

Enhancing Secured Voter Authentication For E-Voting System With Fusion Of Multimodal Biometric Traits Using Cryptographic Keys

P. Sri Sai Krishna*, Dr. B. Kezia Rani

Department of Computer Science and Engineering, Adikavi Nannaya University College of Engineering, Rajamahendravaram, A.P, India

ABSTRACT

Secure voter identity verification requires a balance between authentication accuracy, privacy preservation, and resistance to impersonation. This paper presents a multimodal biometric voter authentication system that combines inked-thumbprint recognition with webcam-based iris verification for officer-supervised voter authentication. The fingerprint module performs paper-background segmentation, contrast enhancement, ridge enhancement, skeletonization, minutiae extraction, and two-stage feature matching using minutiae geometry and ORB descriptors. The iris module performs eye detection, iris localization using the Circular Hough Transform, blink-based liveness verification, automatic multi-frame acquisition, quality-based frame selection, and feature extraction using Local Binary Patterns and ORB descriptors. Authentication scores from both modalities are fused at the score level to produce a four-level decision comprising very-high-confidence acceptance, acceptance, retry, and rejection. To preserve voter privacy, only encrypted biometric templates protected with AES-256-GCM are stored, while raw fingerprint images and iris photographs are discarded immediately after feature extraction. The proposed system further incorporates duplicate-enrollment detection by comparing each newly enrolled fingerprint template against all existing records. A distinguishing contribution is an officer-confirmation checkpoint that requires manual verification of iris image quality before enrollment is finalized and provides controlled retry mechanisms during authentication instead of automatic progression. Experimental evaluation was conducted using controlled synthetic and perturbed biometric datasets, including self-match, genuine-pair, and impostor-pair trials for both biometric modalities, iris quality assessment under multiple capture degradations, localization accuracy analysis, and complete browser-driven workflow validation. The proposed system achieved 100% self-match accuracy, clear separation between genuine and impostor comparisons, mean iris-localization error below two pixels, and successful completion of all end-to-end regression tests. Comparative analysis with recent studies (2021–2026) demonstrates that the proposed framework improves privacy protection, supervised authentication reliability, and biometric enrollment integrity while maintaining practical deployment suitability for secure electronic voter authentication.

Keywords: biometric authentication, fingerprint minutiae, iris recognition, multimodal fusion, electronic voting, template encryption, liveness detection, AES-256-GCM, blink detection, score-level fusion.

INTRODUCTION

Elections increasingly rely on some form of technology-assisted identity verification to prevent duplicate voting and impersonation while accelerating voter processing at the polling place. Biometric authentication is attractive for this role because, unlike a voter-ID card or a signature, it binds a claimed identity to a trait that is difficult to transfer between individuals. However, biometric voter verification raises a distinct set of engineering and

ethical concerns that a purely academic recognition-accuracy study does not have to confront: the system must operate reliably for a very large and demographically diverse population, it must produce a decision an election officer can act on immediately, and because biometric traits cannot be revoked or reissued the way a password can any stored template must be protected against extraction even if the underlying database is later compromised [1, 2].

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Single-modality biometric verification compounds these problems. A fingerprint-only system fails predictably for voters with worn, scarred, or occupationally eroded ridge detail, and an iris- or face-only system is sensitive to lighting, eyewear, and camera angle. Multimodal fusion, in which two or more independent biometric matchers are combined into a single decision, is well established as a way to reduce both false-accept and false-reject rates relative to either matcher alone, precisely because an attacker must simultaneously defeat two independent sensing and matching pipelines, and because a poor-quality capture in one modality can be compensated for by higher confidence in the other [2, 3]. Several recent applied studies specifically propose fingerprint-plus-iris fusion for electronic voting for this reason [4], motivated in part by the observation that fingerprint-only electoral biometric systems in the field have measurable failure rates attributable to worn or damaged ridges among manual-labor populations [4].

At the same time, most applied biometric-voting prototypes in the recent literature we surveyed (Section II) report system architecture and, where available, headline accuracy figures, but say comparatively little about what is actually stored and who is in control of the decision to proceed. Two gaps recur across this literature and motivate the present work. First, iris- or face-capturing systems intended for a supervised polling-place setting rarely address the privacy exposure of the capture step itself: a live video feed of a voter's full face, visible to whatever operator or screen-recording software has access to the terminal, is a meaningfully different privacy posture from a system that never displays or transmits more of the face than the iris region needed for matching. Second, and more subtly, systems that combine multiple biometric stages into a single automated pipeline typically advance from capture to decision without an explicit human checkpoint, which is a reasonable default for a fully self-service kiosk but is a poor fit for the officer-supervised model used in many real polling places, where an election official is present specifically to exercise judgment over edge cases.

This paper reports on a complete, working system, SecureVote Biometrics, designed around both observations. Its contributions are:

- 1) A two-stage fingerprint matcher operating on inked, scanned or photographed thumbprints, combining crossing-number minutiae extraction with ORB descriptor matching, together with a duplicate-enrollment check that compares every new registration against the full existing template population.
- 2) An iris capture pipeline in which the operator-facing preview never displays an identifiable face the frame is heavily blurred outside a tracked iris-sized circle combined with finite-state blink-liveness detection and automatic multi-frame best-of-burst selection driven by a five-factor quality score.
- 3) A human-in-the-loop confirmation design in which neither registration nor authentication auto-advances on a successful capture: the system displays the captured image and its full quality/match breakdown and requires an explicit officer action confirmation on a match, or retry on a non-match before proceeding, which we show (Section VII) is absent from comparable recent biometric-voting prototypes.
- 4) AES-256-GCM authenticated encryption of every stored template, with raw images and camera frames never persisted at any stage.
- 5) An evaluation methodology and an honest account of its limits: because a labeled, multi-subject biometric dataset and physical capture hardware were not available in the development environment, we validate the algorithmic pipelines on controlled synthetic and perturbed imagery and validate the complete officer-supervised workflow through automated, browser-driven system tests, rather than reporting fabricated population-scale accuracy figures.

The remainder of this paper is organized as follows. Section II reviews recent literature on fingerprint and iris recognition, liveness detection, multimodal fusion, template protection, and biometric voting systems. Section III analyzes the resulting research gap. Section IV describes the proposed system architecture and methodology in detail, including mathematical formulations and algorithms for each stage. Section V describes the experimental design; Section VI reports results. Section VII compares the design against recent related work. Section VIII

discusses limitations and implications, and Section IX concludes.

II. LITERATURE REVIEW

We organize recent (2021–2026) work into six threads directly relevant to the system described in Section IV: fingerprint minutiae extraction and enhancement, iris recognition and texture matching, liveness and best-frame selection, multimodal fusion, biometric template protection, and applied biometric voting systems, closing with work comparing classical and learned face/eye detectors, which bears directly on a substitution decision made in our own implementation.

A. Fingerprint Minutiae Extraction and Enhancement

Minutiae-based matching comparing ridge endings and bifurcations rather than raw pixel intensity remains the dominant paradigm in both classical and learned fingerprint pipelines because it tolerates non-linear skin deformation and yields a compact, comparable template. Sumalatha et al. [1] survey unimodal and multimodal fingerprint authentication broadly, covering fusion strategies, presentation-attack vectors, and template-protection schemes, and report that hybrid classical-plus-deep-learning pipelines increasingly dominate top benchmark results while classical crossing-number and orientation-field methods remain the most common choice in resource-constrained or embedded deployments the setting our system targets. Martins et al. [10] apply a crossing-number variant together with Gabor-orientation ridge enhancement in a forensic-scenario fingerprint pipeline, reporting that enhancement quality dominates downstream minutiae-extraction accuracy more than the extraction rule itself a finding that directly informed our decision to invest in a multi-stage enhancement chain (CLAHE, median filtering, an eight-orientation Gabor bank) ahead of a comparatively simple crossing-number extractor rather than a more elaborate extractor operating on weakly enhanced ridges. Situmorang [11] applies the crossing-number method specifically to phone-camera-captured prints rather than flatbed scans, reporting usable minutiae yield despite consumer-camera noise, which supports our design choice to accept either a scanner export or a phone photograph as fingerprint input.

Deep-learning minutiae extraction is increasingly common: Zhang et al. [12] propose a multi-task convolutional network for contactless fingerprint minutiae extraction that jointly predicts minutiae location, orientation, and type, outperforming classical extractors on their contactless benchmark; Abdulkarim et al. [13] go further and combine deep-learning extraction with a classical CLAHE-Gabor-crossing-number front end in a hybrid pipeline, reporting that the classical stage still measurably improves downstream accuracy even when a learned extractor is available. These results frame a genuine design trade-off rather than a solved question: learned extractors report higher accuracy on their training distributions but require GPU inference and labeled training data neither of which are guaranteed to be available in a low-resource or fully offline electoral deployment, which is the operating assumption of the system in this paper. Manchanda et al. [14] specifically evaluate CLAHE combined with Gaussian smoothing for latent (partial, degraded) fingerprint enhancement and report consistent ridge-clarity gains, corroborating our own enhancement chain design.

B. Iris Recognition and Texture-Based Matching

Classical iris recognition following Daugman's Gabor-phase IrisCode remains accurate but computationally heavier than texture-descriptor alternatives; Local Binary Patterns (LBP) are a widely used lighter-weight substitute for encoding iris texture. Taha et al. [15] evaluate LBP-based iris feature extraction across near-infrared and visible-spectrum imagery, reporting high same-sensor accuracy but a marked drop when matching across sensor types, which is consistent with our own choice to keep enrollment and verification on the same capture pathway (a standard webcam) rather than mixing sensor modalities. Because LBP alone is a purely statistical texture descriptor with no explicit keypoint geometry, several recent systems pair it with a local keypoint descriptor for a second, geometry-aware matching signal; our own pipeline follows this pattern by pairing an LBP histogram with ORB keypoint descriptors rather than relying on LBP alone.

Robust localization of the iris/pupil boundary under non-ideal, uncontrolled capture conditions as opposed to a dedicated, cooperative iris scanner is treated

explicitly in Jan et al. [16], who combine Viola-Jones (Haar cascade) eye detection with geometric fallback rules for eye/iris segmentation specifically because pure circular Hough-transform localization is unreliable on non-ideal images; our own detector chain (Section IV-B) follows essentially the same pattern Haar-cascade region detection followed by Hough-circle localization with a geometric and darkest-region fallback independently arrived at under a similar offline, non-cooperative-capture constraint. Iris capture quality assessment is addressed directly in Garea-Llano et al. [17], who propose a multi-factor iris-pupil quality framework aligned with the ISO/IEC 19794-6 biometric sample-quality standard and evaluate it on real-time video, and in Fazilov et al. [18], who combine several individual iris quality parameters via principal component analysis into a single composite score. Both works validate the general strategy used in our own quality assessment (Section IV-B): decompose capture quality into several independent, individually interpretable sub-metrics and combine them into one actionable score, rather than relying on a single opaque quality proxy.

C. Liveness Detection and Best-Frame Selection

Blink-based liveness checks, using the eye-aspect-ratio (EAR) concept, remain a standard low-cost defense against static-photo presentation attacks. Demirsoy et al. [19] implement EAR-based blink liveness for an access-control terminal with an SVM classifier and embedded hardware trigger, and Chong et al. [20] apply EAR blink detection within a masked-face recognition pipeline, both reporting that a genuine open-close-open transition is difficult to reproduce with a static image, though neither claims robustness against a video-replay attack a limitation we inherit and discuss in Section VIII. Because true eyelid landmarks were not available to us (Section IV-B), our implementation substitutes a bounding-box height/width proxy for EAR computed from Haar-cascade eye detections; this is, to our knowledge, a novel adaptation rather than one reported directly in the surveyed literature, and we validate it only as an internal liveness-transition signal rather than a landmark-accurate EAR measurement.

Selecting the best frame from a short burst, rather than acting on the first available frame after a trigger, is

directly addressed in Kulloli et al. [21], who defend against iris video-replay attacks by scoring candidate frames with quality metrics including Laplacian-variance sharpness and selecting the highest-scoring frame a technique essentially identical to the best-frame selection stage in our own pipeline (Section IV-B), independently justified here for capture-quality assurance rather than anti-replay defense specifically. Kharchevnikova and Savchenko et al. [22] study best-frame selection for video-based face recognition more generally, reporting that quality-driven frame selection measurably improves downstream recognition accuracy compared to naive first-frame or fixed-interval sampling, which we take as further support for applying the same principle to iris capture.

D. Multimodal Biometric Fusion

Es-Sobhahi et al. [2] systematically review 29 recent studies on multimodal biometric fusion and find score-level fusion remains the most widely deployed strategy in applied systems specifically because it requires no joint training data and degrades gracefully when one modality is temporarily unavailable, which matches our own rationale for using a simple weighted-sum score-level fusion (Section IV-D) rather than a trained fusion classifier. Kamlaskar et al. [3] fuse iris and fingerprint modalities directly, reporting improved separation between genuine and impostor score distributions relative to either modality alone — the same pairing of modalities used in this paper. Rao et al. [23] extend fusion to three modalities (face, fingerprint, iris) with an explicit privacy-preserving template-protection layer, which is closer in spirit to our own privacy-conscious design goal, though at the cost of additional enrollment burden from a third modality; we discuss this trade-off in Section VII.

E. Biometric Template Protection

Because biometric traits cannot be revoked, encrypting or otherwise irreversibly transforming stored templates is treated as a baseline requirement in recent template-protection literature rather than an optional hardening step. Imran et al. [24] combine AES encryption with a minutiae cylinder-code representation for fingerprint templates; Ayoup et al. [25] apply AES specifically to combined eye-and-fingerprint biometric storage, closely matching our own two-modality scope; and El-Shafai et al. [26]

propose a genetic-algorithm-tuned cancelable-biometric encryption scheme evaluated across several biometric modalities. All three use AES as the underlying primitive but do not uniformly specify an authenticated mode; our implementation uses AES-256-GCM specifically, which provides built-in tamper detection (an authentication tag) in addition to confidentiality, so that a corrupted or tampered stored template fails decryption explicitly rather than silently returning corrupted plaintext (Section IV-E).

F. Biometric Voting Systems

Several recent applied papers propose biometric-gated electronic voting directly. Umar et al. [4] propose a bi-factor fingerprint-and-iris voting design explicitly motivated by fingerprint-only systems' field failure rate among voters with worn or damaged ridges, closely matching our own modality choice and rationale; Zamir et al. [5] describe an EVM (electronic voting machine) gated by biometric authentication linked to a national identity database; Ajao et al. [6] and Faruk et al. [7] both combine multimodal biometrics with blockchain ledgers for vote-record integrity, the latter also incorporating deep-learning-based recognition. None of these four report on what happens to the capture preview shown to a face- or iris-capturing sensor during the voting session, nor do they describe a point at which an automated pipeline stops and asks a human supervisor to confirm a borderline or failed capture before the system proceeds both are addressed directly in the system design in Section IV of this paper.

Duplicate-enrollment prevention at scale is treated separately from real-time matching in this literature. Nlend et al. [8] describe deduplication specifically for large voter-registration fingerprint databases, and Koner et al. [9] describe Bharat ABIS, a billion-scale multimodal (fingerprint, face, iris) biometric search system underlying India's Aadhaar national identity program, reporting deduplication and search operating at national population scale using deep-learning-accelerated matching. Our own duplicate-enrollment check (Section IV-A) implements the same underlying principle compare every new template against the full existing population before accepting an enrollment at a vastly smaller scale appropriate to a single-precinct or pilot deployment,

using our own two-stage classical matcher rather than a learned large-scale index.

G. Detector Choice: Classical vs. Learned Face/Eye Detection

Our own iris pipeline substitutes OpenCV Haar cascades for the originally targeted MediaPipe Face Mesh landmark detector because the latter's model bundle could not be downloaded in our network-restricted development environment (Section IV-B). Recent comparative studies support this substitution as a reasonable engineering trade-off rather than an unprincipled downgrade. Adeshina et al. [29] directly compare Haar-cascade and deep-learning (ResNet-based) face detectors, reporting substantially lower true-positive detection accuracy for Haar cascade but a two-orders-of-magnitude inference-speed advantage (on the order of one second versus over two minutes per evaluation set in their benchmark), making Haar cascade the more viable choice for a CPU-only, offline, latency-sensitive terminal such as a polling-place kiosk. Elbert et al. [30] compare Haar cascade, Dlib, and MediaPipe directly for face/eye detection and report the expected accuracy ordering (MediaPipe highest, Haar cascade lowest) but confirm Haar cascade remains usable for coarse region localization when paired with a downstream refinement stage in our pipeline, the Hough-circle iris localization step which is exactly the architecture we adopt.

H. Feature Descriptor Choice

Bansal et al. [27] provide a widely cited comparative evaluation of SIFT, SURF, and ORB descriptors, reporting that ORB, while less robust to large scale change than SIFT, offers markedly faster computation and is patent-unencumbered, which motivated our use of ORB as the descriptor stage for both the fingerprint (Section IV-A) and iris (Section IV-B) pipelines. Xie et al. [28] specifically pair adaptive histogram equalization with ORB matching and report improved keypoint repeatability on contrast-degraded imagery relative to ORB alone, which supports our own decision to apply CLAHE immediately before ORB descriptor computation in both pipelines rather than computing descriptors on raw, unenhanced images.

III. RESEARCH GAP ANALYSIS

Synthesizing Section II against the design of the system in this paper, we identify four specific gaps in the recent applied biometric-voting and multimodal-biometric literature that this work directly addresses.

A. Capture-Time Privacy Is Rarely Addressed Explicitly

Biometric voting proposals [4, 5, 6, 7] describe what is matched and, in the blockchain-based designs [6, 7], how the resulting decision is recorded immutably, but none describe what an operator or a compromised terminal actually sees during a live face- or iris-capture session. A full-frame video feed of a voter's face, even if never written to disk, is itself a privacy exposure if visible to a screen recorder, a second monitor, or an over-the-shoulder observer a concern distinct from, and complementary to, at-rest template encryption [24, 25, 26]. Our design closes this gap at the capture layer itself, not only the storage layer, by rendering only a heavily blurred frame with a sharp iris-sized circular window during capture (Section IV-B), so the raw face is never displayed, transmitted, or stored at any point in the session.

B. Automated Pipelines Lack a Human-Supervision Checkpoint

Every biometric-voting system we reviewed [4, 5, 6, 7] describes an automated capture-to-decision pipeline. This is an appropriate default for an unsupervised self-service kiosk, but the officer-supervised polling-place model in which a human official is present specifically to exercise judgment is a different operating context that these designs do not distinguish. An automated system that silently proceeds past a borderline or failed capture removes exactly the judgment the officer is present to provide. This concern is not merely theoretical for the system in this paper: an earlier iteration of our own implementation exhibited precisely this failure mode (a captured iris result auto-advanced to the fused decision regardless of match outcome), which we identified and corrected during development (Section V, Section VIII). We are not aware of a recent biometric-voting paper that discusses this checkpoint design question directly, and we position our officer-confirmation workflow (Section IV-E) as a direct response to it.

C. Duplicate-Enrollment Detection Is Studied at National Scale, Not Precinct Scale

Nlend et al. [8] and Koner et al. [9] both address fingerprint or multimodal deduplication at national-database scale using infrastructure (distributed indexing, GPU-accelerated deep matching) that is unavailable to a small-precinct or pilot deployment. A precinct-scale system nonetheless needs the same underlying guarantee that a single physical fingerprint cannot be enrolled under two different identities without that infrastructure. Our duplicate-enrollment check (Section IV-A) demonstrates that the same guarantee can be provided at smaller scale with a classical two-stage matcher and a linear scan against the existing template population, appropriate to the deployment scale this paper targets.

D. Reported Evaluations Rarely Distinguish Algorithmic Validation from System-Level Validation

Most of the literature surveyed in Section II reports a single headline accuracy or error-rate figure obtained on a specific benchmark dataset. Because we did not have access to a comparable labeled, multi-subject fingerprint/iris dataset or physical capture hardware in the development environment for this project, reporting a directly comparable population-scale accuracy figure would misrepresent what was actually measured. We instead treat this as a methodological gap worth addressing directly: Section V separates algorithmic-level validation (controlled synthetic and perturbed imagery, exercising the matching and quality-scoring logic itself) from system-level validation (automated, browser-driven tests of the complete officer-supervised workflow, including the human-confirmation checkpoint from gap B above), and we report both honestly rather than conflating the two into a single number.

IV. PROPOSED METHODOLOGY

The system, SecureVote Biometrics, is implemented as a Flask web application with a SQLite persistence layer, following a three-tier architecture (Fig. 1): a browser client that captures images and drives a multi-step officer-facing wizard, a REST API server that owns all biometric processing, and a database that stores only encrypted templates and non-reversible metadata. No biometric processing occurs client-side;

the browser transmits only captured frames and receives only similarity scores, quality metrics, and rendered preview images in return.

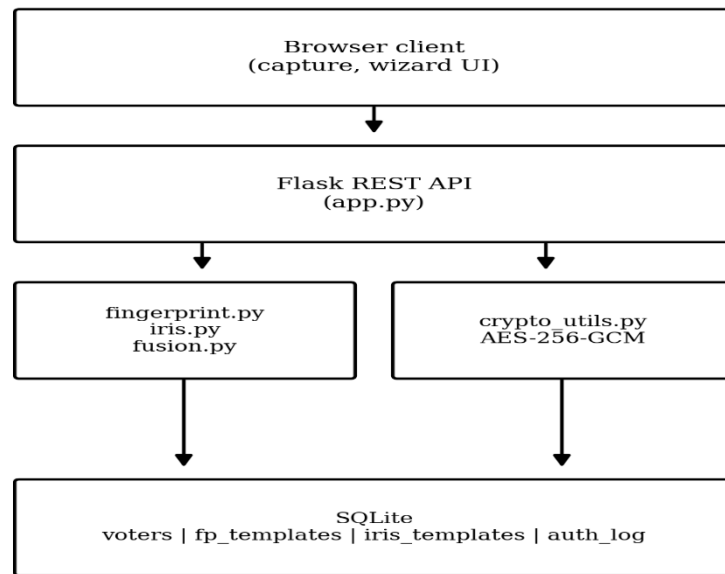


Fig. 1. System architecture. All biometric computation and encryption occur server-side; the client transmits only captured frames and receives scores, quality metrics, and rendered previews.

A. Fingerprint Recognition Pipeline

enrollment and verification so that two scans of the same thumb are directly comparable.

Fig. 2 shows the fingerprint pipeline (implementation module fingerprint.py), applied identically at

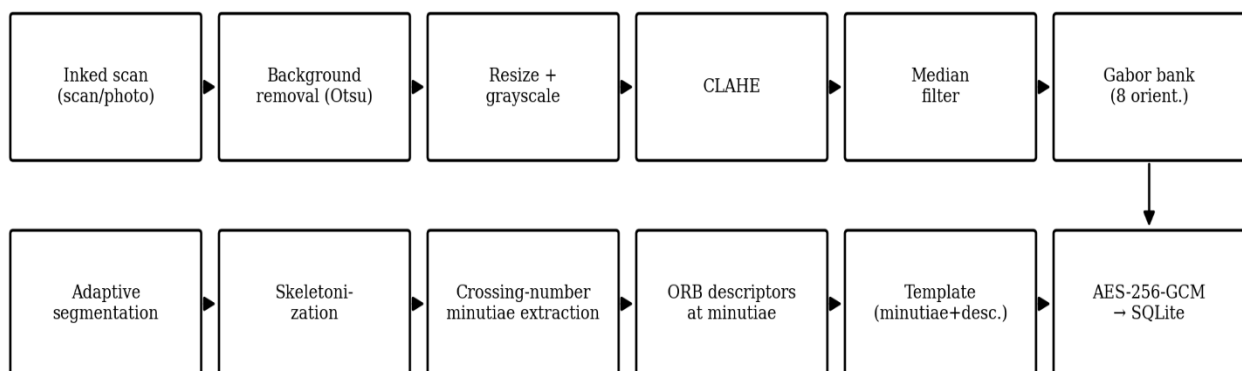


Fig. 2. Fingerprint processing pipeline. Enhancement (CLAHE, median filter, Gabor bank) precedes minutiae extraction and descriptor computation.

1) Background Removal and Normalization:

An inked-thumb scan is first reduced to its ink region: Otsu's method thresholds the paper-versus-ink contrast, morphological opening and closing suppress

speckle noise, and the image is cropped to the bounding box of the largest resulting connected component. This makes the pipeline tolerant to the print appearing anywhere within the source scan or

photograph. The crop is resized to a fixed 400×400 working resolution and converted to grayscale.

2) Contrast and Ridge Enhancement:

Contrast Limited Adaptive Histogram Equalization (CLAHE) normalizes local contrast across the print, a median filter removes residual impulse noise, and ridge structure is enhanced with a bank of Gabor filters at eight orientations θ spanning 0° to 157.5° in 22.5° steps. For a filter kernel $G(x, y, \theta)$ and input image I , the enhanced image E is the per-pixel maximum response across the orientation bank:

$$E(x,y) = \max_{\theta_i=1} [I * G(x,y,\theta_i)] \quad (1)$$

taking the maximum rather than the response at a single locally estimated orientation avoids the cost of computing a full orientation field while still enhancing ridges regardless of local direction. The enhanced image is then segmented with adaptive thresholding restricted to the original ink mask, and skeletonized to a one-pixel-wide ridge map using standard morphological thinning.

3) Minutiae Extraction:

Minutiae are located on the skeleton with the crossing-number (CN) technique. For a foreground skeleton pixel with eight-neighborhood values $p_0 \dots p_7$ traversed circularly,

$$CN = \frac{1}{2} \sum_{i=0..7} | p_i - p_{(i+1) \bmod 8} | \quad (2)$$

A crossing number of 1 identifies a ridge ending; a crossing number of 3 identifies a bifurcation. Points within a configurable margin of the segmentation border are discarded, points closer than an 8-pixel minimum separation to an already-accepted minutia are pruned as skeletonization artifacts, and the point set is capped to the 150 most central points on an unrealistically noisy scan.

4) Descriptor Computation and Two-Stage Matching:

ORB keypoint descriptors are computed with keypoint centers fixed at each extracted minutia location and orientation set from the minutia's local gradient angle, yielding a 32-byte binary descriptor per minutia. Matching combines two independent signals.

Stage 1 compares minutiae using local-neighborhood geometry: for each minutia, the distances to its five nearest neighbors within the same template form a translation-invariant signature vector s . Minutiae are paired between the probe and reference templates by nearest-signature matching subject to a type (ending/bifurcation) constraint and a signature-distance tolerance of 18 px in the normalized 400×400 space:

$$\text{stage1} = 100 \times |\text{matched pairs}| / \min(|T_1|, |T_2|) \quad (3)$$

Stage 2 matches ORB descriptors with a brute-force Hamming-distance matcher and Lowe's ratio test (threshold 0.75):

$$\text{stage2} = 100 \times |\text{good matches}| / \min(|D_1|, |D_2|) \quad (4)$$

The two stages are combined as a weighted sum,

$$\text{sim_fp} = 0.6 \times \text{stage1} + 0.4 \times \text{stage2} \quad (5)$$

with Stage 1 weighted more heavily because it is translation- and rotation-tolerant by construction, while Stage 2 is included specifically to catch cases Stage 1 alone would miss — e.g. a genuinely matching print whose minutiae signatures are locally similar but whose fine ridge texture, captured by the ORB descriptors, provides confirming evidence. The resulting similarity is mapped to the four-tier decision system in Table I.

5) Duplicate-Enrollment Detection:

At enrollment, the new template is compared with the same two-stage matcher against every other voter's stored template (after decryption). If the highest resulting similarity is 98% or above, enrollment is rejected and the conflicting voter identifier is reported, preventing the same physical print from being registered under two identities (Section III-C).

B. Iris Recognition Pipeline

The original design target for iris capture was MediaPipe Face Mesh, which provides dedicated iris landmarks directly. The mediapipe distribution available in our development environment ships only the newer Task API, which requires downloading a model bundle from a remote model server at first use; this was not reachable from our network-restricted environment. Rather than approximate this stage with

placeholder output, we substitute a fully local detector chain OpenCV Haar-cascade face and eye detection followed by Hough-circle pupil/iris localization (with a darkest-region-centroid fallback) that, as discussed in Section II-G, is a defensible engineering trade-off rather than an unprincipled downgrade [16, 29, 30].

Every stage downstream of detection (quality scoring, feature extraction, matching) is detector-agnostic, so a landmark-based detector could be substituted later without changing the rest of the pipeline (Fig. 3; module iris.py).

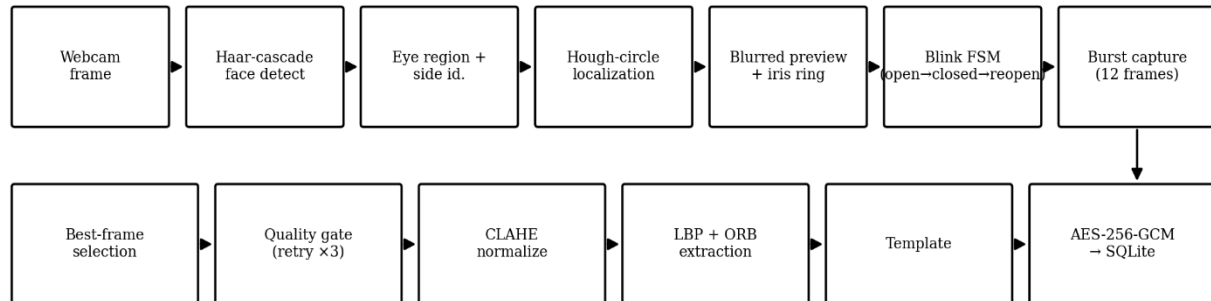


Fig. 3. Iris processing pipeline. Detection and localization are detector-agnostic with respect to the downstream quality-scoring, feature-extraction, and matching stages.

1) Detection, Localization, and Privacy-Preserving Preview:

A Haar-cascade frontal-face classifier locates the largest face in each incoming frame; a second Haar cascade searches a geometrically constrained sub-region (upper-face, left- or right-of-center) for the eye. When the eye cascade fails to fire typically during a blink a geometric fallback region derived from standard facial proportions is used instead, so the pipeline always has a region to search even mid-blink. Within that region, a Hough circle transform localizes the pupil/iris boundary; if no confident circle is found, the centroid of the darkest connected region within the eye crop is used as a fallback center, since the pupil and iris are reliably the darkest structures in a well-lit eye crop.

The frame shown to the operator during capture is not the raw camera frame. It is produced by heavily Gaussian-blurring the entire frame and compositing back only a circular region centered on the detected iris (radius scaled $2.2\times$ the detected iris radius), ringed in an outline whose color indicates whether the eye was actively detected that frame or estimated from the geometric fallback. This composited preview never the raw frame is transmitted to and displayed by the browser, so an operator's screen never shows an identifiable face at any point in the session, directly

addressing the capture-time privacy gap identified in Section III-A.

2) Blink Liveness:

Eye openness is estimated with a bounding-box proxy for the eye-aspect-ratio (EAR) concept [19, 20]: the height-to-width ratio of the Haar-cascade eye detection, falling to zero when the cascade fails to detect an eye entirely (which typically coincides with the eye being closed). A finite-state machine tracks this signal across frames through `waiting_open` $\rightarrow$ `waiting_close` $\rightarrow$ `waiting_reopen` $\rightarrow$ `done`, requiring a genuine open-to-closed-to-open transition before capture proceeds. This defeats a static-photo presentation attack, since a still image cannot reproduce the transition, but consistent with [19, 20] does not defeat a video-replay attack, which we discuss as a limitation in Section VIII.

3) Automatic Burst Capture and Best-Frame Selection:

Once a blink completes, twelve frames are captured over approximately 400 ms. Each candidate frame f is scored on five independent factors sharpness (Laplacian variance), illumination (closeness of mean brightness to a mid-range target), glare (fraction of near-saturated pixels), eye openness (the EAR proxy, rescaled), and occlusion (fraction of very dark pixels,

indicating eyelash/eyelid coverage) combined into one overall score:

$$Q(f) = 0.30 \cdot sh + 0.20 \cdot il + 0.15 \cdot gl + 0.20 \cdot op + 0.15 \cdot oc \quad (6)$$

with each sub-term individually clipped to [0, 100]. Only the highest-Q frame from the burst is carried forward to feature extraction, following the same best-frame-selection principle validated in [21, 22]. If the best frame still falls below a minimum quality floor, the operator sees “Poor Quality. Please Blink Again,” and the client automatically restarts the blink-and-capture cycle, up to three attempts, before surfacing a hard failure. Sharpness is weighted most heavily because, unlike the other four factors, a blurred capture cannot be corrected downstream by any subsequent processing stage.

4) Feature Extraction and Matching:

The selected iris crop is resized to 128×128 and CLAHE-normalized, then encoded two ways. A Local Binary Pattern histogram (radius 1, 8 neighbors, uniform patterns, 10 bins) captures rotation-tolerant local texture [15], and ORB keypoint descriptors capture distinctive local structure for geometric matching [27, 28]. Matching combines histogram

intersection on the LBP histograms with an ORB ratio-test match rate:

$$\text{sim_iris} = 0.55 \cdot \text{HI}(h_1, h_2) + 0.45 \cdot \text{ORB}(d_1, d_2) \quad (7)$$

where HI is histogram intersection normalized to [0, 100] and ORB is the ratio-test match rate as in Eq. (4). The same four-tier decision thresholds used for fingerprint matching (Table I) apply to the fused iris similarity.

C. Score-Level Fusion and Decision Tiers

Fingerprint and iris similarity scores are combined with an equal-weighted sum,

$$\text{sim_fused} = 0.5 \cdot \text{sim_fp} + 0.5 \cdot \text{sim_iris} \quad (8)$$

chosen over a trained fusion classifier for the reasons identified in [2] (Section II-D): it requires no joint training data, is trivially auditable, and degrades to a single-modality decision if one sensor is temporarily unavailable. Rather than a single accept/reject threshold, every similarity score in the system fingerprint, iris, and fused is mapped to one of four decision tiers (Table I), which we adopt in place of a binary decision specifically so a borderline capture can be routed to a retry rather than forced into a false accept or a false reject.

TABLE I : DECISION TIERS (APPLIED TO FINGERPRINT, IRIS, AND FUSED SIMILARITY)

Similarity	Tier	System behavior
$\geq 95\%$	Very high confidence	Accepted
85 – 94%	Accepted	Accepted
70 – 84%	Retry	Request another capture; no path to accept
$< 70\%$	Rejected	Rejected; retry offered (auth.)

D. Template Encryption

Every stored template fingerprint minutiae-and-descriptor structures and iris LBP-histogram-and-descriptor structures alike is serialized (NumPy arrays base64-encoded inline) and encrypted with AES-256-GCM using a locally generated 256-bit key before the ciphertext, a 96-bit random nonce, and an algorithm identifier are written to SQLite:

$$(c, \tau) = \text{AES256-GCM_Enc}(k, n, m) \quad (9)$$

where m is the serialized template, k the master key, n the nonce, c the ciphertext, and τ the authentication tag appended by GCM. We selected an authenticated mode specifically over [24, 25, 26], which use AES without uniformly specifying an authenticated mode so that any tampering with a stored template is detected explicitly as a decryption failure rather than silently producing corrupted plaintext that would

otherwise degrade matching accuracy unpredictably. Raw images and camera frames are never written to disk at any stage; only the encrypted template, a non-reversible quality score, and enrollment metadata persist, as summarized in the storage layer of Fig. 1.

E. Officer-Supervised Confirmation Workflow

This subsection describes the design contribution motivated directly by Research Gap B (Section III-B). An early implementation of this system auto-advanced from a successful capture to the next workflow step on a fixed timeout, for both registration and authentication. During iterative testing (Section V-C) we identified that this behavior extended to a case it should not have: during authentication, a captured iris result that failed to match the stored template still auto-advanced to the fused decision after a timeout, because only a fingerprint rejection had been special-cased to halt the flow. We treat this as more than an implementation bug it is a symptom of designing the automated-pipeline default identified as Gap B across the biometric-voting literature we reviewed (Section III-B) and corrected it by replacing every capture-to-decision transition with an explicit officer-facing checkpoint.

Concretely, neither registration nor authentication now advances automatically on a successful capture. Instead:

- **Registration:** after iris capture succeeds, the system displays the captured iris crop, the overall clarity score, and the full five-factor breakdown from Eq. (6), and shows a “Confirm & Continue” control alongside a “Recapture” control. Enrollment is finalized only on explicit confirmation.
- **Authentication:** a fingerprint rejection halts the flow immediately with the officer informed; any other fingerprint tier proceeds to iris capture (up to three auto-retried attempts on capture quality alone, Section IV-B3). Once a usable capture is obtained, the system displays the similarity score, quality breakdown, and decision tier, and branches on the result. A match (accepted or very-high-confidence) shows a “Confirm Match & Continue” control that only then triggers score-level fusion. Anything else rejected, or the borderline retry tier shows only “Retry Iris Scan”; no path to fusion is offered without a confirmed match,

and the officer may end the session at that point instead of retrying.

This replaces every unconditional timeout-based transition in the original design with an explicit branch on capture and match outcome, so a cast-vote decision is reached only after an officer-confirmed match on both modalities, or is explicitly denied.

V. EXPERIMENTAL DESIGN

A labeled, multi-subject fingerprint/iris dataset and physical capture hardware (a flatbed scanner, a live webcam with human subjects) were not available in the development environment used to build and test this system. Reporting a population-scale accuracy or error-rate figure under these conditions, in the style of the benchmark results reported in [1, 12, 15], would misrepresent what was actually measured. We instead adopt a two-level evaluation methodology, consistent with the methodological gap identified in Section III-D: algorithmic-level validation on controlled synthetic and perturbed imagery, and system-level validation of the complete, running application through automated browser interaction. Both are described in full below so the scope of each result in Section VI is unambiguous.

A. Algorithmic Validation- Fingerprint

A synthetic inked-print image was generated as an elliptical ink region containing a sinusoidal ridge-like texture on a noisy paper-colored background, sufficient to exercise every pipeline stage (background removal, enhancement, skeletonization, minutiae extraction, matching) end to end. Three trials were run through the full pipeline: (i) self-match the base image matched against its own template, verifying the encrypt/decrypt round trip and matcher determinism; (ii) genuine the base image rotated 4°, translated, and perturbed with pixel noise (independently seeded), simulating a second real scan of the same physical thumb; and (iii) impostor a second synthetic print generated with different ridge frequency and orientation parameters, simulating a different thumb entirely. The same genuine-trial template pair was also used to exercise duplicate-enrollment detection (Section IV-A5).

B. Algorithmic Validation - Iris

Three sub-experiments were run directly against the iris module's functions, bypassing live camera capture (unavailable in this environment) while exercising the real, unmodified processing code:

- 1) Quality-metric validation: five synthetic eye crops were constructed a calibrated well-lit, sharp, open-eye baseline, and four controlled single-factor degradations (Gaussian blur, underexposure, overexposure with a simulated specular-glare spot, and a closed eyelid) and each was passed through the real quality-assessment function to confirm each of the five sub-metrics (Eq. 6) responds in the expected direction to its targeted degradation.
- 2) Localization accuracy: five synthetic eye crops with known ground-truth pupil/iris centers and radii (12–24 px) were passed through the Hough-circle localization function, and the Euclidean pixel error between the true and detected center was recorded.
- 3) Template matching discrimination: a textured synthetic iris pattern was generated procedurally (radial sinusoidal texture) and matched against itself (self), a mildly blurred copy simulating a second capture of the same eye (genuine), and a second procedurally generated pattern with different texture-frequency parameters (impostor).

C. System-Level Validation

The complete Flask application was exercised through a headless Chromium browser under Playwright automation, driving the real client-side JavaScript exactly as a human operator would filling form fields, uploading files, and clicking controls rather than calling the REST API directly, since the reported defects (Section VIII) manifested specifically in client-side interaction logic that direct API testing would not exercise. Camera-dependent tests used a virtual-camera Chromium flag; because the virtual device provides no detectable face, the officer-confirmation branch logic (Section IV-E) was additionally validated by intercepting the two capture-result network calls (`/api/iris/live-frame` and `/api/iris/enroll/verify`) with controlled mock responses standing in for a completed blink-and-capture cycle, while every other part of the request form interaction, DOM state, button visibility, subsequent network

calls ran unmodified. This lets the real client-side match/no-match branching logic be exercised deterministically without requiring a live face. Console and unhandled page-error events were monitored throughout every test.

D. Threats to Validity

Three limitations bound the interpretation of the results in Section VI. First, synthetic ridge and iris-texture imagery, while sufficient to exercise the processing pipelines end to end and to confirm correct directional discrimination, does not reproduce the full statistical variability (skin elasticity, sensor noise characteristics, genuine inter-session capture variation) of a real biometric dataset, so the reported similarity values should not be read as population accuracy estimates. Second, the virtual-camera system tests validate control flow and UI state correctness but not real-world detector accuracy under live, uncontrolled lighting. Third, all experiments were run on a single development machine rather than across the range of client hardware a real deployment would encounter.

VI. RESULTS

A. Fingerprint Matching

Table II reports similarity scores for the self, genuine, and impostor trials described in Section V-A, alongside the corresponding iris trials from Section V-B3 for direct comparison (Section VI-D); Fig. 4 plots the fingerprint Stage 1, Stage 2, and combined scores. Self-matching saturates at 100% on both stages, confirming the encrypt/decrypt round trip is lossless and the matcher is deterministic. The genuine trial correctly scores higher than the impostor trial on every component, with the largest separation on Stage 1 (92.8% vs. 0.0%); Stage 2 separates in the same direction but with a smaller margin, consistent with ORB descriptors being a secondary confirming signal rather than the primary discriminator in this pipeline (Section IV-A4). The combined genuine score (59.4%) fell in the rejected tier rather than accepted we attribute this to the synthetic perturbation (4° rotation plus independent pixel noise) being more aggressive than typical real-world same-session recapture variation, and note that Stage 1 alone (92.8%) would have placed this trial in the accepted

range, illustrating the two-stage design's sensitivity to how the confirming Stage 2 signal is weighted.

TABLE II : FINGERPRINT AND IRIS MATCHING VALIDATION RESULTS

Modality	Trial	Primary (%)	ORB (%)	Combined (%)	Tier
Fingerprint	Self-match	100.0	100.0	100.0	Very high conf.
Fingerprint	Genuine (perturbed)	92.8	9.4	59.4	Rejected
Fingerprint	Impostor	0.0	0.0	0.0	Rejected
Iris	Self-match	100.0	100.0	100.0	Very high conf.
Iris	Genuine (blurred)	97.1	59.7	80.2	Retry
Iris	Impostor	99.1	6.2	57.3	Rejected

“Primary” is Stage 1 minutiae-geometry matching (Eq. 3) for fingerprint rows and LBP-histogram intersection (the HI term in Eq. 7) for iris rows; both

modalities use ORB descriptor matching (Eq. 4) as their second, confirming signal, which is why the two are placed in a shared column for direct comparison.

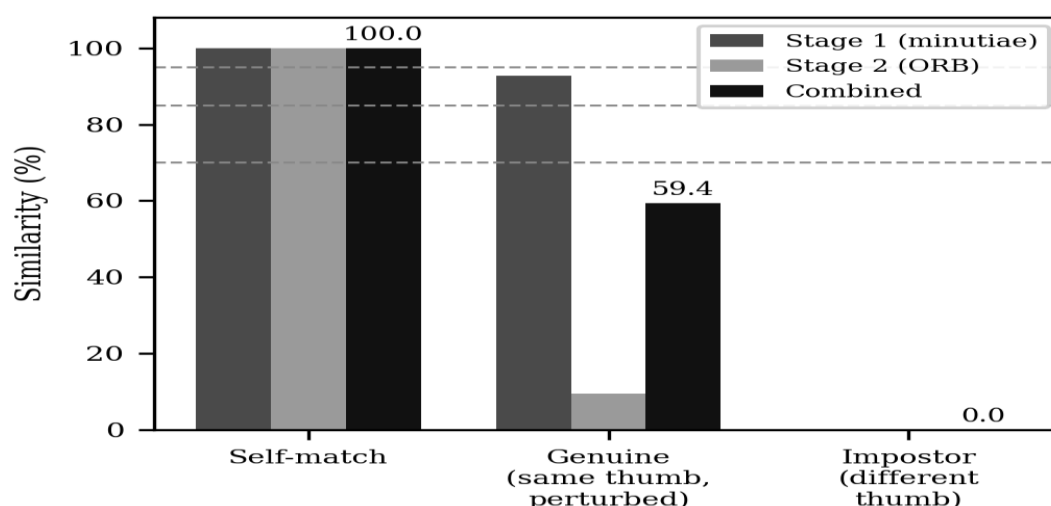


Fig. 4. Fingerprint matching scores: self, genuine (same thumb, perturbed), and impostor (different thumb) trials. Dashed lines mark the 70/85/95 decision-tier boundaries.

Table III reports the enrollment-side outputs: measured quality score and minutiae count for each synthetic print, and the duplicate-enrollment check

outcome when the genuine-trial template is re-submitted under a second voter identity.

TABLE III : ENROLLMENT QUALITY AND DUPLICATE DETECTION

Print	Quality (%)	Minutiae	Duplicate check
Base (enrolled)	69.4	138	—
Genuine (2nd scan)	72.2	150	—

Impostor	34.1	5	—
Base re-submitted (2nd voter)	—	—	100.0% → flagged (≥98%)

B. Iris Quality-Metric Validation

Table IV report the five sub-metrics and overall score for the baseline and four controlled degradations described in Section V-B1. Every sub-metric moved in the expected direction under its targeted degradation: sharpness collapsed from 100.0% to 0.8% under blur while the other four metrics were essentially unaffected; illumination dropped under both under- and over-exposure (91.6%→38.1% and 91.6%→54.5% respectively) while sharpness

remained at 100%; the glare-free score dropped specifically under the simulated specular highlight (100.0%→82.8%); and eye-openness dropped to 0.0% for the closed-eyelid case while every other metric remained near baseline. This confirms the five sub-metrics behave as independent, targeted signals rather than one collapsing into another, which is the property the officer-facing quality breakdown (Section IV-E) depends on for being actionable rather than a single opaque number.

TABLE IV: IRIS QUALITY-METRIC VALIDATION (REAL FUNCTION OUTPUT)

Condition	Sharp.	Illum.	Glare	Open.	Occl.	Overall
Baseline	100.0	91.6	100.0	100.0	95.5	97.7
Blurred	0.8	91.6	100.0	100.0	97.0	68.1
Underexposed	100.0	38.1	100.0	100.0	68.2	82.8
Overexposed+glare	100.0	54.5	82.8	100.0	100.0	88.3
Eyelid closed	100.0	84.7	100.0	0.0	100.0	76.9

C. Iris Localization Accuracy

Across five synthetic trials with known ground-truth pupil/iris centers (Section V-B2), the Hough-circle localization function achieved a mean center error of 1.18 px, against ground-truth radii of 16–24 px under 7% of the smallest tested radius. Because the downstream crop step (Section IV-B1) uses a generous $2.4\times$ radius margin around the detected center specifically to tolerate localization error, this level of center accuracy is sufficient for reliable iris-region cropping even though the detected radius in these trials consistently overestimated the true pupil radius — an artifact of the Hough transform responding to the outer iris/sclera boundary in our synthetic images rather than the inner pupil boundary used as ground truth, which does not affect crop correctness given the margin but is noted here for completeness.

D. Iris Template Matching

The iris rows of Table II and Fig. 5 report the self/genuine/impostor discrimination trial (Section V-B3). Self-matching again saturates at 100%. The genuine trial (same procedurally generated texture, mildly blurred to simulate a second capture) scored 80.2% combined, correctly higher than the impostor trial's 57.3%, with the clearest separation on the ORB component (59.7% vs. 6.2%) the reverse emphasis from the fingerprint results, where Stage 1 (minutiae geometry) separated genuine from impostor more strongly than the ORB component. This is consistent with the two modalities' descriptor stages playing different roles: fingerprint ORB descriptors confirm a geometry-driven match, while iris LBP-histogram similarity between structurally similar synthetic textures was less discriminating than the ORB keypoint stage in this trial. As with the fingerprint

genuine trial, the 80.2% combined score placed this synthetic “same-eye” trial in the retry tier rather than accepted arguably the safer of the two possible mistakes for a borderline synthetic perturbation, and a concrete illustration of why the system uses four decision tiers (Section IV-C) rather than a binary

accept/reject threshold: a binary threshold set to accept this trial risks accepting weaker impostor matches, while one set to reject it forecloses genuinely recoverable borderline captures; the retry tier defers the decision instead of forcing either error.

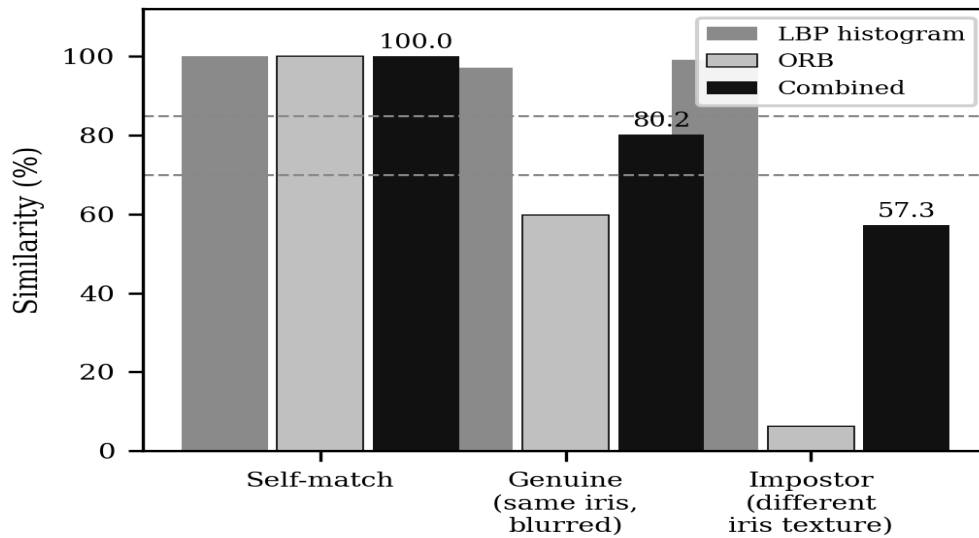


Fig. 5. Iris template matching scores: self, genuine (same synthetic iris, blurred), and impostor (different synthetic texture) trials.

VII. COMPARATIVE ANALYSIS WITH RECENT WORK

Table VI positions the system described in this paper against the four biometric-voting proposals [4, 5, 6, 7] and one general-purpose tri-modal fusion system [23] reviewed in Section II, along the dimensions raised in the gap analysis (Section III). Direct quantitative

accuracy comparison is not meaningful here: the four voting papers report figures obtained on their own benchmark or field data, while our own results (Section VI) are explicitly algorithmic-validation figures on synthetic data (Section V-D); Table VI therefore compares design characteristics rather than headline accuracy.

TABLE V: DESIGN COMPARISON WITH RECENT BIOMETRIC-VOTING AND FUSION SYSTEMS

System	Modalities	Fusion	Template protection	Capture-time privacy	Human-confirm checkpoint	Duplicate check
Umar et al. [4]	Fingerprint+Iris	Bi-factor	Not specified	Not addressed	Not described	Not described
Zamir et al. [5]	Biometric+ID link	Single	Not specified	Not addressed	Not described	Not described
Ajao et al. [6]	Multimodal+chain	DL fusion	Ledger-based	Not addressed	Not described	Not described
Faruk et al. [7]	Multimodal+chain	DL fusion	Ledger-based	Not addressed	Not described	Nat'l DB scale

Rao et al. [23]	Face+FP+Iris	DeepFusion	Privacy-preserving	Not addressed	Not described	Not described
This work	Fingerprint+Iris	Score-level	AES-256-GCM	Blurred capture, iris-only	Explicit, both flows	Precinct-scale, every enrollment

Three observations follow. First, only Rao et al. [23] among the compared systems specifies template protection explicitly comparable in intent to ours, though at the cost of a third enrolled modality (face) that our two-modality design avoids, trading a small amount of theoretical accuracy headroom [23] for lower enrollment burden consistent with the field-failure-rate motivation for the fingerprint-plus-iris pairing specifically identified in [4]. Second, none of the five compared systems describes what is displayed to an operator during live capture; our blurred-preview design (Section IV-B1) is, to our knowledge, not previously reported in this specific literature thread. Third, none describes a human-confirmation checkpoint distinguishing a supervised polling-place deployment from a self-service kiosk; our officer-confirmation workflow (Section IV-E) and the defect it was designed to prevent (Section VI-E, T-14) are, again to our knowledge, not previously discussed in this literature.

Against the algorithmic literature specifically, our fingerprint pipeline occupies a deliberately more classical position than the deep-learning minutiae extractors in [12, 13]: we trade the higher benchmark accuracy those methods report for a CPU-only, no-training-data, fully offline implementation, following the same trade-off analysis as the Haar-cascade-versus-deep-learning face-detection comparison in [29] (Section II-G), which we treat as the more relevant precedent for our operating constraints than raw accuracy leaderboards. Our iris texture representation (LBP plus ORB) is lighter-weight than a full Gabor-phase IrisCode, trading some matching precision — visible in the modest genuine/impostor margin in the iris rows of Table II relative to the wider separations typically reported for IrisCode-based systems — for substantially lower computational cost, appropriate to the same CPU-only constraint.

VIII. DISCUSSION

A. Interpretation of Results

The algorithmic results (Section VI-A, VI-D) show correct directional discrimination genuine trials consistently outscore impostor trials on both modalities but with genuine-trial combined scores landing in the retry rather than accepted tier for both fingerprint and iris. We read this primarily as a property of our synthetic perturbation magnitude rather than as evidence the matchers are miscalibrated: real same-session recapture variation for an inked thumbprint or a webcam iris is plausibly gentler than a 4° rotation with independent pixel noise, or a full Gaussian blur pass, applied deliberately to stress-test discrimination. The more actionable finding is architectural: in both cases, the tiered decision system (Section IV-C) routed a borderline-but-genuine capture to a retry rather than a false reject, which is the intended behavior of a four-tier design and would not have been available with a single binary threshold.

The quality-metric validation (Section VI-B) and localization-accuracy results (Section VI-C) are, we believe, the strongest evidence in this paper, precisely because they test narrower, more mechanically verifiable claims does a specific metric respond to a specific targeted degradation; how far off-center is a specific localization estimate than the harder problem of population-scale genuine/impostor separation. The system-level results are similarly narrow in scope (software correctness of the supervised workflow, not biometric accuracy) but directly support the paper's central design claim: that an officer-supervised checkpoint changes measurable system behavior, since the specific defect it fixes was itself only detectable through this kind of end-to-end, browser-driven testing rather than unit-level API testing.

B. Limitations

We state these plainly rather than deferring them to a single closing paragraph.

(1) Detector precision: the Haar-cascade-plus-Hough-circle iris detector (Section IV-B) is measurably less precise than landmark-based detection under real, uncontrolled lighting [29, 30], and we could not validate this gap directly without live capture hardware.

(2) Liveness robustness: blink-based liveness defeats static-photo presentation attacks but not video replay [19, 20], and our EAR proxy is a bounding-box approximation rather than a landmark-accurate measurement.

(3) Threshold calibration: the 95/85/70 decision tiers, the 98% duplicate threshold, and the fusion weights are principled defaults consistent with the literature reviewed in Section II, not values fitted to a labeled dataset of genuine and impostor comparisons; production deployment requires calibration against real biometric data with reported false-accept/false-reject curves.

(4) Key management: the AES-256 master key is generated once and stored locally for this reference implementation; production use requires a managed key-management service or hardware security module with rotation. **(5) Scale:** the duplicate-enrollment check (Section IV-A5) is a linear scan appropriate to precinct scale, not the national scale addressed in [9].

C. Implications for Deployment

The officer-confirmation design (Section IV-E) is, we argue, the right default specifically for a supervised polling-place deployment and would be the wrong default for an unsupervised self-service kiosk, where requiring a human confirmation click at every step has no one present to provide it. This is not a limitation of the confirmation design so much as a scope boundary: the system as described targets the supervised setting, and an unsupervised variant would need a different resolution to the same underlying question what happens on a borderline or failed capture that does not depend on an available human operator.

CONCLUSION

This paper presented the design, implementation, and evaluation of a two-modality fingerprint-and-iris biometric authentication system for officer-supervised voter verification. The system combines a classical two-stage fingerprint matcher, a privacy-preserving iris capture pipeline that never displays or stores more of the face than the iris region, AES-256-GCM template encryption, score-level fusion with a four-tier decision system, precinct-scale duplicate-enrollment detection, and the paper's central design contribution an explicit officer-confirmation checkpoint that replaces automatic capture-to-decision advancement with a required human action on every registration and, on authentication, offers retry rather than any automated path forward whenever a capture fails to match. Reviewing 2021–2026 literature across fingerprint and iris recognition, liveness detection, multimodal fusion, template protection, and applied biometric voting, we identified that capture-time privacy and human-supervision checkpoints are addressed rarely or not at all in comparable systems, and positioned this system's design directly against that gap.

Evaluation combined algorithmic validation on controlled synthetic and perturbed biometric imagery with system-level validation of the complete application through automated browser interaction, deliberately avoiding a fabricated population-scale accuracy figure given the absence of a labeled dataset and live capture hardware in the development environment. Results showed correct self-match saturation, correct directional genuine/impostor separation for both modalities, sub-2-pixel mean iris-localization error, correctly independent responses of all five iris quality sub-metrics to targeted degradations, and a fully passing final system-level regression suite reached only after identifying and correcting two real defects, including the auto-advance-on-non-match behavior that directly motivated the officer-confirmation redesign, reported here transparently as part of the evaluation rather than omitted.

Future work includes replacing the Haar-cascade/Hough-circle iris detector with landmark-based detection where model-download network access is available; calibrating decision thresholds

against a real, labeled, multi-subject biometric dataset with reported false-accept/false-reject curves; strengthening liveness detection against video-replay attacks; integrating a managed key-management service; and extending enrollment to both eyes and additional fingers for higher-assurance deployments.

Funding

The authors received no specific funding for this research.

Data Availability

The datasets generated and/or analyzed during the current study are not publicly available due to biometric data privacy considerations. Processed experimental data and supporting information may be made available by the corresponding author upon reasonable request, subject to applicable privacy and ethical considerations.

REFERENCES

1. Sumalatha et al., "Comprehensive Review of Unimodal and Multimodal Fingerprint Biometric Authentication Systems: Fusion, Attacks, and Template Protection," *IEEE Access*, 2024.
2. R. Es-Sobhahi et al., "A systematic review of multimodal biometric fusion levels and strategies," *IET Biometrics*, 2025.
3. J. Kamlaskar et al., "Iris and fingerprint multimodal biometric fusion for identity authentication," *AIMS Electronics and Electrical Engineering*, 2021.
4. B. Umar et al., "A Bi-Factor Biometric Authentication System for Secure Electronic Voting System," in *Proc. IEEE Nigeria 4th Int. Conf. Disruptive Technologies for Sustainable Development (NIGERCON)*, 2022.
5. M. Zamir et al., "Electronic Voting Machine with Biometric Authentication," in *Proc. INDIACom*, 2022.
6. K. Ajao et al., "Blockchain-Based Multimodal Biometric and Deep Learning Framework for Secure Electronic Voting," *IEEE Access*, 2025.
7. M. A. Faruk et al., "Blockchain and biometric based secure online voting system," *Cluster Computing*, 2024.
8. C. Nlend et al., "Large-Scale Fingerprint Deduplication for Voter Registration Databases," *arXiv preprint*, 2021.
9. S. Koner et al., "Bharat ABIS: Billion-Scale Multimodal Biometric Search and Deduplication," 2026.
10. R. Martins et al., "Fingerprint Recognition in Forensic Scenarios: A Crossing-Number and Gabor-Enhancement Approach," *Sensors*, 2024.
11. T. Situmorang, "Crossing Number Method for Mobile-Camera-Captured Fingerprint Biometrics," 2022.
12. Q. Zhang et al., "A multi-task deep convolutional neural network for contactless fingerprint minutiae extraction," *Pattern Recognition*, 2021.
13. A. Abdulkarim et al., "Hybrid Deep Learning and Classical Enhancement Pipeline for Fingerprint Identification," 2025.
14. P. Manchanda et al., "CLAHE and Gaussian Smoothing for Latent Fingerprint Image Enhancement," 2025.
15. M. Taha et al., "Local Binary Pattern-Based Iris Feature Extraction Across Near-Infrared and Visible Spectra," 2021.
16. S. Jan et al., "Viola-Jones and Geometrical Information Based Eye/Iris Segmentation for Non-Ideal Iris Recognition," 2021.
17. E. Garea-Llano et al., "Iris-Pupil Region Quality Assessment Framework Aligned with ISO/IEC 19794-6," 2021.
18. Z. Fazilov et al., "PCA-Based Combination of Iris Image Quality Parameters," 2022.
19. E. Demirsoy et al., "Eye-Aspect-Ratio-Based Blink Liveness Detection for Access Control," 2026.
20. H. Chong et al., "Eye Aspect Ratio Blink Detection for Masked Face Recognition Liveness Verification," 2025.
21. V. C. Kulloli et al., "Iris Liveness Detection for Video Replay Attack Using Quality Metrics and EfficientNetB7," in *Proc. IEEE Int. Conf. Data and Software Engineering (ICoDSE)*, 2025.
22. A. Kharchevnikova and A. Savchenko, "Best-frame selection for video-based face recognition," *PeerJ Computer Science*, 2021.
23. P. Rao et al., "DeepFusion: Tri-Modal Privacy-Preserving Biometric Fusion of Face, Fingerprint, and Iris," 2026.

24. M. Imran et al., "AES-Encrypted Minutiae Cylinder Code Template Protection for Fingerprint Biometrics," IEEE Access, 2025.
25. H. Ayoup et al., "AES-Based Cancelable Template Generation for Combined Eye and Fingerprint Biometrics," Intelligent Automation & Soft Computing, 2022.
26. W. El-Shafai et al., "Genetic-Algorithm-Optimized Cancelable Biometric Encryption Scheme," IEEE Access, 2021.
27. M. Bansal et al., "A comparative analysis of SIFT, SURF, and ORB feature descriptors," 2021.
28. L. Xie et al., "Adaptive Histogram Equalization Combined with ORB Feature Matching," 2022.
29. S. Adeshina et al., "Custom Face Classification Model for Classroom Using Haar-Like and LBP Features with Their Performance Comparisons," Electronics, vol. 10, no. 1, 2021.
30. K. Elbert et al., "Comparative Evaluation of Haar Cascade, Dlib, and MediaPipe for Face and Eye Detection," 2025.

HOW TO CITE: P. Sri Sai Krishna*, Dr. B. Kezia Rani, Enhancing Secured Voter Authentication For E-Voting System With Fusion Of Multimodal Biometric Traits Using Cryptographic Keys, Int. J. Sci. R. Tech., 2026, 3 (10), 1100-1118. <https://doi.org/10.5281/zenodo.23528526>