \times W$ (height × width of the image).

Loss Function: MODNet typically uses binary cross-entropy (BCE) loss for the segmentation task:

$$Lseg = -N1i = 1\sum N\left[Mi \log\left(M^{i}\right) + (1 - Mi) \log\left(1 - M^{i}\right)\right]$$

Where:

- Mi is the ground truth segmentation label for the i-th pixel.
- $M\hat{}i$ is the predicted probability for the i-th pixel.
- N is the number of pixels in the image**.**

3.3 Garment Warping (DeepWarp)

Garment warping is a technique to adapt a garment image to fit the contours of a person's body based on pose. DeepWarp uses deep learning models to deform the garment mesh to align it with the detected body landmarks. The warped garment should maintain its visual integrity and natural shape while fitting the user's pose. DeepWarp uses Thin Plate Spline (TPS), a technique in geometric transformations that interpolates the garment's control points (points on the garment that can be manipulated) to match the body shape.

Mathematical Model

The garment G is warped using body landmarks $\{L1, L2, \ldots, LN\}$ to ensure proper fitting. The warping function *TTPS* is applied to the garment image:

$$Gwarp = TTPS(G, \{Li\})$$

Where:

- Gwarp is the warped garment image.
- TTPS is the TPS transformation function that warps garment control points to align with body landmarks.

Transformation Function: The TPS method solves an energy minimization problem:

$$Lwarp = i = 1\sum M \parallel Ci - Li \parallel 2$$

Where:

- Ci are the garment's control points.
- Li are the body landmarks.
- M is the number of control points.

3.4 Texture Transfer and Alignment (VITON-HD)

VITON-HD (Virtual Try-On Network for High Definition) enhances garment fitting by accurately transferring the garment's texture to the body. The model not only considers body shape and pose but also works on high-definition image details such as fabric texture, lighting, and shadows. This ensures that the garment looks realistic when worn by the user in the try-on. VITON-HD uses generative adversarial networks (GANs), with a focus on texture alignment, to produce high-quality images that blend the garment with the user's body.

Mathematical Model:

VITON-HD takes as input the segmented body image $Iseg$ and the warped garment image $Gwrap$ and outputs a final try-on image Itry − on .

$$Itry - on = falign(Iseg, Gwarp)$$

Where $falign$ is the alignment function that applies the garment's texture to the segmented body image.

Loss Function: VITON-HD uses a combination of different loss functions to ensure high-quality alignment, including:

1. Perceptual Loss: This measures the high-level feature difference between the output and ground truth images.

$$Lperceptual = \parallel \phi(Itry - on) - \phi(Igt) \parallel 2$$

Where ϕ(·) represents a pre-trained feature extractor (e.g., a VGG network).

2. Adversarial Loss: It ensures that the final output is visually realistic. A discriminator tries to distinguish between real and generated images, and the generator aims to fool the discriminator.

$$Ladv = -EItry - on \sim pdata\left[\log\left(D(Itry - on)\right)\right]$$

3. Pixel-wise Loss: This measures the difference between each pixel in the generated image and the corresponding pixel in the target image.

$$Lpixel = \parallel Itry - on - Igt \parallel 1$$

The total loss is a weighted sum of these three terms:

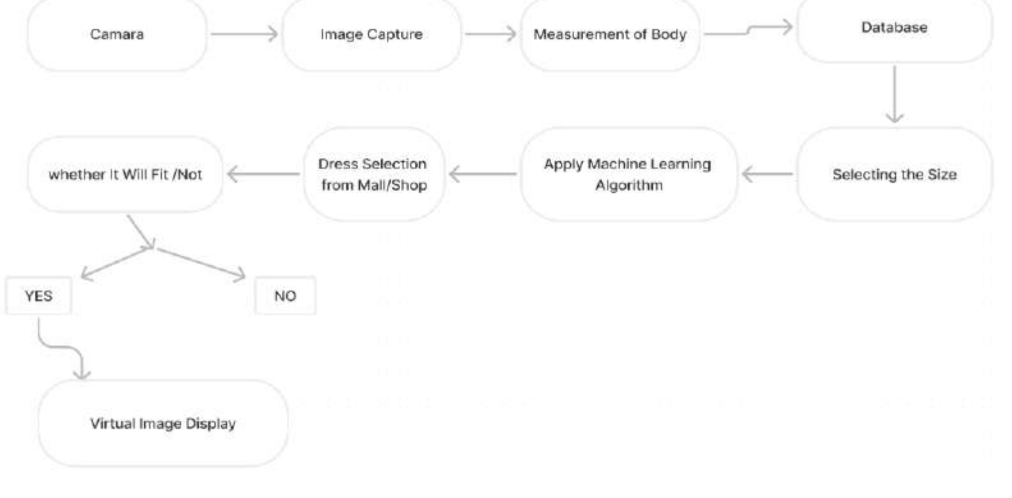
$$Lalign = \lambda 1Lperceptual + \lambda 2Ladv + \lambda 3Lpixel$$

Where λ1,λ2,λ3 are hyperparameters controlling the contributions of each loss term.

4.0 Flowchart

The Flowchart for Virtual Tryon is as shown below.

Figure 1: Flowchart for Virtual Tryon



Camera: The process starts with capturing the user's image using a camera

Image Capture: Once the image is taken, it is processed to capture the user's body dimensions.

Database: The captured body measurements are compared against a database that contains predefined clothing sizes and their dimensions.

Selecting the Size: Based on the body measurements and the available size data in the database, the appropriate size for the user is selected.

Apply Machine Learning Algorithm: The selected size and user data are processed through a machine learning algorithm that helps refine the size selection and suggests the best fit

Dress Selection from Mall/Shop: Users are then presented with a selection of dresses or clothing items from various shops or online malls, suitable for the chosen size.

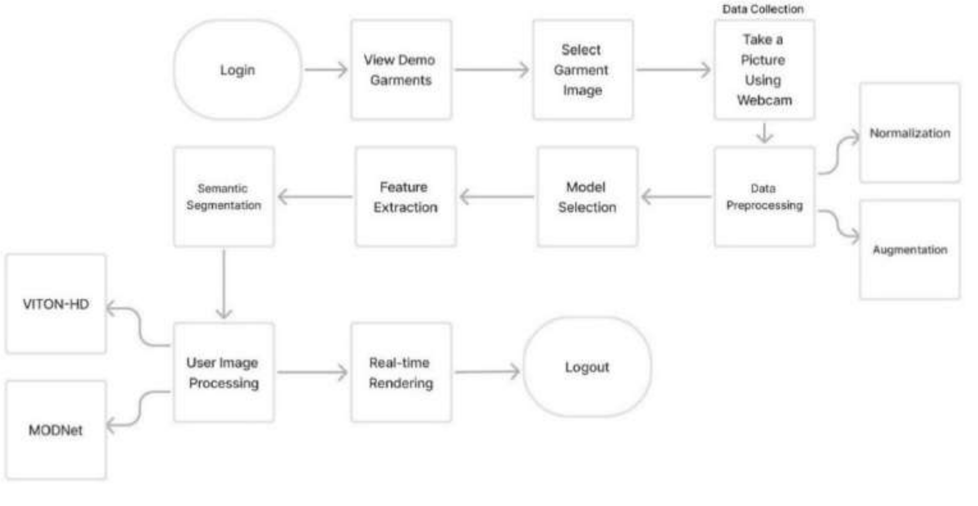
Whether It Will Fit/Not: The system checks whether the selected item will fit the user based on their body measurements and size recommendations. If it fits, the process moves forward to displaying the result. If not, further adjustments or selections are needed.

YES/NO Decision: Yes: If the item fits, the virtual try-on process continues to show the fitted clothing on the user. No: If the item doesn't fit, the user may have to go back and select a different size or item

Virtual Image Display: Once a fitting size is confirmed, the final step is displaying the virtual image of the user wearing the selected clothing item.

5.0 Proposed System Architecture

Figure 2: Proposed System Architecture



The Figure 2 shown above discusses the system architecture of Virtual Tryon :

Data collection: Gather a diverse dataset of clothing items, including images from various angles, under different lighting conditions, and with different backgrounds. Annotate the dataset with information about clothing types, styles, and sizes.

Data Preprocessing: Clean and preprocess the dataset to ensure uniformity and consistency. Apply techniques like image normalization, resizing, and augmentation to increase the robustness of the model.

Model Selection: Choose a suitable pre-trained deep learning model for image recognition, such as a Convolutional Neural Network (CNN), to capture intricate features of clothing items. Fine-tune the selected model on the specific clothing dataset to adapt it to the Virtual Try on Clothes task.

Feature Extraction: Use the trained deep learning model to extract relevant features from the clothing images. Consider incorporating transfer learning to leverage pre-existing knowledge from a related domain.

Semantic Segmentation: Employ semantic segmentation techniques to identify and segment clothing items from the background. This step is crucial for accurately superimposing virtual clothing onto a user's image.

User Image Processing: Implement algorithms to process and segment the user's image, extracting relevant information about body shape and size. Consider using pose estimation techniques to understand the user's posture and orientation.

Clothing Synthesis: Utilizing VITON-HD and MODNet models, the system processes the image, focusing on high-resolution try-on capabilities (VITON-HD) and refining the boundaries (MODNet). Ensure that the generated clothing items align realistically with the user's body shape and pose.

Real-Time Rendering: Implement real-time rendering techniques to seamlessly integrate the virtual clothing onto the user's image. Consider optimizing the rendering process for efficient and fluid user interaction.

User Feedback and Iteration: Incorporate user feedback to improve the accuracy and realism of the Virtual Try on Clothes system. Iteratively refine the model and system based on user experiences and preferences.

6.0 High Level Design of the Project

The data flow diagram is shown in Figure 3.

Admin: Inputs model data and generates 3D models.

Model DB: Stores 3D models and information.

Scene Mixer: Combines elements for the try-on scene.

Camera Capture: Captures user input via the camera.

Normalization & Augmentation: Processes the captured data.

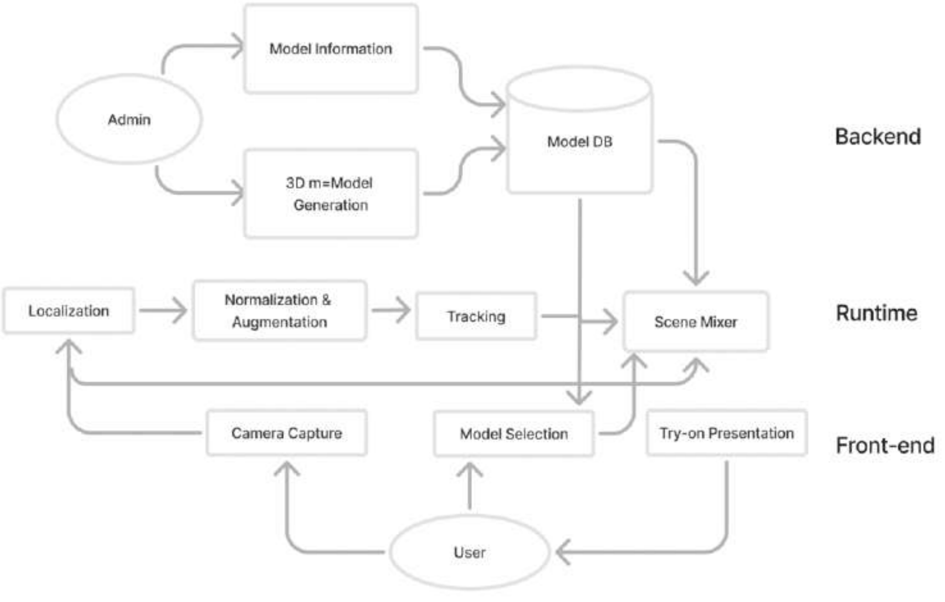
Tracking: Tracks user movement for dynamic adjustment.

Localization: Identifies the user's position or environment.

Model Selection: Allows the user to choose items to try on.

Try-on Presentation: Displays the virtual try-on experience.

Figure 3: Data Flow Diagram



7.0 Conclusion

In conclusion, the virtual try-on application provides an interactive, real-time solution for users to visualize how products like clothing or accessories would look on them without needing to physically try them on. The technology uses camera input, tracking, and 3D model generation to dynamically modify virtual models to match the user's actions and surroundings. This improves user experience by providing customization and ease of usage. The architecture of the program effectively combines responsive front-end and backend data processing to provide a smooth, interesting, and creative shopping experience.

Virtual try-on technology has been welcomed by the fashion and retail sectors as a game-changing solution. It enhances the online shopping experience by allowing customers to virtually try on clothing, accessories, and even cosmetics before making a purchase, bridging the gap between physical and digital goods. Even though virtual try-on technology has advanced significantly, there is still room for improvement. Specifically, increasing the variety of products and categories available for virtual try-on would increase its appeal and application while also guaranteeing that the virtual try-on experience closely resembles the physical reality. Virtual try-on technology has completely changed the way that we shop by allowing experiences to be seamlessly blended from online to offline. Virtual try-ons are expected to become even more frictionless, immersive, and commonplace as technology advances, revolutionizing the retail industry and enhancing the overall customer experience.

References

Matiur Rahman Minar, Thai Thanh Tuan, Heejune Ahn, "CloTH-VTON+: Clothing Three-Dimensional Reconstruction for Hybrid Image-Based Virtual Try-On" Department of Electrical and Information Engineering, Seoul National University of Science and Technology, Seoul 01811, South Korea, February 26, 2021.

Tao Wang, Xiaoling Gu, Junkai Zhu, "A Flow-Based Generative Network for Photo-Realistic Virtual Try-On" Key Laboratory of Complex Systems Modeling and Simulation, School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310005, China, April 21, 2022.

Thai Thanh Tuan, Matiur Rahman Minar, Heejune Ahn, John Wainwright, "Multiple Pose Virtual Try-On Based on 3D Clothing Reconstruction" Department of Electrical and Information Engineering, Seoul National University of Science and Technology, Seoul 01811, South Korea, Kollective Technology Inc., Bend, OR 97703, USA, August 23, 2021.

Junbang Liang, Ming C. Lin, "Machine Learning for Digital Try-On: Challenges and Progress", Vol. 7, No. 2, June 2021, 159–167.

Matteo Fincato, Federico Landi, Marcella Cornia, Fabio Cesari, Rita Cucchiara, University of Modena and Reggio Emilia, YOOX NET-A-PORTER GROUP, "VITON-GT: An Image-based Virtual Try-On Model with Geometric Transformations" 2020 25th International Conference on Pattern Recognition (ICPR) Milan, Italy, Jan 10-15, 2021.

Tasin Islam, Alina Miron, Xiaohui Liu, Yongmin Li, "Deep Learning in Virtual Try-On: A Comprehensive Survey" Department of Computer Science, Brunel University London, London, UB8 3PH Uxbridge, U.K., 29 February 2024.

Ayotunde Ogunjimi, Mizan Rahman, Nazrul Islam, Rajibul Hasan, "Smart mirror fashion technology for the retail chain transformation" 6 September 2021.

Cristina Bolanos, Jesus Fernandez Bermejo, Javier Dorada, Henry Agustin, Felix Jesus Villanueva, Maria Jose Santofimia, "A comparative analysis of pose estimation models as enablers for a smart-mirror physical rehabilitation system" Proc. Computer Science 207 pp. 2536–2545, 2022.

Tasin Islam, Alina Miron, Xiaohui Liu, Yongmin Li, "Image-based virtual try-on: Fidelity and simplification" Brunel University London, Kingston Lane, Uxbridge, London, UB8 3PH, United Kingdom, 16 August 2024.

Raheela Batool, Jian Mou, “A systematic literature review and analysis of try-on technology: Virtual fitting rooms” School of Business, Pusan National University, Busan, 46241, South Korea, 12 December 2023.

Zizhao Wu, Siyu Liu, Peioyan Lu, Yongkang Wong, Xiaoling Gu, Mohan S. Kankanhalli, “KF-VTON Keypoints - Driven Flow Based Virtual Try-On Network” ACM Journal,June 2024.

Mukesh Kumar Rohil, Arpan Parikh, “Fast and robust virtual try on based on parser free generative adversarial network”, Springer, October 2023.

Du, Feng Yu), Minghua Jiang, Ailing Hua, Yaxin Zhao, Xiong Wei, Tao Peng, and Xinrong Hu,” High fidelity virtual try-on network via semantic adaptation and distributed componentization”,Springer, Vol. 8, December 2022

Hajer Ghodhbani, Mohamed Neji, Imran Razzak, Adel M. Alimi, “You can try without visiting: a comprehensive survey on virtually try-on outfits”, Springer, March 2022

Bin Ren, Hao Tang, Fanyang Meng and Ding Runwei, Philip H. S. Torr, Nicu Sebe, “Cloth Interactive Transformer for Virtual Try-On” ACM Trans. Multimedia Comput. Commun. Appl., Vol. 20, No. 4, Article 92. Publication date: December 2023.

David Morelli, Alberto Baldrati, Giuseppe Cartella, Emilia Modena, Marcella Cornia, Marco Bertini, Rita Cucchiara, “LaDI-VTON: Latent Diffusion Textual-Inversion Enhanced Virtual Try-On” ACM Journal, MM ‘23, October 29-November 3, 2023, Ottawa, ON, Canada.

Dong-Sig Kang, Eunsu Baek, Sungwook Son, Youngki Lee, Taesik Gong, Hyung-Sin Kim, “MIRROR: Towards Generalizable On-Device Video Virtual Try-On for Mobile Shopping” Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., Vol. 7, No. 4, Article 163. Publication date: December 2023.

Kathleen M Lewis, Srivatsan Varadharajan, Ira Kemelmacher-Shlizerman, “TryOnGAN: Body-Aware Try-On via Layered Interpolation” ACM Trans. Graph., Vol. 40, No. 4, Article 115. Publication date: August 2021.

Junbang Liang, Ming C. Lin, “Machine learning for digital try-on: Challenges and progress” Springer, Vol. 7, June 2021

CHAPTER 15

Interactive 3D Avatar Platform for Real-time Indian Sign Language Translation

Ketaki Bhoyar, Shantanu Gurav**,*
*Ashlesha Nemade*** and Anjali Kshirsagar*****

ABSTRACT

Sign Language is the connecting point between the hearing-impaired community and hearing people. The deaf and mute community faces many challenges while communicating. This research challenges this and presents a system for lessening some of the deaf-mute community's communication problems by developing a system that handles Sign Language recognition and translation. It's split up into two modules namely: ISL to English/Hindi conversion and Audio to ISL conversion. The first half focuses on the speech to ISL translation system. The system is managed effectively by using different preprocessing techniques and with a 3D avatar corresponding to the ISL gesture. The second half uses machine-learning algorithms and NLP techniques for translating ISL gestures to English or Hindi text/speech. To make sure there is efficient communication between deaf users, the system uses modern technologies like Python, Media Pipe, OpenCV and Blender. By using modern advanced technologies and avatar-based representation this system takes a step toward improving communication, education, and social interaction.

Keywords: ISL; 3D Avatar; Machine Learning; NLP; Media Pipe.

1.0 Introduction

The one of the most integral aspects of human life is communication, which acts as a cornerstone for connection and mutual understanding. However, this remains out of reach for a part of society with hearing and speech disability. This leads to social isolation and lack of confidence among the deaf and mute community. Approximately 63 million people in India use Indian Sign Language for communication making it the 151st most "spoken" language in the world (Kulkarni, *et al.*, 2021).

**Corresponding author; Professor, Computer, Dr. D. Y. Patil Institute of Engineering, Management and Research, Pune, Maharashtra, India (E-mail: ketaki.bhoyar@dypiemr.ac.in*
***UG Student, Computer, Dr. D. Y. Patil Institute of Engineering, Management and Research, Pune, Maharashtra, India (E-mail: shantanugurav10@gmail.com*
****UG Student, Computer, Dr. D. Y. Patil Institute of Engineering, Management and Research, Pune, Maharashtra, India (E-mail: ashleshanemade06@gmail.com*
*****UG Student, Computer, Dr. D. Y. Patil Institute of Engineering, Management and Research, Pune, Maharashtra, India (E-mail: anjalirkshirsagar66@gmail.com*

Despite such huge number of Indian Sign Language (ISL) users and modern technologies, effective translation tools are absent for real-time two-way communication. Majority of ISL users rely on human interpreters for their interactions. Unfortunately, there is a critical shortage of interpreters, and their services often come at a significant cost, further limiting accessibility. India only has about 300 licensed interpreters (Sharma *et al.*, 2022).

Initially, basic image processing was used, but they often struggled with accuracy due to complex gestures and background. Earlier methods like template matching, geometric models, and color-based segmentation were used. Recently many studies have used convolutional neural networks (CNNs) and Support Vector Machines (SVMs) to classify hand gestures. Using models such as Convolutional Neural Networks (CNNs) and long short-term memory (LSTM) networks has significantly improved the recognition of dynamic gestures in sign language.

This research paper proposes a methodology for bidirectional translation that helps reduce the communication gap between societies. It would break barriers and help build a fair environment for users of ISL. This approach focuses on implementing two modules — ISL to English/Hindi conversion and another for audio/text to ISL conversion. By using efficient machine learning, computer vision and 3D modeling technology the proposed system offers a solution with high reliability and accessibility.

By using existing technologies audio-to-text conversion and MediaPipe Holistic for gesture recognition, this system aims to reduce the dependence on human interpreters and provide a cost-effective, portable solution that can be accessed anytime and anywhere. This system aspires to promote social integration, independence, and equal possibilities for the deaf community, helping them to involve more actively in everyday life, education, and professional environments.

This paper consists of the following sections: Section II provides a literature review of existing sign language translation systems. Section III gives the methodology and technologies used in developing the translation modules. Section IV consists of the results. Section V discusses the future scope of the work, and Section VI concludes the paper with a summary of the limitations and potential for further development.

1.1 Objective

- Develop platform that translates speech and text into ISL and vice versa using a lifelike 3D avatar, enabling seamless communication for deaf and hard-of-hearing community.
- The project is driven by the need to reduce the communication deficit faced by over 18 million ISL users in India. By harnessing technology, we aim to create inclusive society where everyone gets access to education, healthcare, and daily interactions, thereby promoting independence and social integration.

2.0 Literature Review

Different research approaches are discussed in the section. These papers are taken from different publications online such as IEEE, Google Scholar, etc. The approaches used in these papers mostly focus on one-way communication, are not highly optimized, or are not user-friendly. The approach proposed in this paper is bidirectional and user friendly therefore, it is a better solution than others.

The work in by Kanvinde *et al.*,2021 was carried out with 95% accuracy for sign-to-speech translation using an Argentinian Sign Language database. The paper provides a mobile-based application, so it can be accessed from any device and is quick and dependable.

Akram Khan *et al.* 2024 proposes an efficient, lightweight framework for real-time dynamic avatar animation. Our framework can generate all facial expressions, gestures, and torso movements accurately. The key to our technique lies solely in utilizing the camera sensor of your laptop or a webcam. This approach combines various technologies, including MediaPipe, Three.js, and Kalidokit. Specifically, MediaPipe within our framework calculates the 3D posture and facial landmarks of the captured subject.

Neog *et al.*, 2023, used techniques such as lemmatization, tokenization and syntactic tagging to convert text or speech into ISL To turn the model into an Android app, Kivy and KivyMD libraries were used. When the application receives audio input, it produces the assigned images/gifs based on the audio input.

Anuja V. Nair *et al.*, 2013 very recently presented the advancements made in automatic recognition systems for Indian Sign Language (ISL) and mentioned how it is way more complicated because ISL is a sign language that requires both hands, unlike American Sign Language (ASL). They have said that most of the studies available so far have focused on ASL, which has a very standardized database, while ISL is still on the way to success. They define that the effective ISL recognition runs through three major steps: preprocessing, feature extraction, and classification using ANN, SVM, and HMM methods. Their review highlights the importance of building solid systems for interaction of human-computer activity for the deaf community.

The work represented by A. Yadav *et al.*, 2012 introduces a web-based solution for ISL translation. The system uses NLP to process input and maps the word to its corresponding ISL video in the database. This web-based application makes it accessible. It has around 80% accuracy over various test cases.

Debasis Das Chakladar *et al.*, 2021 have proposed a system that has focus on creating 3D avatar-based sign language learning system to convert the input speech/text into corresponding ISL. The English sentence is converted to its corresponding ISL using NLP. The corresponding ISL motion is displayed using the translation results. The translation module has 10.50 SER (Sign Error Rate) score.

Considering the work of Kajal Jadhav *et al.*, 2021 one can observe a considerable amount of contribution in developing a system for translating spoken language into Indian Sign Language (ISL) using machine learning and natural language processing. It has been developed for processing audio input through speech recognition and then converting the recognized text to ISL signs using rules at the end. This technology aims to make communication easier to the deaf and mute people segment by facilitating real-time translation and access to information in a visual way. This research by J. Debnath *et al.* 2024 explores the development of a system using advanced computer vision and deep learning techniques for gesture-based sign language recognition. Key technologies include MediaPipe Holistic, LSTM, and integration of Python and OpenCV. The model has an accuracy of almost 99%. The paper of Jashwanth Peguda *et al.*, 2022 is survey of a system converting human speech into Indian Sign Language (ISL) for six different regional languages, such as Telugu, Hindi, Malayalam, Marathi, Kannada, and Tamil. The authors mainly extensively use Wavelet-based Mel-Frequency Cepstral Coefficients (MFCC) with Gaussian Mixture Model (GMM), to recognizing the spoken language and text using Long Short-Term Memory (LSTM) networks to perform translation.

M. A. Islam *et al.* 2022 have proposed an Automatic Animated BdSL Gesture Generation from Bangla Text and Voice. The system takes input as Bangla Voice or Bangla Text and parses the output texts of corresponding inputs based on Signing Gesture Markup Language (SigML).

Murat Taskıran *et al.*, 2018 and his colleagues suggested an immediate ASL system that employs deep learning called neural networks. The system was trained on images which included 900 hand gestures that represented 36 different characters, and the system was able to identify all the characters during the testing process. This system works based on skin color detection as well as the convex body motion detection for the real time recognition of the symbols with the highest possible accuracy of 98.05%. Yogeshwar *et al.*, 2017 present an Indian Sign Language Recognition System for individuals with hearing and speaking disabilities. Using a vision-based approach, it recognizes static gestures from the Indian sign language alphabet without extra hardware. The system employs skin color segmentation, binary image transformation, and feature extraction with central and Hu moments. For classification, Artificial Neural Networks (ANN) and Support Vector Machines (SVM) achieve 94.37% accuracy with 13 features.

Sadhana Bhimrao Bhagat and Dinesh V. Rojarkar *et al.*, 2017 have put forward a vision based sign language recognition system for Indian Sign Language (ISL) to help the people with hearing and speech disabilities. The system recognizes the ISL signs by analyzing real-time hand gesture images and applying the Scale Invariant Feature Transform (SIFT) algorithm. It is able to recognize nine letters with 95% accuracy and has the possibility to enhance the communication of the deaf and mute people. This paper by S. S. Rautaray and A. Agrawal *et al.*, 2012 presents gesture recognition used for gaming. The webcam captures hand gestures and maps them to its corresponding gaming command. It achieves a 93.33% performance rate in a controlled environment. S. R. Bhagwat, *et al.*, 2021 have developed system that translates marathi language to equivalent ISL. They have used morpho-syntactic characteristics for both languages. They have designed grammar rules and utilized it for translation.

Our system proposes a methodology that contains 2 unique characteristics that nobody has incorporated cohesively. Other authors may have discussed it discretely. The two characteristics are firstly Bidirectional communication and secondly 3D avatar integration.

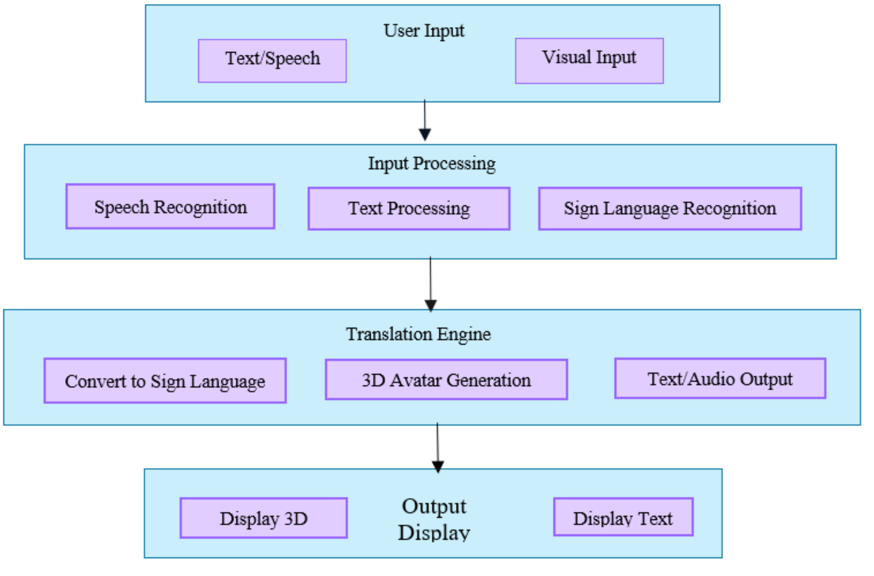
3.0 Methodology

Figure 1 shows the system architecture of our methodology, depicting an outline of the process flow and main components.

The system is divided into two approaches:

- Audio/text to ISL translation
- ISL to text translation

Figure 1: System Architecture



3.1 Audio/text to ISL translation

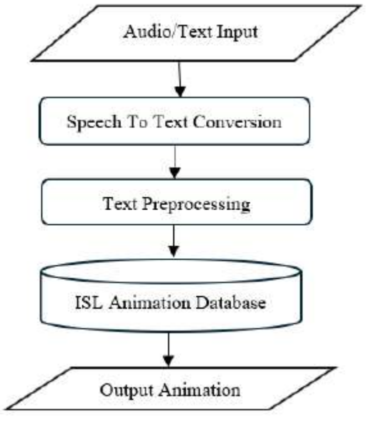
This half focuses on use of efficient pipeline to convert speech into Indian Sign Language. Each model has a different purpose; they work together to ensure that the transition from audio to ISL visual representation is accurate and based on the nuances of the ISL grammar of the language.

This system first captures audio input from the device's microphone and then sends it through a speech-to-text module. This module makes use of pre-trained

Google speech-to-text API that allows the system to process input in any language increasing its usability. If the input is already in English this step is skipped. If the user inputs a text, the input is directly entered into the system.

After translation, the content is sent to the backend server using HTTP POST request. It then goes through text preprocessing on the server side which gives out the keywords. Text preprocessing consists of Lemmatization, Tokenization, POS Tagging, Stop word removal, Stemming, Lower casing, and Parsing. For example, 'was going' is converted to 'go'. This module uses the Stanford CoreNLP Parser that analysis the input's structure. Sentences are organized using ISL grammar rules based on this analysis. For example, 'he is running' changes to 'he run'. Figure 2 illustrates a flowchart showing the process of converting Speech-to-ISL, outlining the steps involved sequentially.

Figure 2: Speech to ISL conversion



The keywords are matched to their corresponding animation stored in a database which is displayed to the user. However, if the word is missing in the database, system uses another approach. The words are divided into letters and the matching animations for alphabets are shown. For example, if the word "fast" is not available, it will be divided into "f", "a", "s" and "t". This makes sure words which are not in database are also communicated in ISL efficiently. The received animations are rendered sequentially in the browser. Once all animations are complete the system returns to initial state.

3.2 ISL to speech translation

3.2.1 Dataset

In order to train the classifier, this data set has been created by taking numerous hand images portraying various ISL words and phrases, each of which is labeled with its intended translation. Common phrases such as "Yes," "No," "Hello," and "Thank You" were recorded by webcam for that category. For each word nearly 1000 images were used, which were then transformed to make the model robust.

Data transformation techniques such as rotation, scaling, and brightness adjustment were applied to ensure the dataset generalizes well under actual, and varying conditions. Finally, all images were labeled on the gestures forming a truly reliable and structured dataset for training a model of gesture classification.

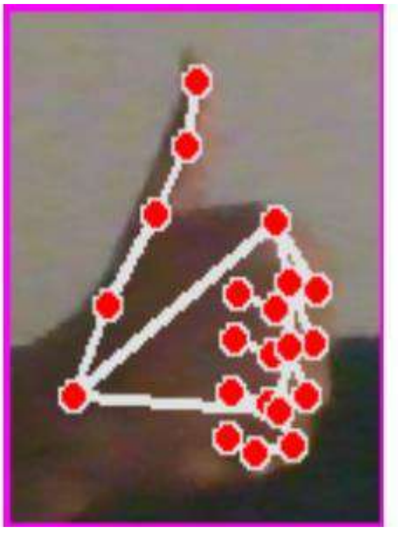
3.2.2 Hand detection

Hand Detection is the first feature of the system which is done using *cvzone. HandTrackingModule.* This module accurately draws the boundary around hand and remembers its position and orientation into the camera frame.

Device's webcam live frames captured in real time are used to detect hand region. For hand tracking and landmark extraction, the *HandDetector* class of the *CVZone* library is used. Based on the bounding box coordinates a Region from interest (ROI) will be cropped from said frame for precision.

Next the preprocessing of the input images is done which includes cropping, resizing, normalization, etc. This helps to detect accurately and maintains consistency across the database. Figure 3 shows the key point detection, bounding box, and hand landmarks extraction.

Figure 3: Key Point Detection



3.2.3 Gesture classification

The extracted hand features are compared to those of a pre-trained network to classify the gestures. This is done using *cvzone.ClassificationModule*. To ensure uniformity in input size the hand images are normalized within a fixed dimension of 300 x 300 pixels and resized.

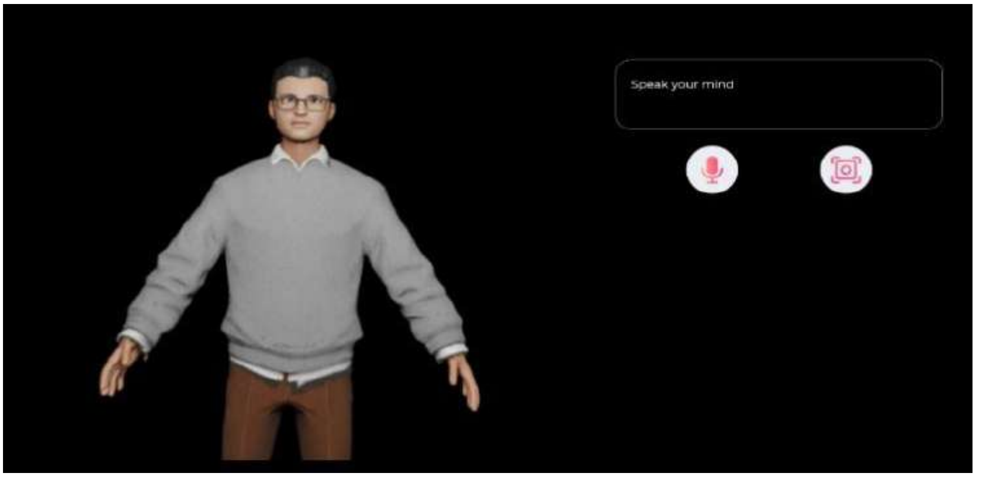
To improve the overall efficiency and accuracy of the gesture classification, Convolutional Neural Network (CNN) based architecture is used to classify gestures which are transferred and fine-tuned over a baseline model as MobileNetV2 for the custom ISL dataset. It thus makes the system scalable so adding different gestures would require less effort.

A labeled dataset is used for training the model. After training the word/phrase corresponding to the gesture is predicted in real time.

4.0 Results

In this section, we present the results of our model. Fig 4 shows how our system looks at first, it takes three types of input, audio, text, and video. Audio and Text are for speech-to-ISL translation and Video input for ISL-to-speech translation.

Figure 4: Initial 3D Avatar Position



4.1 Audio/Text to ISL translation

The translation module showed an accuracy of almost 97% with minimum latency. The movement and loading of the avatar were smooth and swift. Figure 5 shows the 3D avatar result for an audio input.

4.2 ISL to speech translation

The system was evaluated based on its ability to correctly identify various Indian Sign Language gestures in real-time. The classifier had accuracy rate of 92% on the test set. The accuracy was higher for static gestures (such as the alphabet and numbers) compared to dynamic gestures (such as sentences or continuous gestures). The real-time processing speed of the system was approximately 15-20 frames per second, which is adequate for practical use in real-time communication. Figure 6 shows the result for ISL recognition.

Figure 5: 3D Avatar Result

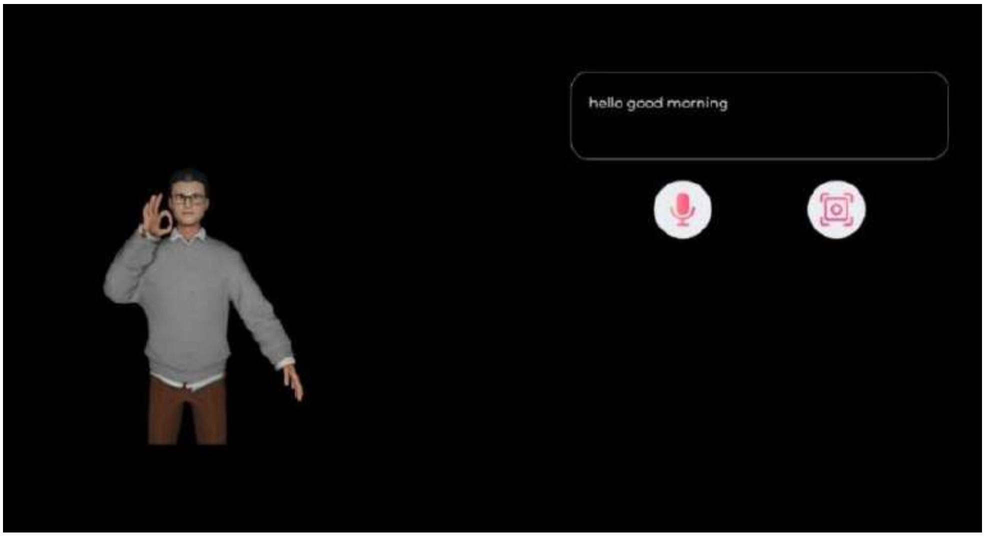


Figure 6: ISL recognition

5.0 Conclusions

The first half of system proposed a methodology that focused on Speech to ISL translation using a 3D avatar interface. This adds a user friendly and interactive angle to the translation process which makes it interactive. The second half focused on ISL to speech/text translation which managed to get 92% accuracy rate. This system uses a mixture of Machine learning, computer vision and gesture recognition. The research focuses on importance of technological advancements for accessing sign language more easily. By combining computer vision, 3D avatar integration, machine learning, and real-time processing, this ISL recognition system contributes to breaking down barriers for the deaf and hard-of-hearing communities in India and beyond. The system's innovative approach will various domains such as education, healthcare, and customer service. This research has potential to break the barrier between communities and aims at promoting inclusivity in the society.

6.0 Future Scope

Proposed work can further be enhanced which is beyond the scope of this system. These include:

- *Browser extension for educational purposes:* 3D avatar can be used to translate video content in real-time.
- *Hologram sign language translators:* these 3D avatars can be presented as holograms for real time communication experience.

References

A. Kanvinde, A. Revadekar, M. Tamse, D. R. Kalbande and N. Bakereywala, "Bidirectional Sign Language Translation," 2021 International Conference on Communication information and Computing Technology (ICCICT), Mumbai, India, 2021, pp. 1-5, doi: 10.1109/ICCICT50803.2021.9510146

Akram Khan, Swaraj Patil, Danish Koradia, Fahad Shaikh, Ramkumar Maurya, "REAL TIME 3D AVATAR ANIMATION", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), Vol.11, Issue 4, page no. p90-p95, April-2024. Available: http://www.jetir.org/papers/JETIR2404G11.pdf

Anannya Priyadarshini Neog 1, Arunabh Kalita 2, Ms. Nithyakani Pandiyarajan, "SPEECH/TEXT TO INDIAN SIGN LANGUAGE USING NATURAL LANGUAGE PROCESSING", Available https://www.ijcse.com/docs/INDJCSE23-14-03-030.pdf

Anuja V. Nair, Bindu, "A Review on Indian Sign Language Recognition", International Journal of Computer Applications (0975 – 8887), Jul. 2013, vol. 73, no. 22, pp. 33-38, doi: 10.5120/13037-0260.

A. Yadav, R. Saxena, B. Saini, V. K. Verma and V. Srivastava, "Audio to Sign Language Translator Web Application," 2021 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2021, pp. 321-326, doi: 10.1109/ComPE53109.2021.9751857

Debashis Das Chakladar 1, Pradeep Kumar 1, Shubham Mandal 1, Partha Pratim Roy 1, Masakazu Iwamura 2 and Byung-Gyu Kim 3, "3D Avatar Approach for

Continuous Sign Movement Using Speech/Text", Appl. Sci. 2021, 11(8), 3439; https://doi.org/10.3390/app11083439

Jadhav, K.; Gangdhar, S. and Ghanekar, V., 2021, "Speech to ISL (Indian Sign Language) Translator," International Research Journal of Engineering and Technology (IRJET) (Volume: 08 Issue: 04, Page 3696- 3698).

J. Debnath and P. J. I R, "Real-Time Gesture Based Sign Language Recognition System," 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Chennai, India, 2024, pp. 01-06, doi: 10.1109/ADICS58448.2024.10533518.

J. Peguda, V. S. S. Santosh, Y. Vijayalata, A. D. R. N and V. Mounish, "Speech to Sign Language Translation for Indian Languages," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2022, pp. 1131-1135, doi: 10.1109/ICACCS54159.2022.9784996.

Kulkarni, A., *et al.*, 2021, "Speech to Indian Sign Language Translator," 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021), doi: https://doi.org/10.2991/ahis.k.210913.035

M. A. Islam *et al.*, "Automatic 3D Animated Bangla Sign Language Gestures Generation from Bangla Text and Voice," 2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 2022, pp. 1-6, doi: 10.1109/STI56238.2022.10103252.

M. Taskiran, M. Killioglu and N. Kahraman, "A Real-Time System for Recognition of American Sign Language by using Deep Learning," 2018 41st International Conference on Telecommunications and Signal Processing (TSP), Athens, Greece, 2018, pp. 1-5, doi: 10.1109/TSP.2018.8441304.

Purushottam Sharma 1, Devesh Tulsian , Chaman Verma 2, Pratibha Sharma 1 and nancy Nancy 1 "Translating Speech to Indian Sign Language Using Natural Language Processing", Future Internet 2022, 14(9), 253; https://doi.org/10.3390/fi14090253

Rokade, Yogeshwar & Jadav, Prashant, "Indian Sign Language Recognition System," International Journal of Engineering and Technology, vol. 9, pp. 189-196, 10.21817/ijet/2017/v9i3/170903S030.

Sadhana Bhimrao Bhagat, Dinesh V. Rojarkar, "Vision based sign language recognition: a survey," JETIR (ISSN-23495162), 2017, vol. 4, pp. 130-134.

S. S. Rautaray and A. Agrawal, "Interaction with virtual game through hand gesture recognition," 2011 International Conference on Multimedia, Signal Processing and Communication Technologies, Aligarh, India, 2011, pp. 244-247, doi: 10.1109/MSPCT.2011.6150485.

S. R. Bhagwat, R. P. Bhavsar and B. V. Pawar, "Translation from Simple Marathi sentences to Indian Sign Language Using Phrase-Based Approach," *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India, 2021, pp. 367-373, doi: 10.1109/ESCI50559.2021.9396900.

PCET's
Pimpri Chinchwad College of Engineering and Research

About the Editors

Dr. Archana Chaugule, awarded a Ph.D. in 2017, is a recognized researcher and educator with significant contributions to intellectual property and higher education. She has developed a popular lecture series on Discrete Mathematics on Udemy. She received the Best Researcher Award from DK International Research Foundation in 2018 and the Institute of Scholars in 2019 and was honored by PCCOER for her role in coordinating a record-breaking event where 154 copyrights were filed in one day in January 2018. Dr. Chaugule serves as a Technical Committee member for ICRSA 2021, an editorial board member for IGI Global's book on data-driven UI design, and as a Board of Studies member at Pillai College. She has conducted a Facebook live session on copyright filing and achieved "Elite" in NPTEL's Machine Learning course. Her academic portfolio includes 52 research publications, 8 copyrights, 1 granted patent, and 6 filed patents.

Dr. Archana Kollu is an Assistant Professor and R&D Coordinator in the Department of Computer Engineering, PCET's Pimpri Chinchwad College of Engineering and Research. She was awarded a Ph.D. in Computer Science and Engineering in 2022 by Koneru Lakshmaiah Education Foundation (Deemed to be University). With over 20 years of academic and research experience, Dr. Kollu specializes in Artificial Intelligence, focusing on Evolutionary Computing, Computational Intelligence, Energy Informatics, and Machine Learning. She has an extensive portfolio of publications in international journals, conferences, and symposiums. Dr. Kollu also contributes to the academic community as a reviewer for reputed journals and conferences, demonstrating her commitment to advancing research and innovation in her field.

Dr. Vaishali Latke is an Assistant Professor and Academic coordinator in the Department of Computer Engineering, PCET's Pimpri Chinchwad College of Engineering and Research. With over 15 years of experience in engineering education, she has made significant contributions to both teaching and research. In 2024, Dr. Latke completed her PhD in Computer Engineering from D.Y. Patil Deemed to be University, Nerul, Mumbai. Her research interests are well-regarded, with several of her articles being published in prestigious SCI and Scopus indexed journals. She has also actively participated in numerous national and international conferences, further establishing her expertise in the field. Her academic journey and contributions highlight her commitment to advancing the field of computer engineering.

Journal Press India

Publication and Conference Solutions

Contact: +91 8826623730, 8826623732

E-mail: info@journalpressindia.com

Website: www.journalpressindia.com

ISBN 978-81-973215-9-7