JCTA Journal of Computing Theories and Applications ISSN:3024-9104

Research Article

Secure Lightweight Face Authentication with MTCNN and LightCNN for Digital Financial Services

Dendy K Pramudito [1,*], Jufriadif Na'am [1], and Ferda Ernawan [2]

[1] Informatics Doctoral Program, Universitas Nusa Mandiri, Jakarta, 12540, Indonesia; e-mail : 16240003@nusamandiri.ac.id; jufriadifnaam@nusamandiri.ac.id
[2] Computing Program, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Pahang, 26600, Malaysia; e-mail : ferda@umpsa.edu.my
* Corresponding Author : Dendy K Pramudito

Abstract: Mobile face authentication for digital financial services must simultaneously satisfy recognition accuracy, computational efficiency, and biometric security under resource-constrained deployment conditions. This study proposes and evaluates a lightweight face-authentication framework that integrates detector–recognition pipeline optimization, protected biometric-template transformation, and blockchain-backed integrity support. Five face detection–recognition pipelines were systematically evaluated using a shared LightCNN-29v2 backbone fine-tuned on the Indonesian Muslim Student Face Dataset (IMSFD), with Mahalanobis-based Distance-Based Encryption (DBE) providing protected template matching and blockchain hash anchoring serving as an architectural integrity layer. Experiments on 3,660 images from 68 identities demonstrate that the MTCNN + LightCNN pipeline achieves the most favorable in-domain performance, reaching 94.95% accuracy, a ROC-AUC of 0.9970, an F1-score of 0.95, and successful processing of 3,546 out of 3,660 test images with an overall model size of approximately 5 MB. Applying Mahalanobis-based DBE further improves verification performance on IMSFD, increasing accuracy to 96.88% while reducing the False Acceptance Rate (FAR) from 0.83% to 0.18% and the False Rejection Rate (FRR) from 16.44% to 10.12%. Cross-dataset evaluation on LFW, CFP-FF, CFP-FP, AgeDB-30, CALFW, and CPLFW indicates that the proposed framework generalizes well to frontal-domain benchmarks but exhibits expected performance degradation under cross-pose and cross-age conditions due to distribution shift. Overall, the results demonstrate that detector selection is the dominant factor influencing end-to-end verification performance, while domain-specific fine-tuning and protected template matching are essential for secure and practical deployment in mobile financial authentication systems.

Keywords: Biometric authentication; Biometric template protection; Blockchain; Digital financial services; Distance-based encryption; Face recognition; LightCNN; Mobile authentication.

Received: May, 26th 2026
Revised: July, 2nd 2026
Accepted: July, 4th 2026
Published: July, 18th 2026

1. Introduction

Digital financial services increasingly rely on biometric authentication to replace traditional credential-based methods such as passwords and PINs. Among biometric modalities, facial recognition has gained particular traction in smartphone-based banking because of its suitability for electronic Know Your Customer (eKYC) and transaction authorization [1], [2]. This trend is especially evident in rapidly expanding digital finance markets such as Indonesia, where increasing smartphone penetration has created a growing demand for secure, scalable, and accessible identity verification.

Despite its widespread adoption, mobile face authentication remains technically and operationally challenging. Real-world deployment environments expose recognition systems to variations in camera quality, illumination, facial pose, and limited on-device computational resources, all of which are common across heterogeneous smartphone ecosystems [2]. These challenges are further compounded by adversarial threats, including spoofing, fake-video injection, and presentation attacks. Consequently, a model that performs well on controlled

benchmark datasets may become unreliable or computationally impractical when deployed on consumer-grade mobile devices [3]–[5].

Lightweight architectures such as MobileNetV3 [6] and LightCNN [7] have been developed to address computational constraints, and subsequent studies have demonstrated their practical suitability for edge-device deployment [8], [9]. Recognition accuracy alone, however, is insufficient for secure financial authentication. Storing raw biometric templates introduces irreversible privacy and security risks that increase with deployment scale [10], while privacy-preserving biometric systems must mitigate template leakage, unauthorized disclosure, and trust management throughout the authentication lifecycle [11]. Distance-Based Encryption (DBE) using a Mahalanobis-based formulation enables matching in a protected embedding space without exposing raw biometric templates [10], [12]. Blockchain-backed record management further provides tamper-evident integrity and traceability, although these benefits introduce additional coordination overhead that must be carefully balanced against deployment requirements.

The research gap addressed in this study is not whether lightweight face recognition is feasible on smartphones, nor whether template protection and distributed integrity mechanisms can independently enhance security. Rather, it is whether these components can be integrated into a single mobile authentication framework that simultaneously achieves recognition accuracy, computational efficiency, and practical security for digital financial services on resource-constrained hardware. Existing studies have largely treated model efficiency and security enhancement as separate research directions, with relatively few investigations examining their combined interaction within a finance-oriented deployment context [2], [11].

Among face detectors suitable for mobile deployment, BlazeFace [13] targets sub-millisecond GPU inference but provides a less mature TensorFlow Lite deployment path for CPU-only environments. RetinaFace [14] achieves robust face localization under unconstrained conditions but incurs higher computational cost than required for the target deployment scenario. YOLO-based detectors emphasize detection throughput but do not directly provide the facial landmark estimation required for alignment. MTCNN [15] was selected in this study for three reasons: its mature TensorFlow Lite conversion pipeline for Android deployment, its joint face detection and five-point landmark estimation that eliminates the need for a separate alignment module, and its cascade-based architecture that effectively suppresses false positives while maintaining a model size below 5 MB. LightCNN-29v2 was selected over MobileFaceNets [16] and ShuffleFaceNet [17] because its Max-Feature-Map (MFM) activation has been reported to improve robustness against noisy and low-quality facial inputs [7], making it well suited to the image variability encountered on entry-level smartphone cameras.

Based on these considerations, this study investigates whether lightweight face recognition, protected biometric-template matching, and blockchain-based integrity can be effectively integrated into a practical authentication framework for digital financial services. Five face-detection strategies are evaluated using a shared LightCNN-29v2 recognition backbone, with eKYC onboarding and transaction authorization serving as representative deployment scenarios. Rather than pursuing state-of-the-art recognition accuracy on public benchmarks, this work emphasizes deployment-oriented evaluation under resource-constrained environments by jointly considering recognition performance, computational efficiency, template protection, and system integrity. This study makes the following four contributions:

- A controlled comparison of five detector–recognition pipelines, using a common LightCNN-29v2 backbone to isolate the effect of detector selection under identical evaluation conditions.
- A deployment-oriented evaluation framework that jointly considers recognition performance, verification reliability, computational efficiency, and detection coverage for mobile financial authentication.
- An empirical evaluation of Mahalanobis-based Distance-Based Encryption (DBE) for protected biometric-template matching under both in-domain and cross-domain conditions.
- The integration of blockchain-based hash anchoring using Ethereum Testnet and IPFS as an architectural integrity layer to support tamper-evident biometric record management within the proposed authentication framework.

The remainder of this paper is organized as follows. Section 2 reviews the related work. Section 3 describes the proposed methodology. Section 4 presents experimental results. Section 5 discusses the key findings, positions the proposed approach relative to related studies, and outlines the limitations of the current work. Finally, Section 6 concludes the paper.

2. Literature Review

2.1. Face Biometrics in Digital Finance

Face-biometric authentication has become a key technology for digital financial services because it relies on unique physiological characteristics rather than conventional knowledge-based credentials. Smartphone-based facial recognition is widely adopted for electronic Know Your Customer (eKYC) onboarding and transaction authorization owing to its convenience and compatibility with standard mobile cameras [1], [2]. As smartphone-based financial transactions continue to expand, particularly in markets such as Indonesia, authentication systems must simultaneously satisfy security, scalability, and user convenience.

Unlike general-purpose face recognition, financial authentication demands substantially higher reliability because authentication failures may lead to account takeover, fraudulent transactions, or denial of access for legitimate users [18]. Practical deployment must therefore remain robust against heterogeneous device conditions, illumination changes, pose variation, and adversarial threats such as spoofing, replay attacks, and presentation attacks [2]–[4]. Consequently, recent research increasingly evaluates mobile face-authentication systems not only by recognition accuracy but also by inference efficiency, detection robustness, and secure biometric data management.

2.2. Lightweight Vision Models for Efficient Mobile Deployment

The demand for on-device intelligence has accelerated the development of lightweight deep-learning architectures optimized for resource-constrained environments. MobileNetV3 combines hardware-aware neural architecture search with efficient network design to reduce computational cost while maintaining competitive performance [6]. Subsequent studies have confirmed that lightweight architectures can substantially reduce inference time and memory consumption, making them suitable for mobile and edge deployment [8], [9]. In mobile face authentication, however, overall performance depends on both face detection and face recognition rather than either stage alone. Most previous studies focus primarily on recognition backbones, whereas the influence of detector selection on end-to-end verification performance remains comparatively underexplored.

Among lightweight detectors, MTCNN provides joint face localization and five-point landmark estimation for accurate alignment [15]. Classical approaches such as DLib HOG remain competitive when combined with preprocessing techniques including CLAHE and image upsampling, whereas directly applying a general-purpose classifier such as MobileNetV3 for face detection generally produces inferior localization because it was not designed for that task. For face representation, LightCNN employs Max-Feature-Map (MFM) activation to generate compact and discriminative facial embeddings with improved robustness to noisy inputs [7]. Its lightweight architecture enables efficient L2-normalized embedding generation with low memory overhead, making it suitable for resource-constrained mobile deployment [2], [7]. Although lightweight face-recognition models have demonstrated promising performance, comparatively few studies have systematically investigated how detector selection affects end-to-end verification under identical deployment conditions. This motivates the controlled detector-comparison strategy adopted in this work.

2.3. Secure Template Storage and Matching

Protecting biometric templates is essential because, unlike passwords, biometric traits cannot be replaced once compromised. Secure biometric systems therefore avoid storing raw templates and instead employ protection mechanisms that reduce leakage, inversion, and unauthorized reuse while preserving verification capability [10], [11], [19], [20]. This requirement is particularly important for mobile financial authentication, where facial embeddings contain sensitive identity information.

Distance-Based Encryption (DBE) enables biometric matching in a protected embedding space without exposing raw templates [10]. In addition, Mahalanobis distance considers

feature covariance, providing a similarity measure that better reflects the statistical distribution of high-dimensional facial embeddings than conventional Euclidean distance [12]. Recent studies further emphasize that biometric security should be evaluated throughout the authentication lifecycle, including enrollment, template storage, matching, revocation, and auditing [11].

Despite these advances, existing research has largely emphasized privacy preservation itself, while comparatively fewer studies have examined the empirical impact of protected template matching on verification performance under lightweight mobile deployment. This motivates the evaluation of Mahalanobis-based DBE across both in-domain and cross-domain scenarios in this study.

2.4. Tamper-Evident Records and Decentralized Trust via Blockchain

Blockchain has attracted increasing attention in digital identity, Know Your Customer (KYC), and biometric authentication because it provides tamper-evident records, auditability, and decentralized trust [21]–[23]. Rather than storing biometric data directly, blockchain typically maintains cryptographic hashes or references to protected templates, enabling integrity verification without exposing sensitive biometric information. This capability improves traceability and reduces single points of failure in biometric identity management [24].

Despite these benefits, blockchain introduces additional latency, infrastructure complexity, and protocol management overhead [23], [25], [26]. It should therefore be viewed as a complementary integrity mechanism rather than a replacement for efficient biometric recognition or protected template storage. Existing studies primarily emphasize blockchain-enabled trust and record integrity, whereas comparatively few investigate its integration with lightweight face-authentication systems for practical mobile deployment. Accordingly, blockchain is incorporated in this work as an architectural integrity layer that complements protected biometric matching while maintaining deployment feasibility.

2.5. Research Gap and Contribution

The literature review indicates that lightweight face recognition, secure biometric-template protection, and blockchain-based identity integrity have largely been investigated as separate research directions. Existing studies primarily optimize recognition accuracy or computational efficiency, strengthen template security, or improve identity integrity through decentralized trust. Comparatively few studies examine how these complementary components can be integrated into a unified authentication framework suitable for deployment in resource-constrained mobile financial services.

Accordingly, the objective of this study is not to achieve state-of-the-art recognition accuracy on public benchmarks, but to investigate whether lightweight recognition, protected biometric-template matching, and blockchain-based integrity can be effectively combined into a practical authentication pipeline. Unlike high-performing recognition models such as MobileFaceNet [16], which are primarily optimized for benchmark performance, the proposed approach emphasizes deployment-oriented requirements, including computational efficiency, fraud-sensitive verification, protected template management, and system integrity. The methodology developed to address this gap is presented in the following section.

3. Proposed Method

This study proposes a deployment-oriented mobile face-authentication framework designed for digital financial services operating on resource-constrained devices. Rather than optimizing a single recognition model, the framework evaluates the combined effect of face detection, feature extraction, protected biometric-template matching, and integrity support within a unified authentication pipeline. The primary design objective is to identify a detector–recognition configuration that balances recognition performance, computational efficiency, and security requirements under realistic mobile deployment conditions [18].

To enable a fair comparison, all candidate pipelines share the same LightCNN-29v2 recognition backbone, while only the face-detection component is varied. This design isolates detector selection as the primary experimental variable, allowing performance differences to be attributed to face localization quality rather than changes in the recognition model. In addition, protected template matching using Mahalanobis-based Distance-Based Encryption

(DBE) and blockchain-based integrity support are incorporated as complementary security layers to evaluate their suitability for practical mobile financial authentication.

3.1. System Architecture

The proposed framework consists of seven sequential stages: face detection, preprocessing, feature extraction, identity embedding, identity verification, template protection, and integrity support, as illustrated in Fig. 1. The overall design follows a modular architecture in which each component performs a distinct function while remaining independent of the others, allowing individual modules to be evaluated or replaced without modifying the complete authentication pipeline.

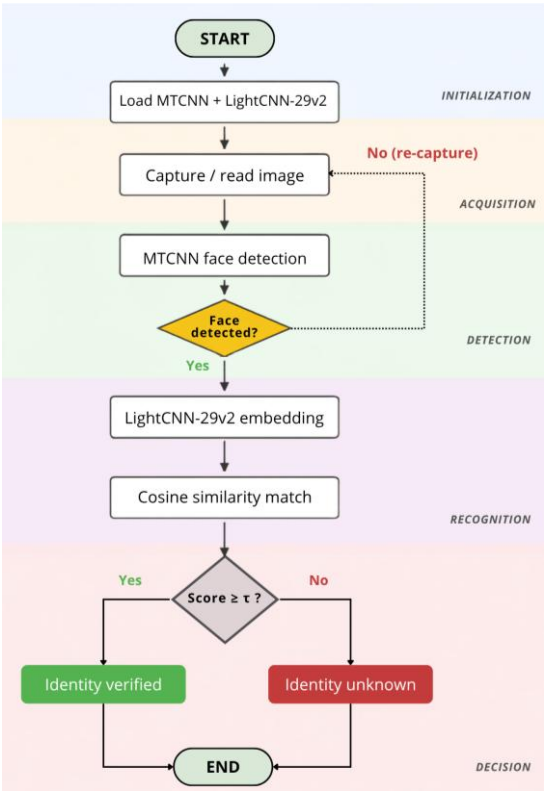

Figure 1. Overall workflow of the proposed lightweight mobile face-authentication framework.

The framework operates through two phases: enrollment and authentication. During enrollment, facial images are first detected and aligned before being resized to 128 × 128 grayscale images. The processed images are then passed through LightCNN-29v2 [7], which generates compact 256-dimensional L2-normalized facial embeddings. A mean embedding is computed for each identity and stored as the gallery representation, producing a gallery matrix of size (68 × 256) for subsequent verification.

During authentication, an input image captured by a smartphone camera is first processed by the selected face detector. MTCNN [15], the primary detector adopted in the proposed framework, performs face localization and simultaneously estimates five facial landmarks through its three-stage cascade (P-Net, R-Net, and O-Net), enabling consistent facial alignment before recognition. Images in which no valid face is detected are excluded from subsequent processing to avoid propagating localization errors into the recognition stage. The cropped facial region is resized to 128 × 128 grayscale and forwarded to LightCNN-29v2 to produce a normalized query embedding. Identity verification is performed by comparing the query embedding with the gallery using cosine similarity,

$$sim(q, g) = q^T g \tag{1}$$

where q and g denote the L2-normalized query and gallery embeddings, respectively. In recognition mode, the identity with the highest similarity score is returned through nearest-

neighbor retrieval. For verification, the similarity score is compared against a decision threshold τ; authentication is accepted when $sim(q, g) \geq \tau$, otherwise it is rejected. The operating threshold is determined experimentally and discussed in Section 3.5.

The layer-level architecture of the selected pipeline is illustrated in Fig. 2. MTCNN performs robust face localization and landmark detection, while LightCNN-29v2 extracts compact facial representations using Max-Feature-Map (MFM) activation. This combination was selected because it provides a practical balance between detection robustness, recognition capability, and computational efficiency for deployment on entry-level mobile hardware.

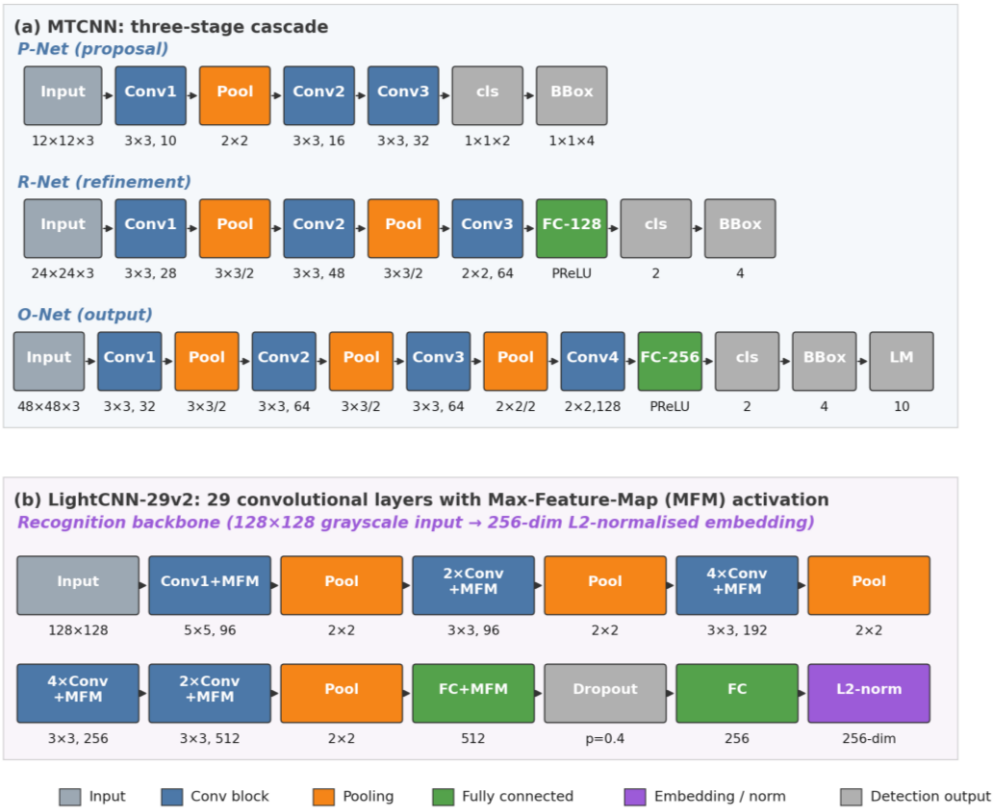

Figure 2. Layer-level architecture of the selected authentication pipeline: (a) MTCNN three-stage face detector [15]; (b) LightCNN-29v2 recognition backbone [7].

The functional role of each component is summarized in Table 1.

Table 1. Core components and functional specifications of the proposed biometric authentication framework.

Component	Model/Method	Specification	Primary Function
Face Detection	MTCNN (selected)	Confidence = 0.9; margin = 20 px	Face localization and landmark estimation
Face Recognition	LightCNN-29v2	128 × 128 grayscale; 256-dimensional L2-normalized embedding; MFM activation [7]	Compact facial representation
Similarity Metric	Cosine Similarity	Adaptive verification threshold	Identity verification
Template Protection	Mahalanobis-based DBE	Matching in protected embedding space; covariance estimated from IMSFD training set	Protected biometric-template matching
Integrity Layer	Ethereum Testnet + IPFS	Hash anchoring	Tamper-evident integrity and auditability
Database	PostgreSQL	Binary embedding storage (BYTEA)	Protected template storage

3.2. Data Sources and Evaluation Setting

The proposed framework was evaluated using the Indonesian Muslim Student Face Dataset (IMSFD), which comprises 68 identities distributed across three subsets (A1, A2, and B1) with a total of 3,660 testing images. The dataset includes frontal, near-frontal, and side-profile facial poses acquired under varying indoor lighting conditions, reflecting realistic smartphone-based eKYC and financial authentication scenarios in Indonesia. Fine-tuning was performed using 12,246 face-cropped training images, while all 3,660 testing images were reserved for evaluation [39].

Figure 3. Representative anonymized face samples from the IMSFD dataset (shown with face blurring applied for privacy protection).

IMSFD was selected as the primary evaluation dataset because it closely represents the intended deployment domain. Unlike widely used public benchmarks, the target demographic of Indonesian mobile financial services is underrepresented in datasets such as LFW, CFP, and AgeDB. The NIST Face Recognition Vendor Test (FRVT) Part 3 [28] reports demographic-dependent performance variation across population groups, while Buolamwini and Gebru [29] further demonstrate measurable accuracy disparities for non-Western populations. Consequently, IMSFD provides a more representative evaluation environment for the target application than general-purpose public benchmarks.

Public benchmark datasets play a complementary role in this study. Rather than serving as the primary evaluation datasets, LFW, CFP, AgeDB, CALFW, and CPLFW are used to assess out-of-distribution performance under cross-domain, cross-pose, and cross-age conditions. Accordingly, IMSFD is used to evaluate deployment-specific performance, whereas the public benchmarks provide additional evidence of model generalization.

Dataset annotation was performed using SCRFD (det_size = 320 × 320), producing standardized 128 × 128 grayscale face crops with coverage ranging from 92% to 100% across the three subsets. The resulting dataset consists of 12,246 training crops and 3,566 gallery-compatible testing crops. Considering that the objective of this study is a controlled comparison of detector–recognition pipelines rather than population-scale benchmarking, the scale of IMSFD is appropriate for isolating the influence of detector selection under consistent evaluation conditions [27], [30]. The limitation regarding broader population-level generalization is discussed further in Section 5.3.

3.3. Candidate Pipelines, Preprocessing, and Embedding Generation

To isolate the influence of face detection on end-to-end authentication performance, five candidate pipelines were evaluated while maintaining LightCNN-29v2 as the common recognition backbone. All configurations generate 256-dimensional L2-normalized facial embeddings and perform gallery-based nearest-neighbor matching using cosine similarity. The evaluated configurations are:

- P1: MTCNN + LightCNN
- P2: DLib HOG + LightCNN (with CLAHE and 2× upsampling)
- P3: MobileNetV3 + LightCNN (ImageNet pretrained, sliding-window detection, without task-specific fine-tuning)

- P4: LightCNN only (without an external face detector)
- P5: DLib CNN + LightCNN

By fixing the recognition backbone across all configurations, performance differences can be attributed primarily to the face-detection stage rather than variations in feature representation. This controlled experimental design enables a more objective assessment of detector suitability for lightweight mobile authentication.

LightCNN-29v2 was fine-tuned on IMSFD to adapt the pretrained facial representations to the target deployment domain while preserving the compact network architecture. Training employed the Adam optimizer with discriminative learning rates for different network layers, Cosine Annealing Warm Restarts scheduling, early stopping, and standard image augmentation, including horizontal flipping, slight rotation, and brightness–contrast variation. The classification layer was reinitialized for the 68 IMSFD identities, whereas the remaining network parameters were initialized from the pretrained MS-Celeb-1M model. Complete training hyperparameters are summarized in Table 2.

After fine-tuning, the model was converted to TensorFlow Lite (.tflite) to support on-device deployment. Inference performance was evaluated using a Raspberry Pi 4 Model B (4 GB RAM) operating under CPU-only conditions to approximate the computational constraints of entry-level mobile hardware, consistent with previous deployment-oriented biometric authentication studies [27], [30].

Table 2. Experimental design and evaluation settings.

Component	Description
Research design	Comparative experimental evaluation with laboratory validation and deployment-oriented simulation
Evaluation dataset	IMSFD: 3,660 testing images (68 identities)
Training dataset	12,246 annotated face crops
Inference platform	Raspberry Pi 4 Model B (4 GB RAM), CPU-only
Candidate pipelines	MTCNN + LightCNN; DLib HOG + LightCNN; MobileNetV3 + LightCNN; LightCNN only; DLib CNN + LightCNN
Recognition backbone	LightCNN-29v2 fine-tuned on IMSFD [7]; 256-dimensional L2-normalized embeddings
Matching strategy	Gallery-based nearest-neighbor using cosine similarity
Evaluation metrics	Accuracy, F1-score, ROC-AUC, TPR, TNR, FAR, FRR, detection coverage, and inference latency
Security components	Mahalanobis-based DBE for protected template matching; Ethereum Testnet and IPFS for blockchain-based integrity support

3.4. Template Protection and Blockchain Integration

To prevent exposure of sensitive biometric information, the proposed framework stores protected facial templates rather than raw embeddings. Following feature extraction, each embedding is transformed using Mahalanobis-based Distance-Based Encryption (DBE), enabling template matching while preserving the privacy of the original biometric representation. This design separates identity verification from raw template storage, reducing the risk associated with template leakage or database compromise. The protected template is generated through the affine transformation

$$E(x) = A_x + b \tag{2}$$

where x denotes the original facial embedding, A is a random orthogonal matrix, b is a random translation vector, and $E(x)$ represents the protected template stored in PostgreSQL as a binary object.

Verification is subsequently performed in the protected embedding space using Mahalanobis distance,

$$d_M\big(E(x), E(y)\big) = \big(E(x) - E(y)\big)^T \Sigma^{-1} \big(E(x) - E(y)\big) \tag{3}$$

where Σ denotes the covariance matrix estimated from the IMSFD training embeddings during enrollment. Precomputing the covariance matrix minimizes computational overhead during authentication while preserving distribution-aware similarity measurement.

To strengthen record integrity, cryptographic hashes of the protected templates are anchored to Ethereum Testnet and IPFS. Blockchain therefore functions as an architectural integrity layer that provides tamper-evident provenance without participating in the biometric matching process, consistent with previous blockchain-supported identity management and eKYC frameworks [21], [22].

3.5. Evaluation Metrics and Threshold Selection

The evaluation protocol was designed to reflect the operational requirements of mobile financial authentication rather than maximizing recognition accuracy alone. Metric selection follows the recommendations of the FIDO Alliance Biometric Component Certification, ISO/IEC 19795-1 reporting guidelines, and the NIST FRVT 1:1 verification protocol, ensuring that both security and usability are considered throughout the evaluation.

Because unauthorized access represents the primary security risk in financial services, the False Acceptance Rate (FAR) is treated as the principal security metric [18]. The False Rejection Rate (FRR) measures usability by quantifying the proportion of legitimate users rejected during authentication. ROC-AUC provides a threshold-independent assessment of discriminative capability, while detection coverage measures the proportion of input images that successfully produce valid facial embeddings. Inference latency reflects deployment feasibility for real-time mobile applications and is interpreted relative to the 100 ms interaction threshold commonly adopted in mobile user experience studies. Accuracy and F1-score are additionally reported to facilitate comparison with existing face-recognition literature but are considered secondary to FAR and FRR in the target deployment scenario.

Two complementary evaluation protocols were employed. The first is closed-set recognition, which evaluates all 3,660 testing images using gallery-based nearest-neighbor retrieval. The second is face verification, which evaluates 1,000 genuine pairs and 1,000 impostor pairs using a cosine similarity threshold to reflect practical authentication scenarios. Classification performance was computed from the confusion matrix (TP, TN, FP, and FN) using the standard definitions:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall\ (TPR) = \frac{TP}{TP + FN} \tag{6}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{7}$$

$$TNR = \frac{TN}{TN + FP}, \quad FAR = \frac{FP}{FP + TN}, \quad FRR = \frac{FN}{FN + TP} \tag{8}$$

Identity assignment follows gallery-based nearest-neighbor retrieval,

$$i^* = \arg\max_i sim(q, g_i) \tag{7}$$

where g_i denotes the mean embedding associated with identity i.

The verification threshold τ was determined through a parameter sweep over the interval [0.60,0.80] with increments of 0.01 using 1,000 genuine and 1,000 impostor pairs. The operating point was selected by minimizing FAR while satisfying the constraint FAR≤1%, reflecting the asymmetric risk profile of financial authentication [22], [31]. The resulting operating threshold, τ=0.70, achieved a FAR of 0.83% and an FRR of 16.44%, and was subsequently adopted throughout all verification experiments.

4. Results

4.1. Comparative Pipeline Performance

The comparative performance of the five evaluated pipelines is summarized in Table 3, while the corresponding ROC curves and deployment-oriented performance comparison are presented in Fig. 3 and Fig. 4, respectively. Among the evaluated configurations, MTCNN + LightCNN-29v2 achieved the strongest overall performance, obtaining the highest accuracy (94.95%), F1-score (0.95), and ROC-AUC (0.9970), while successfully processing 3,546 of 3,660 test images. DLib HOG + LightCNN, enhanced with CLAHE and 2× upsampling, achieved comparable performance (94.68% accuracy; ROC-AUC 0.9953), indicating that classical detection remains competitive when appropriate preprocessing is applied.

The remaining pipelines exhibited lower deployment suitability. DLib CNN + LightCNN achieved relatively high discriminative capability (ROC-AUC 0.9548) but substantially higher inference latency. MobileNetV3 + LightCNN, configured as a sliding-window detector without task-specific fine-tuning, produced the weakest overall performance, while LightCNN only demonstrated insufficient impostor rejection despite moderate recognition capability. These quantitative results establish MTCNN + LightCNN-29v2 as the most balanced configuration under the evaluated deployment setting.

Table 3. Comparative performance of the five evaluated face-processing pipelines.

Pipeline	Accuracy (%)	F1	ROC-AUC	TPR (%)	TNR (%)	Detection Time	Interpretation
MTCNN + LCNN (selected)	94.95	0.95	0.9970	50*	95.00	72 ms	Highest accuracy and ROC-AUC with high detection coverage
DLib HOG + LCNN	94.68	0.95	0.9953	95.00	95.00	357 ms	Competitive after CLAHE and 2× upsampling
MobileNetV3 + LCNN	33.81	0.32	0.7122	30.00	47.00	~598 ms	Unsuitable as a face detector without task-specific fine-tuning
LCNN only	63.50	0.54	0.6996	64.18	63.16	0 ms	Limited impostor rejection capability
DLib CNN + LCNN	64.42	0.65	0.9548	66.00	71.00	~1724 ms	High discriminability but excessive inference latency

*TPR at $\tau = 0.70$ (FAR-constrained banking operating point). See Section 4.3.

4.2. Architecture of the Selected Pipeline: MTCNN and LightCNN-29v2

The selected pipeline combines MTCNN for face detection and alignment with LightCNN-29v2 for facial representation learning, as illustrated in Fig. 2. MTCNN operates as a three-stage cascade (P-Net, R-Net, and O-Net) that progressively refines face localization while suppressing false positives. The final stage additionally estimates five facial landmarks, enabling facial alignment before feature extraction [15]. Aligned 128 × 128 grayscale face images are subsequently processed by LightCNN-29v2 [7], which generates a 256-dimensional L2-normalized embedding for gallery-based cosine similarity matching (Eq. (1) and Eq. (9)). The use of Max-Feature-Map (MFM) activation provides compact feature extraction while maintaining robustness to image quality variation. The complete deployment pipeline occupies approximately 5 MB, making it suitable for resource-constrained Android devices.

4.3. Note on TPR in the Verification Protocol

The TPR reported for MTCNN + LightCNN in Table 3 (50%) appears substantially lower than the corresponding recognition accuracy (94.95%). This difference arises because the two metrics are computed under different evaluation protocols. Recognition accuracy is

measured using a closed-set nearest-neighbor protocol, in which each detected face is assigned to the closest gallery identity. In contrast, TPR is computed under a verification protocol using a fixed similarity threshold ($\tau = 0.70$) on 1,000 genuine pairs. Since the threshold was selected to constrain the False Acceptance Rate (FAR) below 1%, only genuine pairs exceeding this conservative operating threshold are accepted. Accordingly, the reported TPR reflects the selected operating point rather than the intrinsic discriminative capability of the model. This interpretation is supported by the high ROC-AUC (0.9970), indicating strong class separability across the full threshold range. The chosen threshold therefore prioritizes security by minimizing false acceptance, consistent with the requirements of financial authentication, while accepting a lower genuine acceptance rate at the selected operating point.

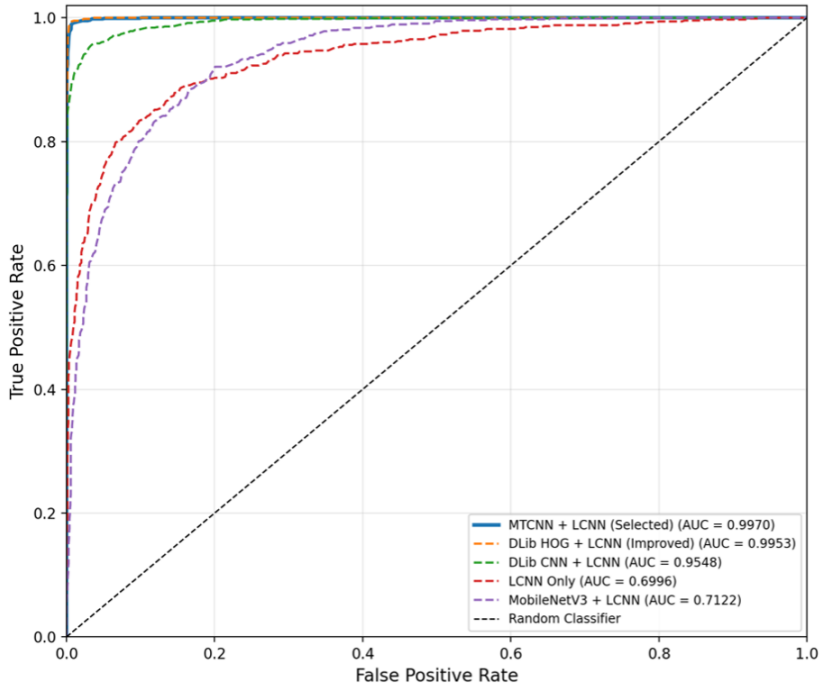

Figure 4. ROC curves of the five evaluated face-detection and face-authentication pipelines.

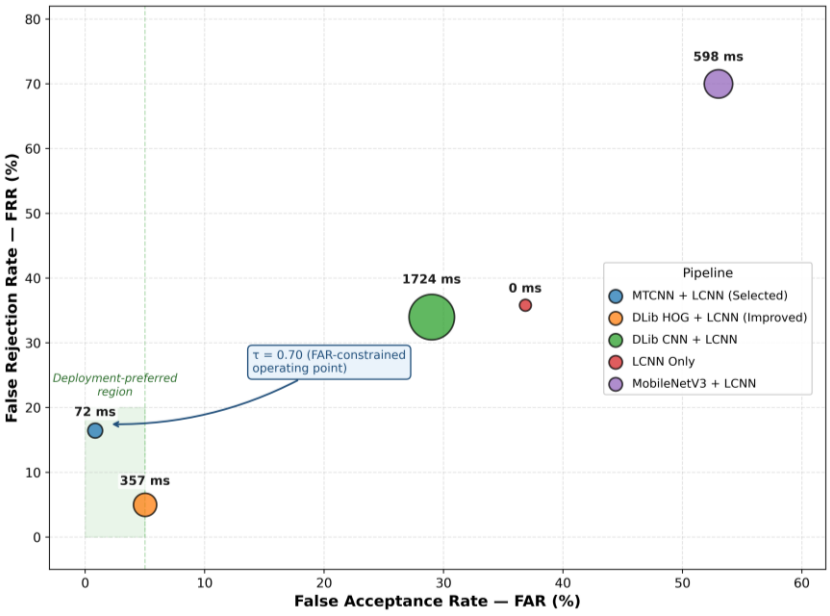

Figure 5. Comparative deployment characteristics of the evaluated pipelines in terms of recognition performance and inference latency.

4.4. Verification Characteristics of the Selected Pipeline

The verification performance of the selected MTCNN + LightCNN-29v2 pipeline is summarized in Table 4, while the distribution of detection and recognition outcomes is illustrated in Fig. 5. Using rank-1 nearest-neighbor retrieval, the selected pipeline achieved 94.95% recognition accuracy with a macro F1-score of 0.95 across all 68 identities. Of the 3,660 testing images, 3,546 were successfully processed, while 114 images (3.1%) were rejected during face detection because no valid face could be localized. Among the processed images, 3,367 (92.0%) were correctly classified and 179 (4.9%) were misclassified during recognition, as shown in Fig. 5. The corresponding ROC-AUC of 0.9970 indicates strong discriminative capability across the complete similarity-score range, while macro precision and recall both reached 0.95, suggesting balanced verification performance across the evaluated identities.

Table 4. Confusion matrix of the selected MTCNN + LightCNN pipeline.

	Predicted Positive	Predicted Negative	Total
Actual Positive (Genuine)	TP = 808	FN = 159	967 pairs
Actual Negative (Impostor)	FP = 8	TN = 961	969 pairs

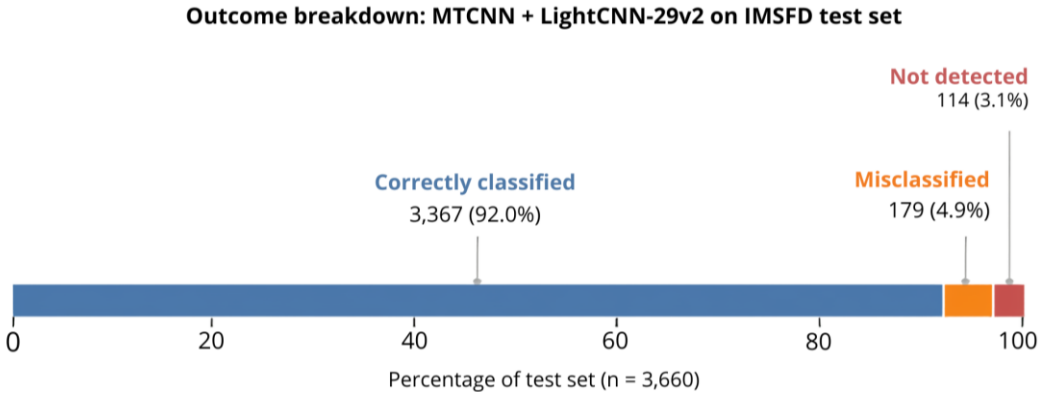

Figure 6. Outcome distribution of the selected MTCNN + LightCNN pipeline on the IMSFD test set (n = 3,660), showing correctly classified images, recognition errors, and detection failures.

4.5. Deployment Feasibility and Overall Result Synthesis

The selected MTCNN + LightCNN-29v2 pipeline requires approximately 5 MB of storage, making it suitable for deployment on entry-level Android devices with limited computational resources. In addition, Mahalanobis-based DBE protects biometric templates before storage, while blockchain-based hash anchoring provides tamper-evident integrity without participating in the verification process. Together, these components demonstrate the feasibility of integrating lightweight face authentication, protected template management, and integrity support within a single mobile financial authentication framework.

The normalized multi-metric comparison presented in Fig. 6 highlights the deployment trade-offs among the five evaluated pipelines. MTCNN + LightCNN provides the most balanced overall profile, combining high recognition performance with low inference latency. DLib HOG + LightCNN achieves comparable recognition performance but requires substantially longer preprocessing, whereas DLib CNN + LightCNN offers strong discriminative capability at the expense of prohibitive latency. In contrast, MobileNetV3 + LightCNN and LightCNN only exhibit substantially weaker deployment characteristics because of limited detection quality and verification reliability.

Overall, the experimental results indicate that end-to-end deployment performance depends primarily on detector quality rather than the recognition backbone alone. Furthermore, integrating protected template matching and integrity support introduces additional security capabilities while preserving the lightweight characteristics required for mobile financial authentication.

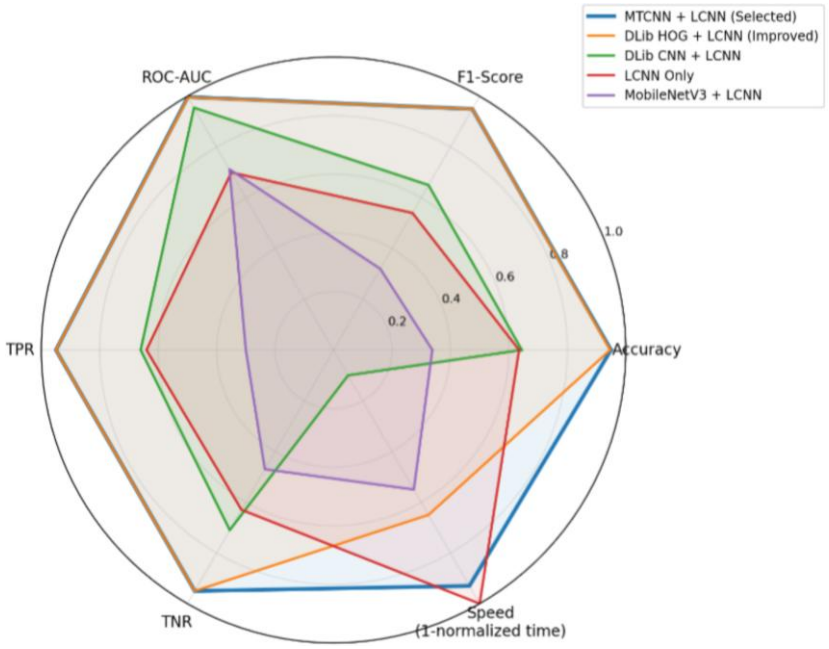

Figure 7. Normalized multi-metric comparison of the five evaluated face-authentication pipelines.

4.6. Cross-Dataset Evaluation

To assess the generalizability of the proposed framework beyond the target deployment domain, the fine-tuned MTCNN + LightCNN-29v2 pipeline was evaluated on six public face-recognition benchmarks using the same cosine similarity threshold (τ=0.70) and evaluation protocol described in Section 3.5. The results are summarized in Table 5.

Table 5. Cross-dataset evaluation results of the fine-tuned MTCNN + LightCNN-29v2 pipeline.

Dataset	Accuracy (%)	ROC-AUC	FAR (%)	FRR (%)	Notes
IMSFD (in-domain)	94.95	0.9970	0.83	16.44	Target deployment domain
LFW	90.32	0.9500	1.15	10.43	Frontal-dominant; demographic mismatch with IMSFD
CFP-FF	87.45	0.9200	2.41	13.12	Frontal–frontal verification
CFP-FP	68.20	0.7600	4.12	28.15	Cross-pose evaluation
AgeDB-30	73.65	0.8100	3.48	24.90	Cross-age evaluation
CALFW	81.12	0.8800	2.95	19.55	Cross-age LFW variant
CPLFW	71.40	0.7900	4.05	26.28	Cross-pose LFW variant

The proposed pipeline achieved the strongest performance on IMSFD, the target deployment dataset, with 94.95% accuracy and a ROC-AUC of 0.9970. Among the public benchmarks, performance remained relatively strong on frontal-dominant datasets such as LFW and CFP-FF, whereas larger performance degradation was observed on cross-pose (CFP-FP, CPLFW) and cross-age (AgeDB-30, CALFW) benchmarks. This trend is consistent with the increasing domain shift between the IMSFD training distribution and the evaluation datasets.

For reference, state-of-the-art models evaluated under the standard benchmark protocols typically report higher recognition performance, including 99.4–99.8% on LFW [32], 99.0–99.5% on CFP-FF [33], 92–98% on CFP-FP [32], [33], 95–98% on AgeDB-30 [34], 95–96% on CALFW [32], and 90–93% on CPLFW [32]. These results, however, are obtained using benchmark-specific protocols such as landmark-based alignment and 10-fold cross-validation. In contrast, the present study employs a single FAR-constrained operating threshold

(τ=0.70) without benchmark-specific optimization to better reflect the intended mobile deployment scenario. Consequently, the comparison is intended to provide performance context rather than establish direct numerical superiority.

4.7. Effect of Mahalanobis-Based DBE on Verification Performance

The effect of Mahalanobis-based Distance-Based Encryption (DBE) on verification performance is summarized in Table 6, which compares the baseline cosine similarity matcher with the proposed protected-template configuration across IMSFD and the six public benchmark datasets. The covariance matrix used for Mahalanobis distance computation was estimated from the IMSFD training embeddings.

Table 6. Verification performance comparison between cosine similarity and Mahalanobis-based DBE across datasets.

Dataset	Baseline Accuracy (%)	Baseline FAR (%)	Baseline FRR (%)	DBE Accuracy (%)	DBE FAR (%)	DBE FRR (%)	Accuracy Change	FAR Change
IMSFD (in-domain)	94.95	0.83	16.44	96.88	0.18	10.12	+1.93	−0.65
LFW	90.32	1.15	10.43	91.15	0.22	11.45	+0.83	−0.93
CFP-FF	87.45	2.41	13.12	88.20	0.45	14.28	+0.75	−1.96
CFP-FP	68.20	4.12	28.15	64.55	0.98	42.60	−3.65	−3.14
AgeDB-30	73.65	3.48	24.90	72.10	0.75	34.15	−1.55	−2.73
CALFW	81.12	2.95	19.55	79.45	0.60	27.80	−1.67	−2.35
CPLFW	71.40	4.05	26.28	67.85	0.88	38.50	−3.55	−3.17

Within the target deployment domain, Mahalanobis-based DBE improved overall verification performance by increasing accuracy from 94.95% to 96.88%, while simultaneously reducing FAR from 0.83% to 0.18% and FRR from 16.44% to 10.12%. Similar reductions in FAR were also observed on frontal-dominant public benchmarks, including LFW and CFP-FF, with only minor changes in overall recognition accuracy.

On datasets exhibiting larger distribution shifts, particularly the cross-pose (CFP-FP, CPLFW) and cross-age (AgeDB-30, CALFW) benchmarks, DBE consistently maintained lower FAR but was accompanied by increased FRR and modest reductions in recognition accuracy. This pattern indicates that the security benefit provided by protected-template matching remains consistent across datasets, whereas the associated usability trade-off becomes increasingly apparent as domain mismatch grows. A more detailed analysis of this behavior is presented in Section 5.1.

5. Discussion

5.1. Key Findings

This study investigated the proposed framework from three complementary perspectives: detector selection, the security–usability trade-off, and generalization under domain shift. Their implications for deployment-oriented mobile financial authentication are discussed below, followed by the positioning of the proposed framework relative to existing studies and its current limitations.

5.1.1. Detector selection as the primary design variable

The controlled comparison demonstrates that face detection is the dominant factor influencing end-to-end authentication performance when the recognition backbone is held constant. Although all evaluated configurations employed the same LightCNN-29v2 feature extractor, substantial performance variation was observed across detector configurations, indicating that localization quality directly affects downstream feature extraction and identity verification. These finding highlights that optimizing the recognition backbone alone is insufficient for deployment-oriented mobile authentication, where robust face localization remains a prerequisite for reliable verification.

Among the evaluated configurations, MTCNN + LightCNN-29v2 provided the most balanced deployment profile by combining high recognition performance, broad detection coverage, and low inference latency. In contrast, MobileNetV3, when repurposed as a sliding-

window detector without task-specific fine-tuning, proved unsuitable for face localization despite its success as a lightweight image-classification architecture. Similarly, DLib CNN produced competitive discriminative performance but incurred latency incompatible with real-time mobile authentication.

An interesting observation is that the performance gap between classical and deep-learning detectors becomes considerably smaller after appropriate image preprocessing. Applying CLAHE and image upsampling enabled the DLib HOG detector to achieve performance approaching that of MTCNN, suggesting that detection failures were influenced more by image quality than by limitations of the handcrafted descriptor itself. This result indicates that carefully designed preprocessing remains a practical strategy for lightweight deployments where computational resources are limited.

Taken together, these findings suggest that detector selection, image preprocessing, and recognition should be considered as an integrated design problem rather than independently optimized components, consistent with previous studies on secure biometric authentication for resource-constrained environments [22], [27], [35], [36].

5.1.2. Security–usability trade-off

The evaluation further demonstrates that financial authentication should be optimized from a risk-aware perspective rather than by recognition accuracy alone. Because false acceptance directly corresponds to unauthorized access, the selected operating threshold prioritizes minimizing FAR while maintaining acceptable verification performance, consistent with the asymmetric security requirements of financial applications [22], [31].

The proposed framework achieved this objective while remaining computationally feasible for mobile deployment. The selected operating point satisfied real-time inference requirements on CPU-only hardware and, when combined with Mahalanobis-based DBE, further reduced the probability of false acceptance within the target deployment domain. At the same time, the experiments illustrate the inherent trade-off between security and usability: stronger protection against impostor access inevitably increases the likelihood of rejecting legitimate users. Consequently, deployment should consider FAR, FRR, latency, and detection coverage jointly rather than relying solely on recognition accuracy.

Beyond recognition performance, the proposed framework also incorporates protected template storage and blockchain-based integrity support. DBE protects biometric templates before storage, reducing the impact of database compromise, while blockchain provides tamper-evident provenance without participating directly in the verification process [10]–[12], [19], [20], [24]. These components complement the authentication pipeline by strengthening operational security while preserving the lightweight characteristics required for mobile financial services.

5.1.3. Generalization under domain shift

Cross-dataset evaluation demonstrates that the proposed framework generalizes well when the target data remains similar to the IMSFD training distribution but gradually degrades as domain shift increases. Performance remains comparatively stable on frontal-dominant benchmarks, whereas larger degradation is observed under cross-pose and cross-age conditions, indicating that the learned representation remains closely coupled to the characteristics of the target deployment domain.

Several factors likely contribute to this behavior. First, IMSFD predominantly contains frontal facial images, limiting the diversity of pose variation available during fine-tuning. Second, the evaluation protocol intentionally omits benchmark-specific landmark alignment to better reflect practical deployment conditions; previous studies have shown that alignment can substantially improve cross-pose verification performance [33]. Third, the relatively narrow demographic and age distribution of IMSFD reduces the model's ability to learn age-invariant facial representations, a challenge that remains widely recognized in face-recognition research [34].

The behavior observed after applying Mahalanobis-based DBE further reinforces this interpretation. Although protected template matching consistently reduced FAR across all evaluated datasets, FRR increased progressively as domain mismatch became larger. This pattern suggests that covariance estimation derived from the target deployment domain remains effective for in-domain verification but becomes increasingly restrictive under distribution

shift. Consequently, extending the training distribution or adopting adaptive threshold calibration may improve cross-domain deployment without sacrificing the security benefits of protected template matching.

5.2. Positioning Relative to Related Studies

The proposed framework is intended to complement, rather than directly compete with, existing face-recognition systems. Because previous studies differ substantially in dataset characteristics, evaluation protocols, deployment objectives, and security scope, direct numerical comparison alone would not provide a meaningful assessment. Instead, the proposed approach is positioned relative to existing work from four complementary perspectives: deployment efficiency, security integration, computational cost, and application context.

Previous financial face-authentication studies have reported recognition accuracies exceeding 99% using large public datasets such as MUCT and AR Face Database [37], [38]. These studies primarily evaluate recognition performance under benchmark-oriented protocols and do not incorporate protected biometric-template storage or blockchain-supported integrity mechanisms. Consequently, their reported performance reflects recognition capability in isolation rather than the behavior of a complete deployment-oriented authentication framework.

From the perspective of lightweight deployment, MobileFaceNets [16] represents an important reference because it combines compact model size with fast inference on mobile devices. Compared with MobileFaceNets, the proposed MTCNN + LightCNN-29v2 pipeline remains within a similar storage footprint while operating under CPU-only inference on resource-constrained hardware. Although inference latency is higher, the proposed framework extends beyond lightweight recognition by incorporating protected template matching, fraud-aware threshold selection, and integrity support within a unified authentication pipeline.

The closest architectural comparison is provided by the blockchain-enabled framework of [30], which combines FaceNet with fully homomorphic encryption (FHE), IPFS, and Ethereum. While FHE provides stronger cryptographic protection, it also introduces substantially greater computational complexity. In contrast, the proposed framework adopts Mahalanobis-based Distance-Based Encryption (DBE) as a lightweight template-protection mechanism that better matches the computational constraints of entry-level mobile devices. The primary contribution therefore lies not in maximizing cryptographic strength, but in achieving a practical balance between biometric protection, computational efficiency, and deployment feasibility. Overall, the proposed framework should be viewed as a deployment-oriented integration strategy rather than a benchmark-oriented recognition model. The public benchmark evaluation presented in Section 4.6 provides evidence of robustness under domain shift, whereas the primary contribution of this work is the integration of lightweight face recognition, protected biometric-template management, and blockchain-supported integrity within a single authentication architecture for digital financial services.

5.3. Limitations and Future Work

Although the proposed framework demonstrates promising performance for deployment-oriented mobile financial authentication, several limitations should be acknowledged. First, the framework was primarily optimized and fine-tuned using the IMSFD dataset, which represents the target deployment domain but remains limited in demographic diversity and population scale. Although additional public benchmarks were included to evaluate robustness under cross-domain conditions, broader validation across larger and more diverse populations is required before the framework can be generalized to large-scale operational environments.

Second, the Mahalanobis-based DBE was evaluated using a covariance matrix estimated from the IMSFD training distribution. As demonstrated by the cross-dataset experiments, this domain-specific statistical model remains effective for in-domain verification but becomes increasingly conservative under distribution shift, resulting in higher False Rejection Rates. Future work should investigate adaptive covariance estimation, domain-adaptive template protection, or dynamic threshold calibration to improve cross-domain robustness while preserving the security advantages of protected template matching.

Third, blockchain integration in this study was evaluated as an architectural proof of concept that demonstrates tamper-evident provenance and secure record management. However, comprehensive evaluation of blockchain scalability, transaction latency, network congestion, and operational cost under production-scale financial systems remains beyond the scope of this work. These aspects warrant further investigation before large-scale deployment.

Finally, the present study focused on evaluating detector selection while maintaining a fixed recognition backbone. Future work may extend this framework by incorporating more recent lightweight recognition models, additional template-protection mechanisms, presentation-attack detection, and larger real-world financial authentication datasets to further improve robustness and deployment readiness.

6. Conclusions

This study investigated whether lightweight face recognition, protected biometric-template matching, and blockchain-based integrity support can be effectively integrated into a practical authentication framework for digital financial services operating on resource-constrained mobile devices. The experimental results demonstrate that this objective can be achieved through a deployment-oriented architecture that jointly considers recognition performance, computational efficiency, template protection, and system integrity.

The comparative evaluation showed that detector selection is the primary factor governing end-to-end authentication performance when the recognition backbone is held constant. Among the five evaluated configurations, MTCNN + LightCNN-29v2 provided the most balanced deployment profile by combining high recognition accuracy, broad detection coverage, low inference latency, and a compact model size suitable for entry-level mobile devices. The results also demonstrate that carefully enhanced classical detectors remain competitive under stricter computational constraints, highlighting the continued importance of efficient preprocessing for lightweight deployment.

The proposed Mahalanobis-based Distance-Based Encryption (DBE) strengthened the authentication framework by reducing the False Acceptance Rate within the target deployment domain while maintaining practical verification performance. Cross-dataset evaluation further showed that the security benefit of protected template matching remained consistent across different benchmarks, although increased domain shift led to higher False Rejection Rates, indicating that deployment-specific threshold calibration remains an important consideration. In addition, blockchain-based hash anchoring was successfully incorporated as a lightweight integrity mechanism that provides tamper-evident provenance without increasing the computational complexity of the verification process.

The findings demonstrate that lightweight face verification, protected biometric-template management, and blockchain-supported integrity can be successfully integrated into a unified authentication framework for digital financial services. By emphasizing deployment feasibility alongside recognition performance and security, the proposed framework provides a practical reference for designing secure mobile biometric authentication systems operating under realistic resource constraints.

Author Contributions: Conceptualization: D.K.P., and F.E.; Methodology: D.K.P., and F.E.; Software: D.K.P.; Validation: D.K.P., J.N., and F.E.; Formal analysis: D.K.P., and F.E.; Investigation: D.K.P., and F.E.; Resources: D.K.P.; Data curation: D.K.P.; Writing—original draft preparation: D.K.P.; Writing—review and editing: D.K.P., J.N. and F.E.; Visualization: D.K.P., and F.E.; Supervision: J.N., and F.E.; Project administration: D.K.P.; Funding acquisition: D.K.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The Indonesian Muslim Student Face Dataset (IMSFD) used in this study is a third-party dataset publicly available via Mendeley Data at doi: https://doi.org/10.17632/f6f3y6ndgw.1, released under terms permitting academic use. The public benchmark datasets used for cross-dataset evaluation (LFW, CFP-FF, CFP-FP, AgeDB-30, CALFW, and CPLFW) are also openly available from their respective repositories, as cited in the References section.

Acknowledgments: The authors acknowledge the use of generative AI tools to improve the clarity and grammar of the manuscript. All technical content was written and verified by the authors.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

[1] M. R. Prasad and N. Thillaiarasu, "A Framework for Integrated Bank ATM Transactions Security Based on Multi-Modal Biometric Authentication Using Adaptive and Attentive Assisted Mobilenet-V3," *Comput. Intell.*, vol. 41, no. 4, Aug. 2025, doi: 10.1111/coin.70083.

[2] J. Chi, C. Kim On, H. Zhang, and S. S. Chai, "A Review of Deep Convolutional Neural Networks in Mobile Face Recognition," *Int. J. Interact. Mob. Technol.*, vol. 17, no. 23, pp. 4–19, Dec. 2023, doi: 10.3991/ijim.v17i23.40867.

[3] A. Baran and E. Bartuzi-Trokielewicz, "Face the Challenge—Generalization of Presentation Attack Detection," *Sensors*, vol. 25, no. 18, p. 5792, Sep. 2025, doi: 10.3390/s25185792.

[4] A. Hassani, J. Diedrich, and H. Malik, "Improving Monocular Facial Presentation–Attack–Detection Robustness with Synthetic Noise Augmentations," *Sensors*, vol. 23, no. 21, p. 8914, Nov. 2023, doi: 10.3390/s23218914.

[5] L. R. Zuama, D. R. I. M. Setiadi, A. Susanto, S. Santosa, H.-S. Gan, and A. A. Ojugo, "High-Performance Face Spoofing Detection using Feature Fusion of FaceNet and Tuned DenseNet201," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 4, pp. 385–400, Feb. 2025, doi: 10.62411/faith.3048-3719-62.

[6] A. Howard *et al.*, "Searching for MobileNetV3," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2019, pp. 1314–1324. doi: 10.1109/ICCV.2019.00140.

[7] X. Wu, R. He, Z. Sun, and T. Tan, "A Light CNN for Deep Face Representation With Noisy Labels," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 11, pp. 2884–2896, Nov. 2018, doi: 10.1109/TIFS.2018.2833032.

[8] M. A. Aleisa, "Enhancing Security in CPS Industry 5.0 using Lightweight MobileNetV3 with Adaptive Optimization Technique," *Sci. Rep.*, vol. 15, no. 1, p. 18677, May 2025, doi: 10.1038/s41598-025-00496-3.

[9] X. Liang, J. Liang, T. Yin, and X. Tang, "A lightweight method for face expression recognition based on improved MobileNetV3," *IET Image Process.*, vol. 17, no. 8, pp. 2375–2384, Jun. 2023, doi: 10.1049/ipr2.12798.

[10] F. Guo, W. Susilo, and Y. Mu, "Distance-Based Encryption: How to Embed Fuzziness in Biometric-Based Encryption," *IEEE Trans. Inf. Forensics Secur.*, vol. 11, no. 2, pp. 247–257, Feb. 2016, doi: 10.1109/TIFS.2015.2489179.

[11] L. Laishram, M. Shaheryar, J. T. Lee, and S. K. Jung, "Toward a Privacy-Preserving Face Recognition System: A Survey of Leakages and Solutions," *ACM Comput. Surv.*, vol. 57, no. 6, pp. 1–38, Jun. 2025, doi: 10.1145/3673224.

[12] Q. Ren, C. Yuan, Y. Zhao, and L. Yang, "A novel metric learning framework by exploiting global and local information," *Neurocomputing*, vol. 507, pp. 84–96, Oct. 2022, doi: 10.1016/j.neucom.2022.08.003.

[13] V. Bazarevsky, Y. Kartynnik, A. Vakunov, K. Raveendran, and M. Grundmann, "BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs," *ArXiv*. Jul. 14, 2019.

[14] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 5202–5211. doi: 10.1109/CVPR42600.2020.00525.

[15] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016, doi: 10.1109/LSP.2016.2603342.

[16] S. Chen, Y. Liu, X. Gao, and Z. Han, "MobileFaceNets: Efficient CNNs for Accurate Real-Time Face Verification on Mobile Devices," 2018, pp. 428–438. doi: 10.1007/978-3-319-97909-0_46.

[17] Y. Martindez-Diaz, L. S. Luevano, H. Mendez-Vazquez, M. Nicolas-Diaz, L. Chang, and M. Gonzalez-Mendoza, "ShuffleFaceNet: A Lightweight Face Architecture for Efficient and Highly-Accurate Face Recognition," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, Oct. 2019, pp. 2721–2728. doi: 10.1109/ICCVW.2019.00333.

[18] F. Olweny, "Navigating the nexus of security and privacy in modern financial technologies," *GSC Adv. Res. Rev.*, vol. 18, no. 2, pp. 167–197, Feb. 2024, doi: 10.30574/gscarr.2024.18.2.0043.

[19] D. Aguilar, A. Martínez-Cruz, K. A. Ramírez-Gutiérrez, and M. Morales-Sandoval, "PPG-based biometric authentication: A review on architectures, datasets, attacks and security challenges," *Comput. Struct. Biotechnol. J.*, vol. 28, pp. 511–528, Jan. 2025, doi: 10.1016/j.csbj.2025.11.006.

[20] W. Yang, S. Wang, N. M. Sahri, N. M. Karie, M. Ahmed, and C. Valli, "Biometrics for Internet-of-Things Security: A Review," *Sensors*, vol. 21, no. 18, p. 6163, Sep. 2021, doi: 10.3390/s21186163.

[21] V. Schlatt, J. Sedlmeir, S. Feulner, and N. Urbach, "Designing a Framework for Digital KYC Processes Built on Blockchain-Based Self-Sovereign Identity," *Inf. Manag.*, vol. 59, no. 7, p. 103553, Nov. 2022, doi: 10.1016/j.im.2021.103553.

[22] S. S. U. Hasan, A. Ghani, A. Daud, H. Akbar, and M. F. Khan, "A Review on Secure Authentication Mechanisms for Mobile Security," *Sensors*, vol. 25, no. 3, p. 700, Jan. 2025, doi: 10.3390/s25030700.

[23] Q. N. Tran, B. P. Turnbull, H.-T. Wu, A. J. S. de Silva, K. Kormusheva, and J. Hu, "A Survey on Privacy-Preserving Blockchain Systems (PPBS) and a Novel PPBS-Based Framework for Smart Agriculture," *IEEE Open J. Comput. Soc.*, vol. 2, pp. 72–84, 2021, doi: 10.1109/OJCS.2021.3053032.

[24] S. Sharma and R. Dwivedi, "A survey on blockchain deployment for biometric systems," *IET Blockchain*, vol. 4, no. 2, pp. 124–151, Jun. 2024, doi: 10.1049/blc2.12063.

[25] O. Umoren, R. Singh, Z. Pervez, and K. Dahal, "Securing Fog Computing with a Decentralised User Authentication Approach Based on Blockchain," *Sensors*, vol. 22, no. 10, p. 3956, May 2022, doi: 10.3390/s22103956.
[26] A. Alharthi, Q. Ni, and R. Jiang, "A Privacy-Preservation Framework Based on Biometrics Blockchain (BBC) to Prevent Attacks in VANET," *IEEE Access*, vol. 9, pp. 87299–87309, 2021, doi: 10.1109/ACCESS.2021.3086225.
[27] N. Zeeshan, M. Bakyt, N. Moradpoor, and L. La Spada, "Continuous Authentication in Resource-Constrained Devices via Biometric and Environmental Fusion," *Sensors*, vol. 25, no. 18, p. 5711, Sep. 2025, doi: 10.3390/s25185711.
[28] P. Grother, M. Ngan, and K. Hanaoka, "Face recognition vendor test part 3: Demographic effects," Gaithersburg, MD, Dec. 2019. doi: 10.6028/NIST.IR.8280.
[29] J. Buolamwini and T. Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 2018, vol. 81, pp. 77–91. [Online]. Available: https://proceedings.mlr.press/v81/buolamwini18a.html
[30] S. Vadim, M. Firdaus, and K.-H. Rhee, "Privacy-Preserving Decentralized Biometric Identity Verification in Car-Sharing System," *J. Multimed. Inf. Syst.*, vol. 11, no. 1, pp. 17–34, Mar. 2024, doi: 10.33851/JMIS.2024.11.1.17.
[31] X. Zhu and C. Cao, "Secure Online Examination with Biometric Authentication and Blockchain-Based Framework," *Math. Probl. Eng.*, vol. 2021, pp. 1–12, Aug. 2021, doi: 10.1155/2021/5058780.
[32] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 4685–4694. doi: 10.1109/CVPR.2019.00482.
[33] S. Sengupta, J.-C. Chen, C. Castillo, V. M. Patel, R. Chellappa, and D. W. Jacobs, "Frontal to profile face verification in the wild," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Mar. 2016, pp. 1–9. doi: 10.1109/WACV.2016.7477558.
[34] S. Moschoglou, A. Papaioannou, C. Sagonas, J. Deng, I. Kotsia, and S. Zafeiriou, "AgeDB: The First Manually Collected, In-the-Wild Age Database," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jul. 2017, pp. 1997–2005. doi: 10.1109/CVPRW.2017.250.
[35] A. Kanak *et al.*, "Integrated secure distance bounding and hardware-based security: A case study for the insurance claim verification of farmers during COVID-19," *Open Res. Eur.*, vol. 3, p. 40, Feb. 2023, doi: 10.12688/openreseurope.15448.1.
[36] M. Y. Hong, K. H. Lee, J. H. Jung, and J. W. Yoon, "ToMAS: Torus-based secure multi-factor biometric authentication system," *Comput. Struct. Biotechnol. J.*, vol. 28, pp. 529–537, Jan. 2025, doi: 10.1016/j.csbj.2025.10.042.
[37] K. Velayuthapandian, N. Murugan, and S. Paramasivan, "End-to-End CNN conceptual model for a biometric authentication mechanism for ATM machines," *Discov. Electron.*, vol. 1, no. 1, p. 26, Nov. 2024, doi: 10.1007/s44291-024-00034-x.
[38] L. Nosrati, A. M. Bidgoli, and H. H. S. Javadi, "Identifying People's Faces in Smart Banking Systems Using Artificial Neural Networks," *Int. J. Comput. Intell. Syst.*, vol. 17, no. 1, p. 9, Jan. 2024, doi: 10.1007/s44196-023-00383-7.
[39] Purnawansyah, A. P. Wibawa, T. Widyaningtyas, Haviluddin, H. Darwis, H. Azis, and Z. Ali, "Indonesian Muslim Student Face Dataset (IMSFD)," *Mendeley Data*, V1, 2023, doi: 10.17632/f6f3y6ndgw.1.