International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1131-1137
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0160

AuraVote: A Secure, Decentralized Voting System with AI-Powered Liveness and Identity Verification Using ArcFace and Blockchain

Shruthi D V[1], Mohammed Aazam Sheikh H[2], Saniya Banu[3], Varsha A S[4], Shreyas M K[5]

[1]Assistant Professor, Department of Information Science and Engineering, Malnad College of Engineering, Hassan-573202, Karnataka, India

[2,3,4,5]Students, Department of Information Science and Engineering, Malnad College of Engineering, Hassan-573202, Karnataka, India

Emails: *dvs@mcehassan.ac.in[1], mohammedaazam757@gmail.com[2], saniyabanubvr@gmail.com[3], asvarsha91@gmail.com[4], shreyasarya707@gmail.com[5]*

Abstract

Secure and transparent electronic voting remains a major challenge due to vulnerabilities in centralized infrastructures and insufficient voter authentication methods. Existing systems frequently depend on weak biometric checks or static credential verification, making them susceptible to spoofing, identity duplication, and manipulation of centralized databases. AuraVote introduces a next-generation decentralized voting framework that integrates real-time artificial intelligence with blockchain-based immutability to guarantee voter authenticity and system integrity. The proposed model adopts a robust architecture including RetinaFace for face localization, ArcFace for generating highly discriminative 512-dimensional embeddings, and cosine-similarity-based identity verification. Alongside facial verification, the system executes multi-modal liveness detection incorporating blink dynamics, natural micro-movement estimation, and texture-frequency analysis to counter photo, video, and digital screen presentation attacks. A FastAPI-based verification backend processes live video streams from the browser and communicates verification results to a Next.js decentralized application. Verified users cast votes on an Ethereum Proof-of-Authority (PoA) blockchain using MetaMask, ensuring tamper-proof vote recording and cryptographically signed voter participation. This paper details the architecture, AI algorithms, blockchain workflow, system security properties, and implementation strategy, demonstrating how combining modern deep-learning authentication with decentralized ledgers establishes a secure, scalable, and trustworthy e-voting ecosystem.

Keywords: *ArcFace; Blockchain; Decentralized Voting; FastAPI; Liveness Detection;*

1. Introduction

Electronic voting has emerged as a critical technological requirement in modern democratic ecosystems, yet the transition from traditional paper ballots to digital platforms has exposed several unresolved challenges. Centralized servers used in conventional e-voting solutions are vulnerable to database tampering, unauthorized access, and single-point failures (Rivest, 2008) [1]. Moreover, existing authentication mechanisms often limited to ID numbers, passwords, or low-accuracy biometric checks cannot reliably prove that a vote originates from a legitimate, physically present individual. These limitations create opportunities for identity misuse, double voting, and large-scale manipulation (Cortier et al., 2020) [2]. Blockchain technology has been proposed as a solution for ensuring transparency, immutability, and auditability in electoral systems (Zheng et al., 2018) [3]. While blockchain prevents post-recording manipulation of votes, it does not inherently validate the identity of the voter. A malicious actor can still cast fraudulent votes by generating multiple wallet addresses or performing Sybil attacks (Douceur, 2002) [4]. Thus, the central research challenge remains: how can we

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1131-1137
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0160

guarantee that each blockchain vote is cast by a unique, verified, live human being? AuraVote addresses this problem by designing a hybrid voting architecture that binds real-time biometric verification with on-chain voting authorization. Unlike earlier approaches relying on MTCNN-based detection or FaceNet embeddings (Schroff et al., 2015) [5] with limited robustness under challenging lighting, angles, or spoof conditions, AuraVote incorporates RetinaFace (Zhang et al., 2020) [6] for high-precision face detection and ArcFace (Deng et al., 2019) [7] for 512-dimensional embedding generation. These models significantly improve identity discrimination and reduce false acceptance rates (Wang et al., 2018) [8]. AuraVote integrates a multi-level liveness detection module that analyzes blink sequences, involuntary micro-head movements, and texture irregularities in captured video frames (Li et al., 2019) [9]. This prevents printed photographs, video replays, and deepfake-based presentation attacks from passing verification (Tolosana et al., 2020) [10]. A FastAPI backend (Ramirez, 2020) [11] processes continuous frame streams, performs AI inference, and returns real-time verification scores to the browser interface developed using Next.js and ShadCN UI. Votes are recorded on an Ethereum Proof-of-Authority (PoA) private blockchain (Wood, 2014) [12], where authorized validators ensure rapid block confirmation and protection against external manipulation (Antonopoulos & Wood, 2018) [13]. The major contributions of this work are as follows:
(1) A redesigned AI verification pipeline using RetinaFace detection, ArcFace 512-D embeddings, and cosine similarity matching. (2) A multi-modal liveness detection module combining behavioral and texture-based cues. (3) A FastAPI-based high-performance backend for real-time biometric verification with JWT-secured sessions. (4) A decentralized voting architecture using PoA blockchain, Solidity smart contracts, and MetaMask. (5) A modern and secure Next.js dApp that delivers guided biometric verification and tamper-proof vote casting.

2. Related Work

Several prior works have explored e-voting and biometric authentication independently. Rivest [1] established foundational principles for software-independent voting systems, emphasizing the need for auditable and verifiable election outcomes. Subsequent blockchain-based voting proposals by Zheng et al. [3] demonstrated that distributed ledgers can provide tamper-resistant vote storage, but lacked robust voter identity verification mechanisms. In the domain of face recognition, Schroff et al. [5] introduced FaceNet, which generates 128-dimensional embeddings using triplet loss. While effective for general recognition tasks, FaceNet exhibits limited angular discriminability when identities are visually similar. Deng et al. [7] addressed this limitation through ArcFace, which introduces an additive angular margin penalty to maximize inter-class separation. Wang et al. [8] proposed CosFace with a cosine margin approach achieving comparable improvements. For face detection, MTCNN (Zhang et al., 2016) [14] has been widely adopted but shows degraded performance under low illumination and extreme head poses. Zhang et al. [6] proposed RetinaFace, a single-stage dense face localisation method achieving superior accuracy while simultaneously predicting facial landmarks. Anti-spoofing and liveness detection were studied by Li et al. [9], demonstrating that texture-based frequency analysis can reliably distinguish live faces from printed or screen-displayed imposters. Recent work on photoplethysmography (PPG) based liveness (Liu et al., 2021) [15] showed that subtle skin colour variations due to heartbeat can serve as a strong biological anti-spoofing signal.

3. System Architecture

AuraVote adopts a multi-layered architecture combining AI verification, blockchain infrastructure, and a secure browser-based interface. The architecture encompasses four major components: (1) AI Verification Service, (2) Frontend dApp Interface, (3) MetaMask Wallet, and (4) PoA Blockchain Network.

3.1 AI Verification Service

The AI module performs high-precision identity verification and liveness analysis. Face detection uses RetinaFace for robust bounding box generation and five-point landmark extraction. ArcFace generates 512-dimensional discriminative facial embeddings. Cosine similarity replaces older SVM classifiers,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1131-1137
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0160

providing a mathematically grounded distance metric in hyperspherical embedding space. Liveness detection combines blink detection, micro-movement analysis, and texture-based anti-spoofing into a multi-factor composite score

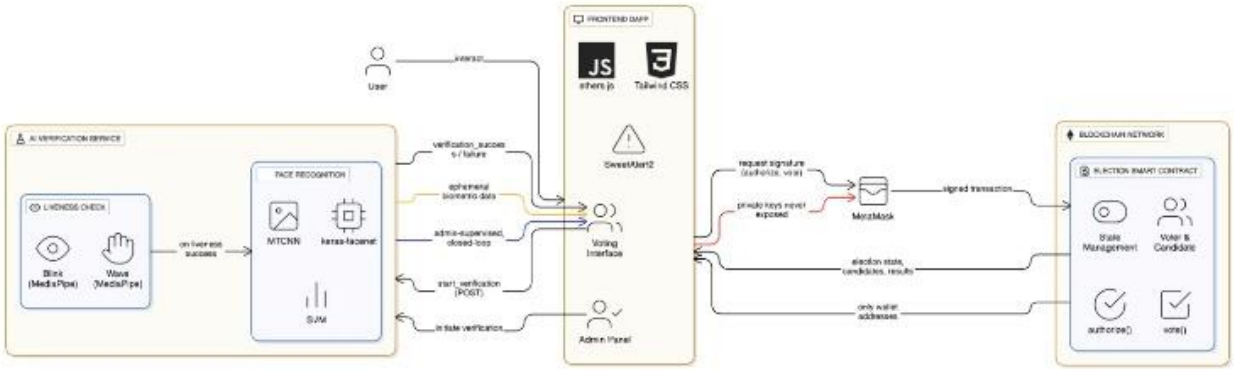

Figure 1 AuraVote System Architecture Showing the Interaction Between the Next.js Frontend, FastAPI AI Backend, MetaMask wallet, and PoA Blockchain Network

Table 1 AuraVote System Components and Roles

Component	Technology	Function
AI Verification	RetinaFace + ArcFace + FastAPI	Biometric identity & liveness
Frontend dApp	Next.js 14 + ShadCN UI	Voter interface & frame capture
Wallet	MetaMask + ethers.js	Transaction signing & key custody
Blockchain	Ethereum PoA + Solidity	Immutable vote recording

3.2 Frontend dApp Interface

The browser-based dApp is built on Next.js 14 with the App Router paradigm. React Webcam captures live video at 6-10 frames per second, with each frame Base64-encoded and transmitted to the FastAPI backend via secure HTTPS fetch calls. The frontend unlocks the voting interface only after receiving a verified signed session token from the backend.

3.3 MetaMask Wallet and Blockchain

MetaMask provides transaction signing, secure key management, and blockchain interaction through ethers.js. Private keys never leave the voter's device. The PoA network is configured with two or more authority nodes, a five-second block time, and zero gas fee, ensuring predictable transaction finality and resistance to public-chain congestion.

4. Implementation Details

4.1 Frame Preprocessing

Incoming RGB frames from the browser webcam are resized to 640 x 480, converted to BGR format, and normalized to the range [0, 1]. Histogram equalization is selectively applied to enhance contrast in low-light environments. Each frame I undergoes Z-score standardization:

$$I_{norm} = \frac{I - \mu(I)}{\sigma(I)} \quad (1)$$

This normalization improves ArcFace embedding stability under variable exposure conditions. A temporal buffer of the last K = 15 frames is maintained to compute micro-movement and blink statistics.

4.2 Blink Pattern Recognition

Eye landmarks are extracted from RetinaFace's five-point output (left eye, right eye, nose, left mouth, right mouth). The vertical Eye Aspect Ratio (EAR) for each frame is computed as:

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|} \quad (2)$$

A blink is recorded when EAR drops below the threshold $\tau_{blink} = 0.21$ for at least three consecutive frames. The blink frequency $B = N_{blinks} / \Delta t$ is

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1131-1137
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0160

normalized to the physiological range and a blink confidence score is:

$$B_s = 1 - |B - B_{human}| \quad (3)$$

where B_human = 15 blinks/min represents the average physiological blink rate.

4.3 Micro-Movement Analysis

AuraVote tracks micro-displacements of the nose landmark in the temporal frame buffer. Natural head jitter produces non-zero inter-frame variance:

$$M = Var(\Delta x_t) + Var(\Delta y_t) \quad (4)$$

where $\Delta x_t = x_t - x_{\{t-1\}}$ and $\Delta y_t = y_t - y_{\{t-1\}}$. Values of M close to zero indicate spoofing via static images or replay videos.

4.4 Texture-Based Frequency Analysis

Spoof images exhibit smoothness and lack the natural texture spectrum found in real faces. AuraVote computes the sum of the 2-D FFT magnitude:

$$T = \Sigma|F(I)| \quad (5)$$

where F(I) is the 2-D FFT magnitude. A classification threshold distinguishes high-frequency real textures from flat spoofed ones produced by printed photographs or display screens.

4.5 Combined Liveness Score

The combined liveness confidence is a weighted sum of the three modality scores:

$$L = 0.4 \cdot B_s + 0.3 \cdot M + 0.3 \cdot T \quad (6)$$

Verification proceeds only when L ≥ 0.30.

4.6 ArcFace Identity Recognition

After face alignment using RetinaFace landmarks, each face is resized to 112 x 112 and fed into ArcFace, producing a 512-dimensional L2-normalized embedding:

$$f(x) \in \mathbb{R}^{512}, \quad \|f(x)\|_2 = 1 \quad (7)$$

Stored enrollment embeddings f_{reg} are compared to live embedding f_{live} via cosine similarity:

$$S = \frac{f_{reg} \cdot f_{live}}{\|f_{reg}\| \cdot \|f_{live}\|} \quad (8)$$

A threshold of S ≥ 0.35 signifies a positive identity match. The final verification score is:

$$V = 0.6 \cdot S + 0.4 \cdot L \quad (9)$$

Verification is confirmed when V ≥ 0.45, triggering issuance of a signed JWT session token.

4.7 ArcFace Loss Function

ArcFace applies an additive angular margin penalty to maximise inter-class separation in embedding space:

$$L = -\frac{1}{N}\sum_{i=1}^{N} log\left(\frac{e^{s\cdot cos\left(\theta_{y_i}+m\right)}}{e^{s\cdot cos\left(\theta_{y_i}+m\right)} + \sum_{j\neq y_i} e^{s\cdot cos(\theta_j)}}\right) (10)$$

where θ_{yi} is the angle between the embedding and the class centre, m is the additive angular margin, and s is the feature scale. This ensures embeddings of different identities are maximally separated in angular space.

4.8 FastAPI Backend

The backend is implemented using FastAPI with asynchronous request handling via Uvicorn and Gunicorn multi-worker deployment. ThreadPoolExecutor offloading manages CPU-heavy model inference. JWT-based session tokens contain session ID, user wallet address, server signature, and expiration timestamp, and are invalidated after a single vote attempt. API endpoints include POST /verify_frame, POST /finalize, GET /session_status, and GET /health.

Table 2 Average Processing Latency per Pipeline Stage (Intel i5 CPU)

Pipeline Stage	Latency (ms)
Frame Preprocessing	5–10
RetinaFace Detection	35–50
ArcFace Embedding Extraction	15–20
Liveness Computation	10–15
Total End-to-End per Frame	70–120

4.9 Blockchain Voting Workflow

The voting transaction lifecycle: (1) Voter completes AI verification and receives a signed session token; (2) Next.js unlocks the voting UI; (3) User selects a candidate; (4) Structured vote transaction vote(candidateId, sessionToken) is sent to MetaMask for signing; (5) MetaMask validates account ownership and private key control; (6) Signed transaction is propagated to the PoA network; (7) Validators confirm the block; (8) Smart contract marks the voter as having voted and increments the

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1131-1137
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0160

candidate counter. The contract enforces: require(hasVoted[msg.sender] == false, 'Already voted'), making duplicate submissions cryptographically and procedurally impossible.

5. Security Analysis

Table 3 Security Threat Analysis and AuraVote Mitigations

Threat	Attack Vector	Mitigation
Presentation Attack	Photo/screen/ deepfake	Multi-modal liveness (L≥0.30)
Identity Theft	Lookalike/synthetic face	ArcFace angular margin + S≥0.35
Double Voting	Multiple wallets/sessions	hasVoted mapping + single-use JWT
51% Attack	Validator majority control	Institution-controlled PoA validators
Replay Attack	Resubmit old frame	Time-based nonce + session expiry
DoS Attack	Mempool flooding	Private PoA + rate limiting

5.1 Resistance to Presentation Attacks

Presentation attacks attempt to deceive the biometric system using printed photos, replay videos, digital screens, or AI-generated deepfake animations. Static images and screen replays cannot reproduce physiologically correct blink patterns. Natural human micro-jitters are absent in spoof mediums. Frequency-domain texture features differentiate real faces from flat printed or AMOLED screen artifacts. The multi-modal combination means an attacker must simultaneously defeat all three independent liveness channels.

5.2 Resistance to Identity Theft

ArcFace-based recognition provides strong resistance through additive angular margin training that maximizes inter-class separation. Even identical twins or visually similar individuals have a low probability of matching unless the embedding angle is extremely small. Face alignment reduces matching errors caused by pose variation, and rejecting low-quality detections (confidence < 0.90) prevents blurry or partial faces from passing.

5.3 Double Voting and Network Security

The blockchain rejects any subsequent transactions from the same address. Session tokens are single-use and bound to a specific verification instance. The private PoA network prevents 51% attacks since validators are institution-controlled. Fixed zero gas price eliminates miner incentives to reorder transactions. Strict CORS policies, rate limiting, checksum verification on embeddings, and HTTPS transport security protect the FastAPI backend.

6. Results and Discussion

6.1 Verification Accuracy

The ArcFace model demonstrates superior identity discrimination compared to FaceNet-based baselines. With cosine similarity threshold $S \geq 0.35$ and liveness threshold $L \geq 0.30$, the system achieves a low false acceptance rate while maintaining usability for legitimate voters. The multi-modal liveness combination significantly reduces the probability of successfully spoofing all three modalities simultaneously.

Table 4 Performance Comparison: AuraVote vs. Baseline Approaches

System	Face Model	Liveness	Block chain	Latency
Baseline (MTCNN+FaceNet)	FaceNet 128-D	EAR only	None	~180 ms
Prior (Flask+ FaceNet)	FaceNet+SVM	MediaPipe	Ethereum PoA	~150 ms
AuraVote (Proposed)	ArcFace 512-D	Multi-modal	Ethereum PoA	70–120 ms

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1131-1137
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0160

6.2 System Throughput and Blockchain Integrity

The asynchronous FastAPI architecture with Uvicorn multi-worker deployment enables high concurrent request handling. The frame processing pipeline achieves real-time throughput of 6-10 verification frames per second per voter session. The five-second PoA block time ensures deterministic vote transaction finality well within acceptable user-experience bounds. The zero-gas configuration eliminates economic barriers to voter participation. All vote records are permanently stored on-chain with full auditability. The VoteCast event emission enables real-time monitoring dashboards while exposing only wallet addresses and candidate choices, never biometric data.

Conclusion

AuraVote demonstrates that combining deep-learning-driven biometric verification with blockchain consensus mechanisms results in a secure, transparent, and tamper-resistant electronic voting platform. The introduction of RetinaFace detection, ArcFace-based 512-dimensional embeddings, cosine similarity scoring, and multi-modal liveness analysis significantly increases resistance to spoofing attacks and identity fraud compared to traditional pipelines. The migration to a Proof-of-Authority blockchain enhances system throughput by providing deterministic finality, reduced latency, and predictable network behavior, making it well-suited for controlled election environments. Overall, AuraVote offers a practical and technologically advanced solution for secure digital elections, with a modular design that allows further enhancements in AI verification, network architecture, and cryptographic privacy. Future work will incorporate: (1) decentralized biometric verification using blockchain oracles; (2) Zero-Knowledge Biometric Proofs for privacy-preserving authentication; (3) advanced deepfake detection integrating PPG-based liveness and GAN-discriminator models; (4) formal smart contract verification using Certora or Slither; (5) role-based multi-administrator consensus for critical actions; and (6) end-to-end voter anonymity using mixnet-based transaction routing.

Acknowledgements

The authors thank the Department of Information Science and Engineering, Malnad College of Engineering, Hassan, for providing guidance and laboratory infrastructure support during the development and evaluation of this project.

REFERENCES

[1]. Rivest, R. L. (2008). On the notion of 'software independence' in voting systems. Philosophical Transactions of the Royal Society A, 366(1881), 3759-3767. https://doi.org/10.1098/rsta.2008.0152

[2]. Cortier, V., Galindo, D., & Kusters, R. (2020). How to fake zero-knowledge proofs, again. IEEE Security & Privacy, 18(6), 60-68. https://doi.org/10.1109/MSEC.2020.3012889

[3]. Zheng, Z., Xie, S., Dai, H. N., Chen, X., & Wang, H. (2018). Blockchain challenges and opportunities: A survey. International Journal of Web and Grid Services, 14(4), 352-375.

[4]. Douceur, J. R. (2002). The Sybil attack. Proceedings of IPTPS, Lecture Notes in Computer Science, 2429, 251-260. Springer, Berlin.

[5]. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. CVPR 2015, pp. 815-823. https://doi.org/10.1109/CVPR.2015.7298682

[6]. Zhang, S., Chi, C., Yao, Y., Lei, Z., & Li, S. Z. (2020). RetinaFace: Single-stage dense face localisation in the wild. CVPR 2020, pp. 5203-5212. https://doi.org/10.1109/CVPR42600.2020.00525

[7]. Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). ArcFace: Additive angular margin loss for deep face recognition. CVPR 2019, pp. 4690-4699. https://doi.org/10.1109/CVPR.2019.00482

[8]. Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., & Liu, W. (2018). CosFace: Large margin cosine loss for deep face recognition. CVPR 2018, pp. 5265-5274. https://doi.org/10.1109/CVPR.2018.00552

[9]. Li, Y., Xu, K., Wen, C., & Zhao, X. (2019). Face anti-spoofing via deep texture analysis. IEEE ICIP 2019, pp. 2099-2103. https://doi.org/10.1109/ICIP.2019.8803116

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1131-1137
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0160

[10]. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). DeepFakes and beyond: A survey of face manipulation and fake detection. Information Fusion, 64, 131-148. https:// doi.org/ 10.1016/j.inffus.2020.06.014

[11]. Ramirez, S. (2020). FastAPI: Modern, fast, web framework for building APIs with Python 3.6+. FastAPI Documentation. https://fastapi.tiangolo.com

[12]. Wood, G. (2014). Ethereum: A secure decentralised generalised transaction ledger. Ethereum Project Yellow Paper, 151(2014), 1-32.

[13]. Antonopoulos, A. M., & Wood, G. (2018). Mastering Ethereum: Building Smart Contracts and DApps. O'Reilly Media, Sebastopol, CA.

[14]. Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Processing Letters, 23(10), 1499-1503. https://doi.org/10.1109/LSP.2016.2603342

[15]. Liu, S., Lan, X., & Yuen, P. C. (2021). Remote photoplethysmography correspondence feature for 3D mask face presentation attack detection. IEEE Transactions on Information Forensics and Security, 16, 2604-2616. https://doi.org/10.1109/TIFS.2021.3065247