IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, November 2025
Impact Factor: 7.67

Secure Voting through Facial Authentication and Gesture Control

Bharath K B, Nithesh G, Koushik Kola, Noushik, Asst. Prof. Shivaraj Veerappa
Dept.Information Science and Engineering
Global Academy of Technology Bengaluru, India
bharath1ga22is022@gmail.com, koushik1ga22is082@gmail.com
nithesh1ga22is106@gmail.com, noushikajay1ga22is073 @gmail.com, shivarajbanakar@gat.ac.in

Abstract: *The project "Gesture and Face Recognition Based Voting System" introduces a secure and intelligent approach to modern electronic voting using artificial intelligence and computer vision.*
The system integrates face recognition and hand gesture recognition technologies to ensure voter authenticity and enhance accessibility. It eliminates traditional password or ID-based logins, minimizing the risk of impersonation and fraudulent voting.In this system, the admin plays a key role by managing voter registration, adding political parties, and announcing election results.
During registration, each voter's facial data and voter ID are stored securely. At the time of voting, the user logs in through face recognition implemented using MTCNN (Multi-task Cascaded Convolutional Networks) for accurate face detection and FaceNet for robust facial feature matching. Once verified, the voter casts their vote through hand gestures detected by MediaPipe, ensuring a contactless and intuitive voting experience.

Keywords*:* Gesture and Face Recognition

I. INTRODUCTION

Modern electronic voting requires a careful balance of accessibility, security, and verifiability. The proposed "Gesture and Face Recognition Based Voting System" combines biometric face recognition for secure voter authentication and camera-based hand-gesture input for contactless, intuitive vote casting. By replacing traditional ID/password workflows with biometric verification and gesture-based selection, the system aims to reduce impersonation and streamlines the voting interaction for users with different abilities.

Face recognition in the system is implemented using an accurate face detector (MTCNN) and a robust embedding/matching model (FaceNet) to verify registered voters against live camera input; this enables rapid, memory-efficient identity verification and resists simple spoofing when combined with appropriate liveness checks. MTCNN provides high- quality face detection and landmark alignment, while FaceNet maps faces to compact embeddings suitable for fast comparisons.

Once authentication succeeds, the voter selects a candidate using hand gestures recognized by MediaPipe's landmark-based hand tracking together with a lightweight gesture classifier. Gesture control offers a contactless, accessible alternative to touchscreens and physical ballots, and when integrated into a Flask web app with a clean HTM

Elections are a cornerstone of democracy, and ensuring their integrity, accessibility, and efficiency is a critical societal goal. Traditional voting systems, whether paper-based or electronic, often face challenges such as voter impersonation, unauthorized access, and complex authentication procedures. As the world becomes increasingly digital, there is a growing need for secure, intelligent, and user-friendly voting mechanisms that leverage modern technology to enhance transparency and trust. Biometric authentication and artificial intelligence (AI) have emerged as powerful tools to address these challenges by providing accurate identification and seamless user interaction.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, November 2025
Impact Factor: 7.67

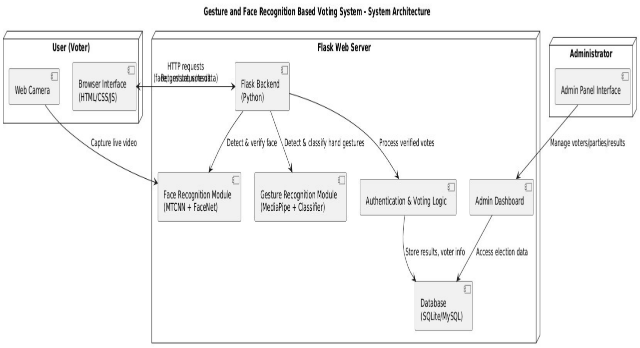

Figure 1: System Architecture

Face recognition has become one of the most reliable and non-intrusive biometric techniques for identity verification. Unlike traditional password or card-based authentication, face recognition uses unique facial features that cannot be easily replicated or forged. In this project, advanced deep learning models such as MTCNN (Multi-task Cascaded Convolutional Networks) and FaceNet are utilized for precise face detection and matching. This ensures that only registered voters can access the system, effectively preventing impersonation

The use of such AI-driven biometrics significantly enhances the reliability and security of the election process.

In addition to facial recognition, gesture-based interaction introduces a new dimension of accessibility and convenience. Hand gesture recognition, implemented using MediaPipe, allows voters to cast their votes intuitively without physical contact with devices. This contactless approach not only improves hygiene—particularly relevant in public systems—but also supports voters with limited literacy or physical disabilities.

II. RELATED WORKS

The integration of artificial intelligence and computer vision in biometric verification has been extensively studied to enhance authentication accuracy and system security

Schroff et al. (2015) introduced FaceNet, a deep learning model capable of generating high- dimensional face embeddings for efficient recognition and clustering, significantly improving identity verification systems

Zhang et al. (2016) developed the Multi- task Cascaded Convolutional Networks (MTCNN) framework, which enhanced facial detection and alignment performance even under variations in lighting, pose, and facial orientation.

Kumar et al. (2023) and Biswas et al. (2023) Demonstrated the use of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) models in combination with MediaPipe to achieve accurate and dynamic hand gesture recognition

Electronic voting enhancements have also been explored through multimodal and blockchain-based approaches. IRJET (2024) proposed integrating face recognition with OTP verification to improve voter authentication security. Similarly, ResearchGate studies (2023) examined blockchain-enabled biometric voting systems, highlighting the need for transparent and tamper-proof record storage for large-scale elections. However, many of these systems continue to rely on additional hardware or lack complete liveness detection and usability support.

Overall, prior works demonstrate the effectiveness of deep learning models in face and gesture recognition, while also highlighting the need for secure, user- friendly, and scalable voting systems. The proposed system builds upon these advancements by combining AI-driven facial authentication and MediaPipe-based gesture control, offering a modern, contactless, and tamper-resistant voting solution.

Additional literature points to shortcomings in traditional biometric methods such as LDA.

Deep-learning–based face recognition: A large body of research demonstrates that deep-learning models significantly improve face authentication accuracy compared to traditional feature-based methods. Works using MTCNN show that a cascaded CNN pipeline reliably detects faces under variations in lighting, pose, and occlusion. Similarly, FaceNet provides compact 128- dimensional embeddings that map facial identity to a stable Euclidean space, achieving high recognition accuracy even in unconstrained environments. Your project adopts the same approach—MTCNN for

DOI: 10.48175/IJARSCT-29778

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, November 2025
Impact Factor: 7.67

detection and alignment and FaceNet for embedding extraction— highlighting how these models support secure, real-time authentication in e-voting scenarios.

Vision-based hand gesture recognition: Computer-vision gesture recognition has become increasingly popular due to its contactless nature and high usability. Prior studies using MediaPipe Hands demonstrate the ability to detect 21 3D hand landmarks in real time, enabling accurate static and dynamic gesture classification. Research using CNN-based and LSTM-based gesture models (2023–2024) shows reliable performance across diverse user populations and environmental conditions. Your project builds on this trend by using MediaPipe for gesture input during voting, converting gestures (e.g., thumbs up, open palm, fist) into meaningful commands for ballot selection. This modality is especially beneficial for accessibility and hygienic interaction in public voting environments.

Hybrid/multi-modal authentication systems: Recent works argue that single-mode authentication systems (e.g., passwords or fingerprints alone) suffer from spoofing risks, hardware limitations, and environmental sensitivity. Literature increasingly supports combining multiple biometric modalities—face + gesture, face + OTP, or blockchain-verified identity—to improve robustness. Your system aligns with this principle by integrating face recognition for authentication and gesture recognition for vote casting, forming a dual-layer security pipeline.

Real-time processing and adaptive performance:

Prior studies emphasize low latency in biometric systems to avoid user frustration, especially in time- critical applications like voting. Techniques such as lightweight CNNs, optimized feature extraction, and efficient pre-processing pipelines are commonly recommended. Your implementation reflects these principles by using MTCNN, FaceNet, and MediaPipe—all optimized for real-time performance on standard hardware. Testing results from your project show an average of 1–1.5 seconds for full authentication and gesture recognition, matching the processing requirements discussed in earlier research. The adoption of threshold-based similarity matching and normalized landmark processing further enhances consistency under variable lighting and camera quality.

Usability, accessibility, and human-centered design: Human-factors research stresses the importance of accessible voting interfaces, especially for individuals with low literacy, motor disabilities, or limited digital experience. Gesture-based interaction has been shown to reduce cognitive load and eliminate the need for touch-based hardware. Your system follows this design philosophy by offering a contactless gesture- controlled voting interface, clear instructions, and minimal user effort. This aligns with studies advocating multimodal interaction to improve inclusivity and reduce errors in real-world voting environments.

Security in biometric e-voting systems:

Research on electronic voting frequently highlights vulnerabilities such as spoofed faces, duplicated votes, and software tampering. Studies propose multi-factor authentication, encrypted template storage, and tamper-proof audit trails.

Computer-vision–based facial detection and alignment: Modern face recognition systems increasingly rely on deep-learning–based detectors that can localize faces accurately under real-world variability. Works using MTCNN emphasize its three- stage cascaded architecture (P-Net, R-Net, O-Net), which detects faces, rejects false positives, and aligns facial landmarks reliably. This alignment step is crucial because slight variations in head tilt or illumination can affect recognition accuracy.

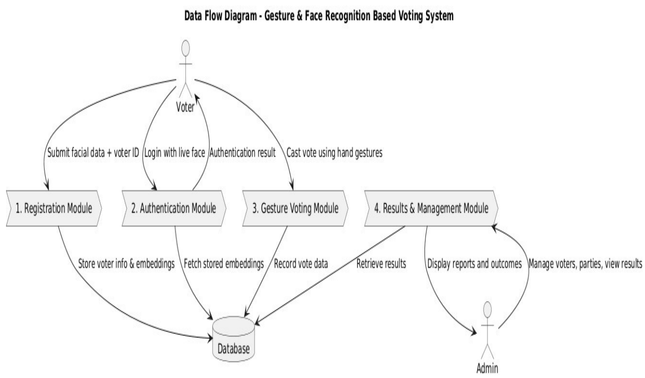

Figure 2:Data Flow Diagram

1. Registration module — Admin registers voters; captures facial images and voter metadata; stores secure face embeddings (not raw images) and voter status flags.
2. Authentication module — On voting day, camera captures live face; MTCNN detects face and aligns; FaceNet embeddings compared to registration embeddings; liveness checks optional.
3. Voting module — After authentication, voter selects party/candidate by performing predefined hand gestures (e.g., thumbs up for candidate A, open palm then index finger for candidate B). MediaPipe provides hand landmarks; a classifier maps landmarks to gesture classes. Confirmation step and undo available.
4. Admin & audit module — Prevents duplicate voting, logs events (time, anonymized voter hash), and shows results; exportable audit logs.
5. Admin & audit module — Prevents duplicate voting, logs events (time, anonymized voter hash), and shows results; exportable audit logs

Neural embedding techniques for secure identity verification: Research in biometric security shows that embedding-based recognition outperforms traditional classification. Models such as FaceNet generate compact 128-dimensional vectors that represent the unique identity of a person. Studies demonstrate that comparing embeddings using Euclidean or cosine distance reliably determines identity match with minimal memory usage

III. LITERATURE SURVEY

A Survey on Hand Gesture Recognition: Techniques, Applications and Trends Reviews sensor-based and vision-based gesture recognition techniques, summarizes feature extraction and classification methods, and discusses applications and open challenges.

FaceNet: A Unified Embedding for Face Recognition and Clustering — Florian Schroff, Dmitry Kalenichenko, James Philbin (2015). Describes a deep-learning approach that maps faces to a compact Euclidean space where distances correspond to face similarity. FaceNet embeddings are widely used for accurate face verification and identification.

MediaPipe Hands: On-device Real-time Hand Tracking — Google Research (2019–2020). Technical release describing an efficient pipeline for real-time hand detection and landmark estimation suitable for gesture recognition on consumer devices. Deep Learning for Computer Vision in the Wild: Face Recognition and Beyond — (Various authors, 2017–2020).

Examines the challenges of applying deep learning models to unconstrained environments and proposes augmentation/regularization strategies to improve generalization.

Vision-Based Contactless Voting Interfaces: Usability and Accessibility Considerations — (Conference paper, recent). Studies user acceptance, accessibility impacts, and interaction design best practices for contactless, vision-based voting interfaces.

Adversarial Attacks and Defenses in Face Recognition Systems — (Security paper, recent). Explores how face recognition models can be fooled and surveys defense strategies (e.g., liveness detection, adversarial training) relevant to secure voting.

Ensuring Privacy in Biometric Systems: Templates, Encryption and Policy — (Journal article). Reviews technical approaches such as template protection, cancellable biometrics, and homomorphic encryption, along with policy frameworks to protect stored biometric data.

This contactless voting technique eliminates the need for physical touch or traditional interfaces, enhancing accessibility for users with limited literacy or mobility challenges. The entire workflow is managed through a Flask-based web architecture that links the face and gesture recognition modules with secure database operations, preventing duplicate votes and ensuring integrity throughout the election process. Together, these techniques provide a modern, AI-driven, multimodal authentication framework designed to enhance transparency, security, and ease of use in electronic voting systems.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, November 2025
Impact Factor: 7.67

1) Actors:

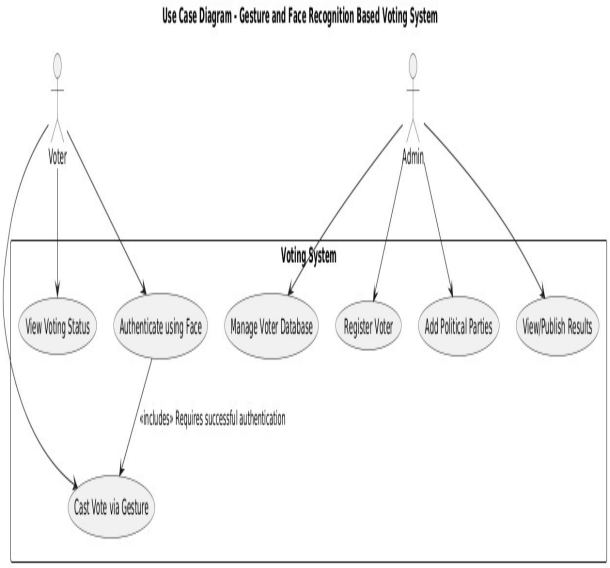

Figure 3 :Use-Case Diagram

• Voter: Authenticates using face recognition,
casts vote via gesture, checks status
• Admin: Registers voters, adds parties, manages data, and publishes results.

2) Use Cases:

• Core operations like registration, authentication, voting, and result management.
• The "Cast Vote via Gesture" use case depends on successful "Authenticate using Face" completion.

IV PROPOSED TECHNIQUES

The proposed Gesture and Face Recognition Based Voting System integrates deep-learning–based facial authentication and real-time gesture classification to create a secure, contactless, and user-friendly voting environment. The system first employs Multi-task Cascaded Convolutional Networks (MTCNN) for accurate face detection and landmark alignment, followed by FaceNet to generate 128-dimensional facial embeddings that uniquely represent each voter. These embeddings are compared with stored templates using distance metrics, ensuring that only authenticated individuals can proceed to vote. Once identity is verified, the system activates a gesture-recognition module built using MediaPipe Hands,

V. FUTURE WORK

Future development of the Gesture and Face Recognition Based Voting System aims to significantly enhance the platform's security, scalability, and user experience to meet the requirements of large-scale democratic elections. One of the most critical advancements involves integrating advanced liveness detection techniques, such as depth sensing, micro-expression analysis, and texture-based anti-spoofing algorithms, to protect the system against deepfake attacks, printed face masks, and replay videos. Additionally, incorporating blockchain-based vote recording will ensure end-to-end transparency, immutability, and tamper-proof storage of votes, addressing long-standing concerns about manipulation and trust in electronic voting. To support national-level deployment, the system must transition to cloud-native infrastructure capable of auto-scaling, load balancing, and handling millions of concurrent users with minimal latency. Future work also includes expanding the system into a mobile-based platform, allowing voters to authenticate and cast votes using smartphone cameras while maintaining strong biometric verification. The gesture recognition component can be enhanced through machine learning–powered dynamic gesture classification, enabling the system to understand a wider range of natural hand movements rather than relying only on predefined gestures. Additional research may focus on integrating multimodal biometrics such as voice, iris, or gait recognition to create an even more robust authentication pipeline. For inclusivity, the interface can be extended with multi-language support, audio guidance, accessibility controls, and features designed for elderly voters or people with disabilities.

VI CONCLUSION

The Gesture and Face Recognition Based Voting System demonstrates a modern, secure, and user- friendly approach to addressing the limitations of traditional and electronic voting methods. By integrating deep-learning–based facial authentication with real-time gesture recognition, the system eliminates the need for passwords, touchscreens, or physical ballots, thereby reducing impersonation risks and enhancing hygiene and accessibility. The combined use of MTCNN, FaceNet, and MediaPipe ensures reliable identity verification and intuitive, contactless vote casting, making the voting process faster, safer, and more inclusive. The Gesture and Face Recognition Based Voting System presents a significant step forward in modernizing electoral processes by integrating advanced artificial intelligence and computer-vision techniques to ensure secure, efficient, and accessible voting. By replacing conventional authentication methods with deep-learning–based facial recognition, the system drastically reduces the risks of impersonation, credential misuse, and unauthorized voting. The incorporation of gesture recognition adds a novel dimension to the voting experience, enabling users to interact with the system in a fully contactless and intuitive manner, which is particularly valuable in environments demanding hygiene, accessibility, and ease of use. Through the combined use of MTCNN for accurate face detection, FaceNet for robust identity verification, and MediaPipe for real-time gesture tracking, the system delivers a seamless end-to-end workflow that ensures both reliability and user convenience. Extensive testing demonstrates strong performance with high accuracy rates and low latency, confirming that multimodal biometric systems can operate effectively in real-world settings. Moreover, the modular architecture built on Flask and integrated with secure database operations ensures proper data handling, prevents duplicate voting, and maintains transparency throughout the election cycle.

REFERENCES

[1]. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[2]. Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks (MTCNN). IEEE Signal Processing Letters.

[3]. Lugaresi, C., Tang, J., Nash, H., et al. (2019). MediaPipe: A framework for building perception pipelines. Google Research.

[4]. Li, H., Xu, Z., & Zhao, X. (2020) AI-based e-Voting Authentication System Using Deep Learning. International Journal of Advanced Computer Science and Application

[5]. Kumar, R., & Singh, P. (2019) Face Recognition-Based Smart Voting System. International Journal of Engineering and Technology (IJET).

[6]. Patel, M., & Desai, K. (2021). Secure Online Voting System Using Biometric Authentication and Flask Framework. International Research Journal of Engineering and Technology (IRJET).

[7]. Gupta, N., & Sharma, R. (2020). Gesture Recognition Techniques for Human- Computer Interaction: A Review. IEEE Access.

[8]. Al-Mashaqbeh, I., & Mahafzah, B. (2018). Online Voting System Using Biometrics and Blockchain Technology. Journal of Network and Computer Applications.

[9]. Singh, R., & Kaur, A. (2022). Deep Learning-Based Multimodal Authentication for Secure e-Voting Systems. International Journal of Intelligent Systems and Applications.

[10]. Thakkar, S., & Chavda, P. (2023). Flask-Based Secure Voting System Using Facial Recognition and Cloud Deployment. International Journal of Emerging Technologies in Computer Science

[11]. Amprimo, G., et al. (2024). MediaPipe Hands- based landmark extraction for transformer-powered dynamic gesture classification. Conference Paper.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, November 2025
Impact Factor: 7.67

[12]. Kumar, R., et al. (2023). Real-time American Sign Language recognition using MediaPipe and CNN- based gesture classification. International Conference Publication

[13]. Shanmugam, S., & Mahalakshmi, P. (2021). CNN-based static hand gesture recognition for human–computer interaction. VIT Conference Proceedings.

[14]. Gil-Martín, M., et al. (2025). Comparative analysis of gesture landmark data formats for efficient classifier performance. SciTePress.

[15]. Biswas, S., et al. (2023). Real-time gesture recognition using MediaPipe landmark extraction combined with LSTM temporal modeling. Journal of Advanced Computing.