

A Survey on multimodal face- voice authentication system

Keerthana R¹, Nithin S², Prema S³, M Sreenath Reddy⁴, Dr. Shruthi M⁵

^{1,2,3,4}Department of ECE, Vemana Institute of Technology, Bengaluru

⁵Associate Professor, Department of ECE, Vemana Institute of Technology, Bengaluru

Abstract—This work explain about a smart security system that uses both face and voice recognition for user authentication. The main aim is to improve security and make the system more reliable than normal single biometric methods. Different AI techniques, cameras and microphones are used to identify users and allow secure access. This survey reviews various systems used in smart homes, attendance systems, banking and security applications. From the reviewed papers, it is understood that combining face and voice recognition gives better accuracy and reduces unauthorized access. Some challenges like lighting problems, background noise and fake image attacks still affect system performance. Future improvements can focus on improving accuracy, security and system speed in real-time applications.

Index Terms—Smart AI Robot, Face Recognition, Voice Recognition, Human-Robot Interaction, Live Streaming, Computer Vision, Natural Language Processing, Security System, Deep Learning, Real-time Processing

I. INTRODUCTION

This work explains a smart multimodal authentication system that uses both face and voice recognition for secure user verification. The system combines image processing and speech analysis techniques to identify authorized users and improve security compared to single biometric methods. Different AI-based models, cameras, microphones and embedded devices are used for authentication and monitoring applications. The survey reviews recent methods developed for smart

homes, attendance systems, banking security, vehicles and humanoid robots. From the reviewed studies, it is observed that combining facial and speech features improves reliability, reduces unauthorized access and provides better performance in real-world conditions. However, factors such as lighting conditions, background noise and spoofing attacks still affect system accuracy. Future improvements can focus on faster processing, stronger spoof detection and better performance in practical environments.

II. LITERATURE SURVEY

[1]S. S. Girajala, D. Putty, V. N. S. K. Cherukuri and S. Chinnadurai, “Dual-Model Biometric Authentication Using Face and Voice Recognition”, 2025.

The study explains a security authentication system that uses both facial and speech verification to improve security and handle real-world issues like masks and background noise. It uses AI-based models such as face recognition model and facial authentication model for face, voice processing models for voice It fuses them with a combined score of 60% face and 40% voice, allowing access if the total exceeds 80%. It was tested on 25 people with masked and unmasked images plus voice samples. Face Net performed better than Arc Face, providing good accuracy and 89.70% masked, while voice recognition stayed stable at 98.20%. Overall accuracy reached 95.63% for unmasked and 93.10% for masked faces using Face Net showing that combined biometric method is more reliable than single biometrics. The system uses contrast enhancement methods and noise filtering for initial processing, works in live conditions

with webcam and mic and is suggested for safe access in banking, healthcare and offices. Future work includes quicker system response, deepfake detection and including additional security methods like iris or fingerprints.

- [2] D. L, A. P. Shetty, S. S. Shetty, V. K. S, M. M. V and Fai rooza, "A Multimodal Biometric Approach for Physical Locker Security: Integration of Face and Voice Recognition", 2025.

The study explains a smart locker security systems that uses both face and voice verification to replace insecure keys and PINs. It uses AI-based methods with OpenCV for facial identification and Mel-Frequency Cepstral Coefficients (MFCC) for speech features, with both face and voice required to meet verification conditions before locker is opened. Tested on 50 users with 500 sessions, gave good face recognition results, 93.7% voice accuracy and 92.1% combined accuracy with 1.8% FAR, 3.5% FRR and 3.2 seconds average authentication time. The backend is built using Python 3.9, Flask, PyQt5 GUI and Firebase for live data recording, while an ESP32 microcontroller triggers the locker. Real-time live detection and automatic adjustment method are used to handle noise and lighting issues.

The GUI screenshots confirm it works practically, granting access only when both biometrics match and denying it for unregistered users or voice mismatch. The system offers high better protection against fake access and is suitable for personal, institutional and banking locker systems.

- [3] S. Venkatesha, V. V. Hiremath, S. Rani G, C. M. Sharvani and C. H. Suvarna Shree, "Voice E-Mailer System with Face Recognition using Machine Learning Techniques", 2025.

This paper explains an easy-to-use email system that combines face recognition and voice-based email functions to improve security and accessibility. The system identifies users using machine learning techniques and allows them to send emails, snooze messages and set reminders through voice commands. It also provides audio responses and supports multiple languages such as Kannada, Hindi, English, Tamil,

Malayalam and Marathi for better user interaction. The system is developed using Flask, HTML, CSS and JavaScript for smooth working and user-friendly operation. After successful face verification, voice commands are activated to perform different email tasks. The results showed good face recognition accuracy and reliable voice command performance, making the system useful especially for users who prefer voice communication or have difficulty using normal typing methods.

- [4] S. Supriya, K. Madhu, M. Sreekrishna, K. Durga Prasad and G. S. Gopika, "A Data Science-Driven Face Recognition Attendance System Integrated with a Voice Assistant", 2025.

This paper explains a contactless attendance system that uses face recognition and a voice assistant to improve security and accessibility. The system captures facial images of students under different lighting conditions and angles for better identification. AI-based methods are used to extract facial features and identify students through a live camera feed. During operation, the system detects faces, compares them with stored student data and records attendance along with details such as name, ID and time. It also identifies absenteeism when attendance is missed beyond a fixed limit. The voice assistant allows hands-free interaction using voice commands, making the system useful for elderly and visually challenged users. The results showed reliable face identification in different conditions, and the IoT-based database helps in updating attendance instantly, providing a faster and more hygienic alternative to manual attendance methods.

- [5] Le Yang WeiTang, Xi hu Wang, "Design and Implementation of a Multimodal Face Recognition Access Control System Based on Raspberry Pi 4B", 2025.

This paper presents an IoT-enabled smart security framework for a 3D printed smart home. The system combines sensors, facial recognition, surveillance cameras, door lock mechanisms, Arduino, and Raspberry Pi to improve home security. The smart home can detect motion, monitor temperature, humidity, air quality, and recognize faces using the

LBPH facial recognition algorithm. Residents can unlock the door using fingerprints or passwords after successful face recognition. The surveillance system records videos based on the confidence level of recognition. The main advantage of this system is the integration of security, monitoring, and automation into a single smart home framework. However, the facial recognition system is sensitive to lighting conditions and does not always provide very high accuracy. The project was successfully implemented and tested, but future improvements are needed for better recognition performance and advanced smart home features.

[6]BOUKERCHE Alaa Eddine ,CHELALI Fatma Zohra, “Real Time Face Recognition System using Raspberry Pi4”, 2024.

This paper presents a real-time face recognition system using Raspberry Pi 4 and Convolutional Neural Networks (CNN). The system uses video sequences, OpenCV, Viola-Jones algorithm, and deep learning techniques to recognize faces in real time. It was mainly developed for attendance monitoring in institutions and organizations to replace traditional manual attendance systems. The system achieved about 94% accuracy using CNN models and supports features such as GUI interface, automated attendance recording, email sharing, and mobile access through Raspberry Pi. One major advantage of the system is its ability to perform real-time recognition with good accuracy using low-cost hardware. However, the system performance can be affected by lighting conditions, camera quality, viewing angles, and physical appearance changes like glasses or hairstyles. Although Raspberry Pi execution time is slightly slower compared to a normal PC, the project successfully demonstrates an efficient and portable real-time face recognition system for security and attendance applications.

[7] Praneeth D, Rekha Sree T, Viswanatha V, M Abhiram Sai, “AI Enabled Home Security System Using Object Detection and Face Recognition with Haar Cascade Algorithm” ,2024.

This paper presents an AI-enabled smart home security system using face recognition and object detection with the Haar Cascade algorithm. The

system uses Raspberry Pi 4, IoT technology, cameras, sensors, and email alerts to improve home security and automation. When a known person is detected, the door unlocks automatically, while unknown persons trigger an email alert with their photo to the homeowner. The system also includes automatic light and fan control, fire and smoke detection, and energy-saving features. One major advantage of the system is its fast response, low computational requirement, and easy implementation using OpenCV and Raspberry Pi. The model performs well under different lighting conditions and provides better convenience and security for smart homes. However, the Haar Cascade algorithm gives lower accuracy compared to advanced deep learning models like CNN and FaceNet. The study concludes that the system is effective for basic smart home security applications, but future improvements can use advanced deep learning techniques for better accuracy and performance.

[8]Viswanatha V, Athreya Hebbar ,Nishant Anand, Ashish Kumar, “AI-Enabled Smart Vehicle with CNN and Deep Face for Anti-Theft Detection System”,2024.

This paper explains a smart vehicle security system that uses AI and face recognition technology to prevent unauthorized access. The system is developed using Raspberry Pi and a camera that captures the driver’s face before the vehicle starts. Facial identification methods are used to detect and compare the captured face with stored user data. If the face matches the authorized database, the vehicle is allowed to start; otherwise, access is blocked and a security alert may be sent to the owner. The system improves vehicle safety by allowing only authorized users to operate the vehicle and helps reduce theft and misuse. Overall, it provides a reliable and intelligent solution for modern vehicle security applications.

[9]Dr. Prashant Ranjan et al ,“TensorFlow, Computer Vision, and Edge AI-based Face Recognition Attendance Monitoring System Using Raspberry Pi 5”, 2025.

This paper explains a smart attendance system developed using Raspberry Pi 5 and Edge AI with face recognition technology. The system automatically records attendance by detecting and identifying student faces through a camera, reducing the problems

found in traditional roll-call methods. All processing is performed directly on the device, which improves privacy and avoids the need for internet connectivity. AI-based face recognition models are used for accurate student identification, and attendance is marked automatically after successful face matching. The system is fast, contactless and easier to use compared to fingerprint or QR-code methods. It also works effectively under different conditions, although factors like face covering and power issues may affect performance. Overall, the system provides a simple, affordable and efficient solution for modern attendance management.

[10] Gokul Prasad C et al, “A DT Framework based AI Robot For Facial Recognition using IoT”, 2023.

This paper explains an AI-based robot system that can recognize human faces and emotions using IoT technology. The system captures images through a camera and analyzes facial expressions to improve interaction between humans and robots. Raspberry Pi and image processing methods are used for face analysis and data communication through IoT. Compared to traditional systems, this method provides faster processing, better efficiency and live monitoring capabilities. The system can be applied in areas such as security, healthcare and smart home environments. Overall, it offers a smart and automated approach for face and emotion recognition, making technology more interactive and useful in everyday life.

[11] Justin Paul Kolengadan et al, “Development of Interactive Visual Recognition Assistant Bot”,2023.

This paper explains the development of an interactive assistant robot that can communicate with people and recognize them using face recognition technology. The main objective of the system is to support elderly people and patients by helping them with daily activities and providing companionship. The robot can understand voice commands, give reminders and answer user questions for better interaction. It uses voice processing and image recognition techniques to identify users and respond accordingly. The system is developed using Raspberry Pi along with a camera, microphone and speaker for communication and

monitoring. It can recognize users and provide personalized interaction based on user input. The system is mainly useful in healthcare and home environments, although lighting conditions and face angles may affect performance. Overall, it provides a smart and supportive solution for human assistance applications.

[12] Lambert Igene et al , “Development of Novel 3D Face Masks for Assessment of Face Recognition Systems”,2024.

This paper explains the security risks present in face recognition systems by testing how fake faces such as 3D masks can fool authentication systems. The researchers developed different low-cost masks, including 3D printed models, to examine how easily face recognition systems can be misled. The results showed that realistic masks can sometimes confuse the system and affect identification accuracy. Different detection methods were also tested, but none of them completely identified all fake faces correctly. The study highlights that current face recognition systems still face security challenges, and using more diverse training data and improved detection methods can help make these systems safer and more reliable.

[13] Alauddin Al-Omary, Maryam Mohammad Akram, Vidya Dhamodharan,” Design and Implementation of Intelligent Socializing 3D Humanoid Robo”,2021.

This paper explains a security attack called DREAM, which can fool 3D face recognition systems using only a single 2D face image. The method reconstructs a 3D face model from a normal photo and uses deep learning and GAN techniques to bypass face authentication systems. The main advantage of this method is that attackers do not need expensive 3D scanners because a simple RGB image is enough to generate fake depth information. The system achieved high attack success rates, especially in authentication models. However, the method also has some limitations because it depends on deep learning models, multiple training stages, and repeated optimization processes, which increase complexity and computational cost. The attack may also fail when the target system has strong security measures or

limited query attempts. Overall, the research shows that even advanced 3D face recognition systems can still be vulnerable to reconstruction-based spoofing attacks.

[14] Jonathan, Adhi Kusnadi, Daud Julio, "Security System With 3Dimensional Face Recognition Using PCA Method and Neural Networks Algorithm",2017.

This paper presents RUBEX, an intelligent 3D humanoid robot designed using 3D printing, AI, and machine learning techniques. The robot can recognize faces, detect emotions, answer questions, track objects, and interact with people using voice communication. RUBEX was designed to behave in a friendly and human-like way using sensors, cameras, microphones, and servo motors. The system worked successfully in real-time testing, but future improvements are needed to make the robot more intelligent and more human-like in movement and interaction.

[15] Zhihan Xu, Shuja Ansari, Amir M. Abdulghani, Muhammad Ali Imran, Qammer H. Abbasi, "IoT Enabled Smart Security Framework for 3DPrinted Smart Home",2020.

This paper presents an IoT-enabled smart security framework for a 3D printed smart home. The system combines sensors, facial recognition, surveillance cameras, door lock mechanisms, Arduino, and Raspberry Pi to improve home security. The smart home can detect motion, monitor temperature, humidity, air quality, and recognize faces using the LBPH facial recognition algorithm. Residents can unlock the door using fingerprints or passwords after successful face recognition. The surveillance system records videos based on the confidence level of recognition. The main advantage of this system is the integration of security, monitoring, and automation into a single smart home framework. However, the facial recognition system is sensitive to lighting conditions and does not always provide very high accuracy. The project was successfully implemented and tested, but future improvements are needed for better recognition performance and advanced smart home features.

[16] Shizong Yan, Huixiang Wen, Shan Chang, "Closed-Box 3-D Face Reconstruction Attack on Face Recognition From a Single Image", 2025.

This research presents a 3D face recognition security system using PCA and Neural Network algorithms. The system uses a Kinect Xbox One camera to capture facial depth data for authentication. PCA helps in extracting important facial features, while the neural network recognizes the face accurately. The system provides better security than normal 2D face recognition because printed photos cannot easily fool it. It also works in low-light conditions and uses low-cost hardware. The system achieved about 95% accuracy with faster recognition speed. However, the training process takes more time, and the performance depends on the learning rate and hidden nodes used in the neural network.

III. LITERATURE SURVEY SUMMARY

Several recent studies have explored AI-based face recognition for smart security applications using Raspberry Pi and IoT platforms. Most implementations combine CNN or PCA with Haar Cascade for face detection, while neural networks handle recognition and classification. These systems are commonly deployed for smart home security, attendance tracking, surveillance, and access control. In practice, the system can identify registered users, flag unknown faces, trigger alerts, and automate actions such as unlocking doors or marking attendance in real time. The main strengths are cost-effectiveness, ease of deployment, and reasonably good accuracy for small-scale setups. That said, reliability often drops under poor lighting, extreme face angles, or with low-resolution cameras. Hardware constraints of Raspberry Pi also limit processing speed when handling multiple faces simultaneously. While the results show that AI-driven face recognition is practical for low-cost security solutions, further work is needed to improve robustness and reduce false positives in real-world environments.

IV. CONCLUSION

This review shows that AI-based face recognition system integrated with IoT provides a practical and

cost-effective solution for smart and security applications such as attendance monitoring, access control and surveillance. The use of CNN for face recognition enables real-time performance with low-cost hardware, but making it suitable only for small-scale deployments. However, the system's accuracy and reliability are still dependent on environmental factors like lighting and camera quality as well as the processing constraints of Raspberry Pi. Addressing these challenges through optimized algorithms and improved hardware will be key to enhancing robustness. Overall, the current implementations demonstrate promising results for basic security needs, further research is required to improve accuracy and ensure consistency in real time.

REFERENCES

- [1] S. S. Girajala, D. Putty, V. N. S. K. Cherukuri and S. Chinnadurai, "Dual-Model Biometric Authentication Using Face and Voice Recognition", 2025.
- [2] D. L. A. P. Shetty, S. S. Shetty, V. K. S. M. M. V and Fairooza, "A Multimodal Biometric Approach for Physical Locker Security: Integration of Face and Voice Recognition", 2025.
- [3] S. Venkatesha, V. V. Hiremath, S. Rani G, C. M. Sharvani and C. H. Suvarna Shree, "Voice E-Mailer System with Face Recognition using Machine Learning Techniques", 2025.
- [4] [4]S. Supriya, K. Madhu, M. Sreekrishna, K. Durga Prasad and G. S. Gopika, "A Data Science-Driven Face Recognition Attendance System Integrated with a Voice Assistant", 2025.
- [5] Le Yang, WeiTang,Xihu Wang, "Design and Implementation of a Multimodal Face Recognition Access Control System Based on Raspberry Pi 4B", 2025.
- [6] BOUKERCHE Alaa Eddine, CHELALI Fatma Zohra, "Real Time Face Recognition System using Raspberry Pi4", 2024.
- [7] Praneeth D, Rekha Sree T, Viswanatha V,M Abhiram Sai, "AI Enabled Home Security System Using Object Detection and Face Recognition with Haar Cascade Algorithm", 2024.
- [8] Viswanatha V, Athreya Hebbar, Nishant Anand ,Ashish Kumar, "AI-Enabled Smart Vehicle with CNN and Deep Face for Anti-Theft Detection System", 2024.
- [9] Dr. Prashant Ranjan et al, "TensorFlow, Computer Vision, and Edge AI-based Face Recognition Attendance Monitoring System Using Raspberry Pi 5", 2025.
- [10] Gokul Prasad C et al, "A DT Framework based AI Robot For Facial Recognition using IoT", 2023.
- [11] Justin Paul Kolengadan et al, "Development of Interactive Visual Recognition Assistant Bot", 2023.
- [12] Lambert Igene et al , "Development of Novel 3D Face Masks for Assessment of Face Recognition Systems", 2024.
- [13] Alauddin Al-Omary, Maryam Mohammad Akram, Vidya Dhamodharan, "Design and Implementation of Intelligent Socializing 3D Humanoid Robo", 2021.
- [14] Jonathan, Adhi Kusnadi, Daud Julio, "Security System With 3Dimensional Face Recognition Using PCA Method and Neural Networks Algorithm", 2017.
- [15] Zhihan Xu, Shuja Ansari, Amir M. Abdulghani, Muhammad Ali Imran, Qammer H. Abbasi, "IoT Enabled Smart Security Framework for 3D Printed Smart Home", 2020.
- [16] Shizong Yan, Huixiang Wen, Shan Chang, "Closed-Box 3-D Face Reconstruction Attack on Face Recognition From a Single Image", 2025.