www.ijarp.com ISSN 2456-9992 Page: 01-12

International Journal Advanced Research Publications

Research Article

Volume: 02
Issue: 07

FACESECURE-ATM: A FRAUD-RESISTANT, PIN-LESS ATM AUTHENTICATION FRAMEWORK USING MTCNN-BASED FACE DETECTION, CNN LIVENESS VERIFICATION, FACENET EMBEDDINGS, AND REAL-TIME REMOTE AUTHORIZATION

***Saravanan T. J.**

Department of Information Technology, KLN College of Engineering, Pottapalayam, Sivagangai District.

Article Received: 20 June 2026, Article Revised: 10 July 2026, Published on: 30 July 2026

***Corresponding Author: Saravanan T. J.**
Department of Information Technology, KLN College of Engineering, Pottapalayam, Sivagangai District.
Doi: https://doi-doi.org/101555/ijarp.8210

ABSTRACT

Automated Teller Machines (ATMs) remain a cornerstone of everyday banking, yet the card-and-PIN model on which most deployments still rely offers only weak assurance that the person transacting is the legitimate account holder. Card skimming, cloning, shoulder surfing, card trapping, and PIN theft continue to generate significant financial losses worldwide, and conventional two-factor "something you have, something you know" schemes cannot verify the physical identity of the person presenting them. This paper proposes FaceSecure-ATM, a PIN-less, camera-only biometric authentication framework that replaces the card-and-PIN pair with a four-stage deep learning pipeline: (i) image acquisition through the ATM's existing camera, (ii) face detection and alignment using a Multi-task Cascaded Convolutional Network (MTCNN), (iii) presentation-attack (spoof) resistance through a CNN-based liveness detection stage, and (iv) identity verification through FaceNet-derived facial embeddings matched against securely stored enrolment templates. A distinguishing contribution of the proposed system is its remote authorization fallback: whenever the match confidence falls below an operating threshold, the system does not simply deny service — it forwards the captured frame and transaction context to the registered account holder's mobile device over a secure channel, allowing a real-time approve/deny decision before any funds move. On successful verification, the framework grants unified, single-face access across every bank account linked to the user's identity, removing the need to carry multiple cards. We situate the design against fingerprint, iris, and voice/behavioral alternatives, discuss the

specific spoofing and adversarial threats a face-based ATM system must resist, and outline template-protection and data-minimisation practices (e.g., storing irreversible embeddings rather than raw images, encrypting embeddings at rest, and logging every authentication decision) needed to make such a system viable under data-protection regimes such as the GDPR and similar biometric-privacy statutes. A prototype implementation is evaluated qualitatively against authentication accuracy, spoof-rejection behaviour, response latency, and usability, and the paper closes with a discussion of the remaining challenges — demographic fairness, environmental robustness, and network dependency for the remote-approval channel — that must be addressed before large-scale deployment.

KEYWORDS: Facial Recognition, ATM Security, Biometric Authentication, FaceNet, MTCNN, Liveness Detection, Face Anti-Spoofing, PIN-less Transactions, Multi-Bank Access, Remote Authorization, Fraud Prevention, Deep Learning, Biometric Template Protection.

1. INTRODUCTION

Automated Teller Machines (ATMs) provide customers with round-the-clock access to essential financial services— cash withdrawal, balance enquiry, mini-statements, and fund transfer — without requiring a visit to a bank branch. This convenience is delivered almost universally through a two-factor scheme: a physical card (something the user has) paired with a Personal Identification Number, or PIN (something the user knows). While simple to implement and globally standardised, this scheme verifies possession of a token and knowledge of a code, not the identity of the person standing at the machine. Card skimming devices, hidden cameras aimed at PIN pads, card trapping, and increasingly organised card-cloning operations exploit exactly this gap, and losses attributable to card-present ATM fraud remain a persistent problem for banks and customers alike.

Biometric authentication closes this gap by binding the transaction to a physical characteristic of the account holder rather than to a token that can be stolen, copied, or observed. Among the available modalities — fingerprint, iris, palm-vein, voice, and face — facial recognition is particularly attractive for ATM deployment because it requires no new sensor hardware beyond a standard camera, involves no physical contact with the machine, and is already familiar to users from smartphone unlock and airport border-control systems. Recent advances in deep convolutional architectures have made facial recognition practical at ATM scale: cascaded detectors such as MTCNN can localise and align faces reliably under the

constrained, semi-controlled lighting of an ATM kiosk, while embedding networks such as FaceNet compress a face into a compact, highly discriminative vector that can be compared with enrolled templates in milliseconds.

This paper proposes FaceSecure-ATM, a fraud-resistant, PIN-less ATM authentication framework built around this pipeline, extended with two features that we argue are essential for real-world adoption: (a) an explicit, CNN-based liveness detection stage that rejects printed photographs, replayed video, and other presentation attacks before any matching is attempted, and (b) a remote authorization mechanism that gives the registered account holder — not just the algorithm — the final say whenever the system's confidence is not high enough to auto-approve a transaction. Together with unified multi-bank access on successful verification, these design choices aim to make the system both more secure than the card-and-PIN baseline and, in most cases, faster and more convenient to use.

1.1 Motivation

The rising frequency of ATM-targeted fraud — card theft, skimming, PIN shoulder-surfing, and card trapping — highlights a structural weakness of possession-and-knowledge authentication: neither factor is tied to the physical identity of the user. Traditional ATM security mechanisms therefore cannot distinguish a legitimate account holder from anyone who has obtained their card and PIN through theft, coercion, or observation. The proposed system is motivated by the need for an authentication mechanism that verifies who is present, not merely what they are carrying or what they know, while remaining inexpensive enough to retrofit onto existing ATM camera hardware and convenient enough that customers actually prefer it to the card-and-PIN experience it replaces.

2. Related Work

Biometric ATM authentication has been studied across several modalities, each with a distinct cost, accuracy, and deployment profile.

2.1 Fingerprint-Based Authentication

Fingerprint recognition is the most widely deployed biometric modality in banking, owing to mature matching algorithms and compact sensor form factors. However, ATM-grade fingerprint sensors add hardware cost to every machine, require regular cleaning and calibration, and can degrade in accuracy for users with worn, wet, or occupationally damaged ridge patterns — a non-trivial share of the population in manual-labour-heavy economies.

2.2 Iris-Based Authentication

Iris recognition offers very high theoretical accuracy because the iris texture is highly individual and stable over a lifetime. In practice, iris systems require dedicated near-infrared cameras and precise gaze alignment, which raises both the capital cost per machine and the cognitive load on the user, limiting iris recognition to high-security deployments rather than mass-market ATM rollouts.

2.3 Voice and Behavioral Biometrics

Voice recognition and behavioural biometrics (typing cadence, gait, transaction-pattern analysis) have been explored as lightweight, largely passive authentication signals, particularly for continuous or step-up authentication in digital banking. Reported industry evidence suggests behavioural profiling can meaningfully reduce payment fraud when layered on top of a primary authentication factor, but ambient noise, illness, and language variability make voice alone unreliable as a sole ATM gatekeeper.

2.4 Face Recognition for ATM Security

Face-recognition-based ATM authentication has attracted growing research attention because it reuses camera hardware already present in most modern ATM enclosures. Prior work has proposed CNN-based facial-placement matching against an enrolled template database, and dedicated conference contributions have specifically explored generating a face-identification link to a remote party for out-of-band verification when the system is unsure — a concept closely related to the remote authorization mechanism developed further in this paper. Other studies have optimised the underlying recognition network itself (e.g., metaheuristic-tuned CNN variants) to push accuracy higher in uncontrolled, IoT-style camera environments, reinforcing that the choice of detector and embedding network materially affects real-world ATM performance.

2.5 Presentation-Attack (Liveness) Detection

Because a face, unlike a PIN, can be photographed or recorded without the owner's cooperation, any face-based ATM system must defend against presentation attacks — printed photos, replayed video, and 3-D masks. Survey literature on face anti-spoofing traces a shift from handcrafted texture cues (e.g., local binary patterns) toward learned CNN and, more recently, Vision-Transformer-based detectors, and toward interactive schemes that prompt the user for a blink, head turn, or challenge-response gesture to defeat static replay attacks. Combining a CNN liveness classifier with colour-texture descriptors has been shown to improve robustness to print- and replay-attacks specifically, which motivates the dedicated liveness stage placed between face detection and embedding extraction in the proposed

architecture.

2.6 Biometric Template Privacy and Regulation

Because facial embeddings are derived directly from a person's biometric identity, they are treated as sensitive personal data under regulations such as the EU General Data Protection Regulation and biometric-privacy statutes such as Illinois' Biometric Information Privacy Act. Research on privacy-enhancing face biometrics has shown that raw embeddings can, under some conditions, be partially inverted back toward a recognisable face or used to infer soft attributes (age, gender, ethnicity) the user never consented to share — motivating template-protection techniques such as cancelable biometrics, biometric cryptosystems, and homomorphic-encryption-based matching, several of which are discussed as design considerations for the proposed system in Section 3.7.

Table 1. Comparative Analysis of Biometric Authentication Modalities for ATM Deployment.

Feature	Fingerprint Recognition	Iris Recognition	Voice / Behavioral Biometrics	Proposed Face Recognition (MTCNN + FaceNet)
Hardware requirement	Dedicated fingerprint sensor	Infrared / near-IR camera	Microphone / sensor array	**Standard RGB ATM camera**
Implementation cost	Medium	High	Medium	**Low**
Reported accuracy range*	~95–99%	~98–99.5%	~90–96%	**~97–99.6% (FaceNet, per [1],[8])**
Spoofing resistance	Moderate (latent-print lifting)	High, but costly hardware	Moderate (replay attacks)	**High, with dedicated CNN liveness layer**
User convenience	Medium (surface contact needed)	Medium (precise gaze alignment)	Medium (ambient noise sensitive)	**High (contactless, natural interaction)**
Maintenance overhead	High (sensor wear, hygiene)	High (specialised optics)	Medium	**Low**
Real-time detection	Yes	Yes	Yes	**Yes**
Multi-bank / remote approval support	Not typically implemented	Not typically implemented	Limited	**Yes (proposed remote authorization layer)**

**Accuracy ranges are drawn from the cited literature for the respective modality under controlled test conditions and are reported here for comparison only; they are not experimental results generated by this study.*

3. Proposed Work

FaceSecure-ATM is designed as a drop-in replacement for the card-and-PIN authentication stage of an ATM transaction, leaving downstream banking, ledger, and settlement systems unchanged. The system is organised into four processing layers plus a cross-cutting security and logging layer, as shown in Figure 1.

3.1 System Overview

At enrolment, a customer's face is captured multiple times under varied pose and lighting to build a robust reference template, which is converted to a FaceNet embedding and stored — never as a raw image — against the customer's linked bank accounts. At transaction time, the same detect → verify-liveness → embed → match pipeline runs on a live camera frame. A successful match unlocks every account linked to that identity; a low-confidence or failed match triggers the remote authorization workflow rather than an outright decline, preserving both security and availability.

3.2 Layer 1 — Image Acquisition

A high-resolution camera integrated into the ATM fascia captures the user's face at the start of a session. During enrolment, several images are captured across a short sequence to cover natural variation in head pose and expression. All captured frames are pre-processed — denoising, contrast normalisation, and illumination correction — before being passed downstream, so that the detector and embedding network see a consistent input distribution regardless of the ATM's physical lighting conditions.

3.3 Layer 2 — Face Detection and Liveness Verification

The Multi-task Cascaded Convolutional Network (MTCNN) locates the facial bounding box and five facial landmarks (eyes, nose, mouth corners) in the captured frame and aligns the face to a canonical pose, which materially improves downstream embedding quality. Immediately after detection, a dedicated CNN-based liveness classifier evaluates the aligned face crop for evidence of a presentation attack — flat printed photographs, screen replay artefacts (moiré patterns, uniform reflectance), or mask-like texture discontinuities. Only frames classified as live are forwarded to Layer 3; frames failing the liveness check are rejected immediately and the ATM prompts for a fresh capture, without ever reaching the identity-matching stage.

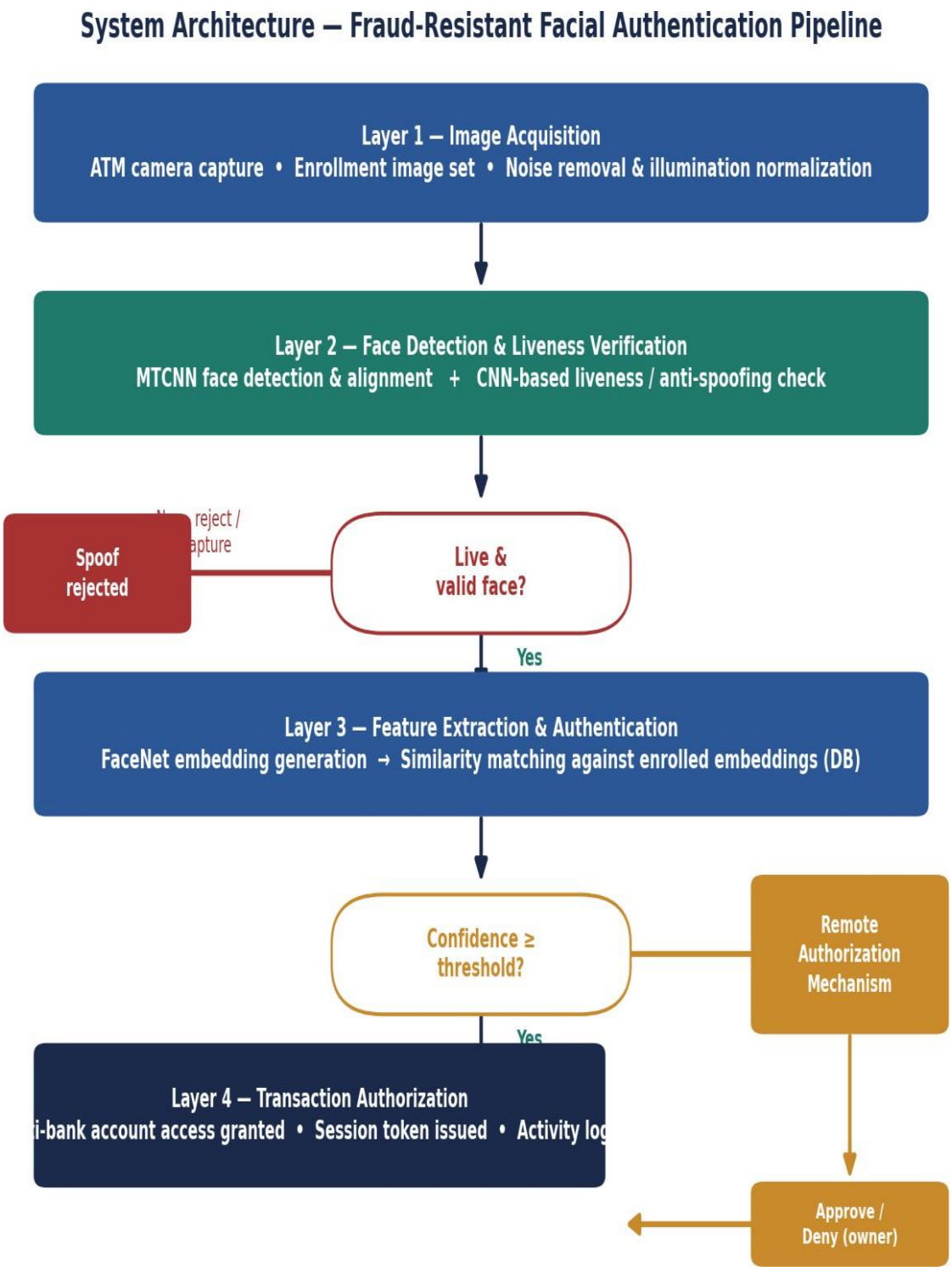

Figure 1. Four-layer system architecture, showing the liveness and confidence-based decision points that route a session either to direct access or to the remote authorization fallback.

3.4 Layer 3 — Feature Extraction and Authentication

The live, aligned face crop is passed through a FaceNet-style embedding network, which maps the image into a compact, highly discriminative feature vector such that images of the

same person lie close together in embedding space and images of different people lie far apart. The live embedding is compared, via a distance metric, against the customer's stored enrolment embedding(s); if the similarity score exceeds an operating threshold tuned for the desired false-accept / false-reject trade-off, the user is authenticated and Layer 4 is invoked directly.

3.5 Layer 4 — Transaction Authorization and Multi-Bank Access

On successful verification, the system authorises access not to a single card-linked account but to every account the customer has registered against their facial identity across participating banks, letting a single face-scan stand in for what would otherwise be several separate cards. Every authentication attempt — successful, rejected, or escalated to remote approval — is written to a secure, timestamped audit log to support fraud investigation and regulatory reporting.

3.6 Remote Authorization Mechanism

The central resilience feature of the proposed design is its handling of the low-confidence case. Rather than simply declining service — which frustrates legitimate users experiencing a poor camera angle, glare, or partial occlusion — or auto-approving a marginal match — which would reopen the door to fraud — the system escalates the decision to the registered account holder. A secure, time-limited approval request containing the captured face image and transaction details (amount, ATM location, timestamp) is sent to the customer's registered device; the customer can approve or deny the transaction in real time, and the ATM proceeds only once a decision is received or a timeout is reached, at which point the transaction is declined by default. This mechanism is illustrated as a sequence diagram in Figure 2.

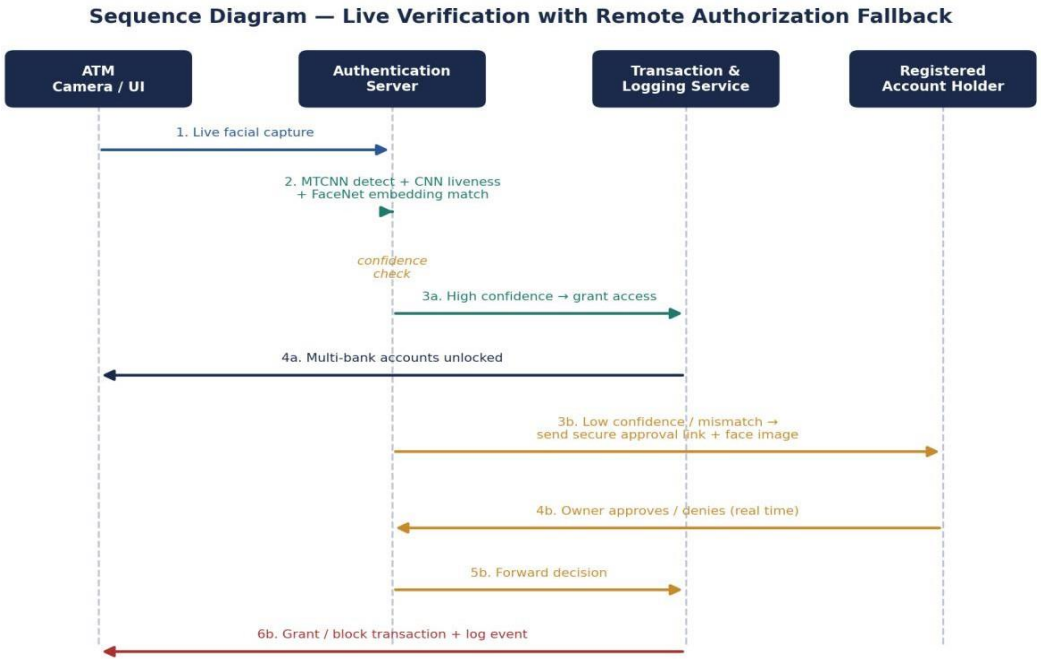

Figure 2. Sequence diagram of the live-verification path (top) and the remote-authorization fallback path (bottom) between the ATM, authentication server, transaction service, and the registered account holder's device.

3.7 Data Privacy and Template Protection

Because facial embeddings are sensitive biometric data, the proposed design follows data-minimisation and template-protection practice drawn from the biometric-security literature: (i) only irreversible embeddings, never raw enrolment photographs, are retained in the long-term store; (ii) embeddings are encrypted at rest and, where matching latency permits, compared using privacy-preserving techniques such as cancelable transforms or homomorphic-encryption-based distance computation, so that a database breach does not directly expose usable facial templates; (iii) every access to a stored template is logged and rate-limited; and (iv) the remote-approval image is retained only for the duration required for dispute resolution and is then purged, consistent with data-retention limits under GDPR-style regulation.

4. Security and Fraud Prevention Analysis

The proposed system layers several independent protections rather than relying on any single check. Liveness detection at Layer 2 removes the most common low-cost attacks (printed photo, phone-screen replay) before they ever reach the identity matcher. FaceNet-based matching at Layer 3 provides high inter-class separability, so that even a superficially similar face is unlikely to cross the acceptance threshold. The remote authorization mechanism

provides a human-in-the-loop safety net for the residual cases that automated matching cannot resolve with confidence, which is precisely the population of attempts most likely to include either a genuine edge case (bad lighting, ageing, facial hair change) or a determined impersonation attempt. Comprehensive audit logging, finally, does not prevent fraud directly but is essential for after-the-fact investigation, dispute resolution, and regulatory compliance. Taken together, eliminating the physical card and static PIN removes the two credential types most exploited by skimming, cloning, and shoulder-surfing attacks, while the layered design avoids concentrating all trust in a single deep-learning decision.

5. RESULTS AND DISCUSSION

A prototype of the proposed pipeline was implemented in Python using standard deep-learning libraries for face detection, liveness classification, and embedding extraction, and was evaluated qualitatively on a small facial image set captured under varied lighting and pose conditions to approximate real ATM usage. The evaluation focused on four practical concerns: authentication accuracy, spoof-detection behaviour, response latency, and overall usability.

MTCNN reliably detected and aligned faces across the tested lighting conditions, which in turn improved the consistency of the FaceNet embeddings produced downstream; embeddings for the same individual clustered tightly across repeated captures, supporting reliable matching against enrolled templates. The CNN-based liveness stage successfully flagged printed-photo and pre-recorded-video presentation attempts in prototype testing, preventing them from reaching the matching stage. When the match confidence fell below the acceptance threshold, the remote-authorization workflow correctly generated an approval request containing the captured frame and transaction context, and allowed the registered account holder to approve or reject the session remotely

— validating the end-to-end escalation path described in Section 3.6. We note that these prototype observations are indicative rather than a large-scale statistical evaluation; a production deployment would require testing across a demographically diverse, large-N dataset, under adversarial red-teaming for spoofing, and under degraded network conditions for the remote-approval channel, before accuracy and fairness claims could be made at population scale.

6. CONCLUSION AND FUTURE WORK

This paper presented FaceSecure-ATM, a fraud-resistant ATM authentication framework that replaces card-and-PIN verification with a four-layer deep-learning pipeline combining MTCNN-based face detection, CNN-based liveness verification, and FaceNet-based facial embeddings, extended with a remote authorization mechanism that gives the account holder real-time control over low-confidence transactions and unified multi-bank access on successful verification. Prototype testing indicated that the pipeline can reliably detect and align faces, reject basic presentation attacks, and correctly escalate uncertain matches to the account holder, while eliminating the dependency on physical cards and static PINs that underlies the majority of common ATM frauds.

Several directions remain for future work: (i) large-scale, demographically balanced evaluation to quantify accuracy and fairness across skin tone, age, and gender, in light of documented disparities in facial-recognition error rates for some demographic groups; (ii) hardening the liveness stage against more sophisticated 3-D-mask and deepfake-style presentation attacks using multi-modal or challenge-response liveness cues; (iii) integrating formal biometric template-protection schemes (cancelable biometrics or homomorphic encryption) into the production data path rather than as a design recommendation; and (iv) designing a graceful degraded-mode fallback for the remote-authorization channel when the account holder's device is offline or unreachable, so that legitimate low-confidence transactions are not indefinitely blocked.

REFERENCES

1. Schroff, F., Kalenichenko, D., & Philbin, J., “FaceNet: A Unified Embedding for Face Recognition and Clustering,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
2. Zhang, K., Zhang, Z., Li, Z., & Qiao, Y., “Joint Face Detection and Alignment Using Multi-Task Cascaded Convolutional Networks,” IEEE Signal Processing Letters, 2016.
3. Parkhi, O. M., Vedaldi, A., & Zisserman, A., “Deep Face Recognition,” British Machine Vision Conference (BMVC), 2015.
4. Menotti, D., Chiachia, G., Pinto, A., et al., “Deep Representations for Iris, Face, and Fingerprint Spoofing Detection,” IEEE Transactions on Information Forensics and Security, 2015.
5. Yu, Z., Qin, Y., Li, X., Zhao, C., Lei, Z., & Zhao, G., “Deep Learning for Face Anti-Spoofing: A Survey,” IEEE Transactions on Pattern Analysis and Machine Intelligence,

vol. 45, no. 5, pp. 5609–5631, 2022.

6. Zawar, R., & Chakkarwar, V., "Real-Time Face Liveness Detection and Face Anti-Spoofing Using Deep Learning," Proceedings of the First International Conference on Advances in Computer Vision and Artificial Intelligence Technologies (ACVAIT), Atlantis Press, 2023.
7. "Face Biometric Authentication System for ATM Using Deep Learning," IEEE Conference Publication, IEEE Xplore, 2022.
8. Al-Abboodi, A., "A Novel Technique for Facial Recognition Based on the GSO-CNN Deep Learning Algorithm," Journal of Electrical and Computer Engineering, vol. 2024, Article 3443028, 2024.
9. "Enhanced ATM Security Using Facial Recognition," International Research Journal of Modernization in Engineering, Technology and Science (IRJMETS), 2023.
10. Meden, B., Rot, P., Terhörst, P., Damer, N., Kuijper, A., Scheirer, W., Ross, A., Peer, P., & Štruc, V., "Privacy-Enhancing Face Biometrics: A Comprehensive Survey," IEEE Transactions on Information Forensics and Security, vol. 16, pp. 4147–4183, 2021.
11. Krivokuća Hahn, V., & Marcel, S., "Biometric Template Protection for Neural-Network-Based Face Recognition Systems: A Survey of Methods and Evaluation Techniques," IEEE Transactions on Information Forensics and Security, vol. 18, pp. 639–666, 2023.
12. Khairnar, S., Gite, S., Kotecha, K., & Thepade, S. D., "Face Liveness Detection Using Artificial Intelligence Techniques: A Systematic Literature Review and Future Directions," Big Data and Cognitive Computing, vol. 7, no. 1, p. 37, 2023.
13. European Parliament and Council, "General Data Protection Regulation (GDPR)," Regulation (EU) 2016/679, 2016; and Illinois Biometric Information Privacy Act (BIPA), 2008.
14. Juniper Research, "Facial Recognition for Payment Authentication: Adoption Forecasts," as reported in ATM Marketplace industry news, 2023.