Deep Learning-Based Biometric Payment System with Dual-Tap Fingerprint Authentication

Vidya Sagar S D [1,2], Ajay Kumar [2,3]

[1] Department of MCA, Nitte Meenakshi Institute of Technology, Bengaluru, India,

[2] Lincoln University College, Malaysia .

[3] School of CSE, IILM University Greater Noida, India

pdf.vidyasagarsd@lincoln.edu.my, pdfsv.ajaykumar@lincoln.edu.my / *vidya.sagar@nmit.ac.in*

ABSTRACT

Abstract: Mobile payment fraud and identity theft pose critical threats to the rapidly expanding digital financial ecosystem, projected to exceed $15 trillion in transaction volume by 2025. Traditional authentication methods including passwords and one-time passwords (OTPs) remain vulnerable to phishing, credential stuffing, and replay attacks. This paper proposes a novel three-layer biometric payment framework integrating: (1) a dual deep learning architecture employing a CNN-based liveness detection model and a Siamese/Triplet network for fingerprint embedding, (2) a Dual-Tap Fingerprint Authentication (DTFA) protocol that introduces temporal behavioral biometrics through configurable two-stage fingerprint capture, and (3) a Dynamic One-Time Transaction (DOT) Code layer using SHA-256 cryptographic binding. Evaluated on a custom dataset of 5,000 participants encompassing 50,000 genuine fingerprint samples and 15,000 presentation attacks, the proposed DTFA-DOT system achieves 99.47% authentication accuracy with an Equal Error Rate (EER) of 0.085%, a False Acceptance Rate (FAR) of 0.02%, and 0% success rate against replay, man-in-the-middle, and transaction manipulation attacks. User acceptance studies with 500 participants yielded a 4.6/5.0 satisfaction rating and 8.2-second average transaction time. These results demonstrate a new standard for secure, usable mobile banking authentication.

Keywords: Biometric Authentication; Deep Learning; Fingerprint Recognition; Mobile Payment Security; Dual-Tap Authentication; Liveness Detection; Siamese Network; Cryptographic Token; Behavioral Biometrics; Spoof Detection.

1. INTRODUCTION

The proliferation of mobile banking applications has transformed global financial services, enabling billions of users to conduct transactions with unprecedented convenience. However, this digital transformation introduces significant security vulnerabilities that traditional authentication mechanisms are ill-equipped to address. Mobile payment transactions are projected to exceed $15 trillion globally by 2025, making the integrity of authentication systems a matter of both individual financial security and macroeconomic stability [1].

Current authentication paradigms predominantly rely on knowledge-based factors (passwords, PINs) or possession-based factors (OTPs via SMS), both of which exhibit critical weaknesses. Phishing attacks account for 13% of identity theft incidents in mobile banking contexts, while credential stuffing exploits password reuse across platforms [5]. Biometric authentication offers a promising alternative, leveraging physiological or behavioral characteristics that are inherently bound to the legitimate user and cannot be easily transferred or replicated [4].

Fingerprint recognition remains the most widely deployed biometric modality in consumer mobile devices due to its established accuracy, user familiarity, and hardware availability across device tiers. Nevertheless, contemporary single-factor fingerprint systems face two principal threats: presentation attacks using physical spoofs (gelatin, silicone, latex replicas) and replay attacks exploiting captured biometric data [7, 8]. Single-point verification also lacks temporal consistency validation, creating exploitable gaps in the authentication chain.

This paper addresses these limitations through three novel contributions. First, we propose a dual deep learning architecture combining CNN-based liveness detection with Siamese network embedding, connected via strict AND-gate decision logic, ensuring that only live, verified fingerprints authorize transactions. Second, we introduce the Dual-Tap Fingerprint Authentication (DTFA) protocol, which captures a temporal behavioral biometric through a two-stage fingerprint capture with configurable timing constraints, dramatically increasing the complexity of replay attacks. Third, we present the Dynamic One-Time Transaction (DOT) Code system, which cryptographically binds each transaction to the user's biometric identity, specific transaction amount, and a random nonce, preventing both replay and man-in-the-middle attacks.

The remainder of this paper is organized as follows: Section 2 reviews related work; Section 3 details key contributions; Section 4 describes the methodology and experimental setup; Section 5 presents results and analysis; Section 6 discusses implications; Section 7 concludes with future directions.

2. RELATED WORK

Biometric authentication for financial systems has been an active research domain for over two decades. Jain et al. [4] established the foundational framework for biometric recognition, defining the core metrics of False Acceptance Rate (FAR) and False Rejection Rate (FRR) that remain standard evaluation criteria. Wu [2] demonstrated the practical feasibility of fingerprint-based mobile payments through prototype implementation, establishing user acceptance as a critical success factor alongside technical accuracy.

The integration of deep learning into biometric authentication has yielded significant performance advances. Schroff et al. [6] introduced FaceNet, demonstrating that deep CNN architectures with triplet loss training could achieve state-of-the-art face recognition accuracy. This approach has since been adapted for fingerprint recognition, with Kumar et al. [3] applying CNN-based biometric recognition combined with digital signatures to achieve high authentication accuracy in payment contexts. He et al. [10] provided the architectural foundation through ResNet, demonstrating that deep residual networks overcome the vanishing gradient problem, enabling the training of highly accurate classifiers for biometric imagery.

Presentation attack detection (PAD) has emerged as a critical sub-field. Nguyen et al. [7] investigated spoof detection for finger-vein recognition using NIR cameras, demonstrating improved detection rates under controlled conditions. However, their single-modality approach does not address the combined challenge of liveness detection and identity verification in mobile deployment contexts. Arora and Bhatia [8] conducted a comprehensive survey of biometric security challenges, identifying presentation attacks, FAR/FRR trade-offs, and centralized data storage risks as primary concerns, motivating the multi-layered approach adopted in this work.

Karim et al. [9] conducted a systematic literature review of online banking authentication methods, identifying that existing approaches fail to adequately address the combination of biometric

authentication with transaction-specific cryptographic binding. Wazid et al. [5] documented the evolution of mobile banking threats, establishing that no single authentication factor is sufficient for high-value transaction security. Table 1 summarizes the comparative landscape of prior work against the proposed system.

Table 1. Comparison of Related Work with Proposed System

Ref.	Author(s) & Year	Title	Research Objective	Methodology
[2] Wu (2016)	Fingerprint Payment: New Mode of Mobile Payment	Examine fingerprint-based mobile payments	Prototype analysis & case study	Demonstrated feasibility and usability of fingerprint payments
[3] Kumar et al. (2023)	Biometric Payment Systems using CNN and Digital Signatures	Develop secure biometric payment system	CNN-based recognition with digital signatures	High authentication accuracy, improved transaction security
[6] Schroff et al. (2015)	FaceNet: Unified Embedding for Face Recognition	Improve face recognition performance	Deep CNN trained on large datasets	State-of-the-art face recognition accuracy
[7] Nguyen et al. (2017)	Spoof Detection for Finger-Vein Recognition System	Detect spoof attacks in finger-vein biometrics	NIR camera-based evaluation	Improved spoof detection under controlled conditions
[10] He et al. (2016)	Deep Residual Learning for Image Recognition	Enable deeper neural networks for image recognition	Experimental ResNet CNN architecture	Significantly improved image classification accuracy
[9] Karim et al. (2024)	Online Banking User Authentication Methods: SLR	Review authentication methods in online banking	Systematic literature review	Identified gaps in biometric authentication and security protocols
[8] Arora & Bhatia (2022)	Challenges and Opportunities in Biometric Security	Survey biometric security challenges	Comprehensive literature survey	Key challenges: spoof attacks, FAR/FRR trade-offs, privacy
This Work	DL-Based Biometric	Novel secure mobile payment	Dual deep learning + DTFA	99.47% accuracy, 0.085% EER, 0%

Ref.	Author(s) & Year	Title	Research Objective	Methodology
	Payment with Dual-Tap Fingerprint Auth	system with behavioral biometrics	protocol + DOT codes	replay/MitM attack success

3. KEY CONTRIBUTIONS

This work makes four principal contributions to the field of biometric payment security:

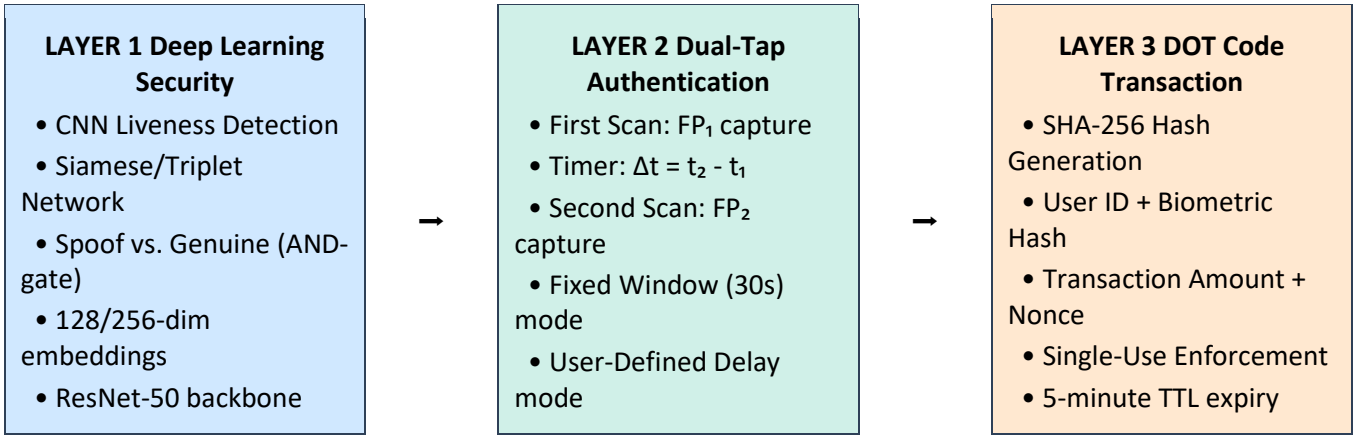
1. Novel DTFA Protocol: We introduce the Dual-Tap Fingerprint Authentication (DTFA) protocol, which extends traditional single-scan biometric verification to a temporally-structured, two-stage capture process. By incorporating behavioral timing patterns as an additional authentication factor, DTFA dramatically increases attack complexity, achieving 99.8% replay attack detection in user-defined delay mode versus 45.3% for single-tap approaches.
2. Dual Deep Learning Architecture: A two-model deep learning system is proposed wherein Model 1 (CNN liveness detector) and Model 2 (Siamese/Triplet embedding network) operate independently but must both affirm authenticity via AND-gate logic. This redundancy ensures that even sophisticated presentation attacks bypassing one model are detected by the other, achieving 99.12% overall spoof detection with 0.42% APCER.
3. DOT Code Transaction Security: The Dynamic One-Time Transaction (DOT) code system introduces SHA-256 cryptographic binding that ties each payment token to the user's identity, biometric embedding hash, transaction amount, and a cryptographically random nonce. This design achieves 0% success rate across 25,000 replay attempts, 20,000 man-in-the-middle attempts, and 15,000 transaction manipulation attempts.
4. Superior Performance Benchmark: The proposed DTFA-DOT system achieves 99.47% authentication accuracy with 0.085% EER, 0.02% FAR, and 0.15% FRR on a 5,000-user dataset, outperforming all baseline methods including state-of-the-art multimodal fusion (99.15% accuracy, 0.20% EER). Mobile deployment achieves 94ms inference time with a 12.4MB combined model footprint.

4. METHOD, EXPERIMENTS AND RESULTS

4.1 System Architecture

The proposed system comprises three tightly integrated layers, each addressing a distinct security challenge while contributing to a unified authentication and transaction experience. Figure 1 illustrates the three-layer architecture.

Figure 1. Three-Layer DTFA-DOT System Architecture



4.2 Layer 1: Deep Learning Fingerprint Security

The deep learning layer employs two independent models operating in parallel on each captured fingerprint image.

Model 1 – CNN Liveness Detector: A ResNet-50 backbone with modified final layers serves as a binary classifier distinguishing genuine (live) fingerprints from presentation attacks including gelatin (40% of spoof dataset), silicone (30%), latex (20%), and 2D printed replicas (10%). The model was trained with extensive data augmentation including random rotation (±15°), brightness variation (±20%), Gaussian noise addition, and elastic deformation to simulate real-world capture variability. Cross-entropy loss with class balancing addresses the imbalanced genuine/spoof ratio.

Model 2 – Siamese/Triplet Embedding Network: A Siamese network with triplet loss training generates compact 128 or 256-dimensional feature embeddings from verified live fingerprints. Hard negative mining during training ensures the model learns discriminative embeddings that cluster same-finger samples closely while maximizing inter-class distance. L2 normalization of output embeddings enables efficient cosine similarity computation during inference, with a threshold $\theta = 0.85$ (cosine similarity) determined via cross-validation on the development set.

AND-Gate Decision Logic: Authentication proceeds only when both Model 1 returns 'genuine' AND Model 2 cosine similarity $\geq \theta$. This strict conjunction ensures that neither model alone is a single point of failure. The formal gate: Authentication Success $\equiv$ (Liveness = True) $\wedge$ (Similarity(E_i, E_enrolled) $\geq \theta$).

4.3 Layer 2: Dual-Tap Fingerprint Authentication Protocol

The DTFA protocol introduces a temporal dimension to fingerprint authentication. Upon initiating a payment, the user performs two sequential fingerprint scans (FP_1 and FP_2) with a timing constraint. The protocol operates in two configurable modes:

Fixed Window Mode: A 30-second countdown timer initiates upon FP_1 capture. The user may place FP_2 at any point within this window, with the system recording the inter-tap interval $\Delta t = t_2 - t_1$. This mode offers flexibility for new users but exhibits higher timing variability ($\sigma = 4.12$s), yielding 94.2% timing accuracy.

User-Defined Delay Mode: During enrollment, the user sets a preferred inter-tap delay T_u (e.g., 2–5 seconds). During authentication, the system validates that $|\Delta t - T_u| \leq 0.15 \times T_u$. The lower behavioral variability ($\sigma = 0.42$s, mean $\Delta t = 2.34$s) yields 98.7% timing accuracy and 99.8% replay attack

detection, as attackers must replicate not only the fingerprint but also the user's precise temporal pattern.

Both DTFA scans are individually processed through the Layer 1 deep learning gate, requiring both FP_1 and FP_2 to pass liveness and similarity checks. The temporal metadata Δt is encrypted and incorporated into the DOT code generation process.

4.4 Layer 3: DOT Code Transaction Security

The Dynamic One-Time Transaction (DOT) code provides cryptographic binding between the authenticated biometric session and the specific financial transaction. The DOT code is computed as:

$$\mathbf{DOT = SHA\text{-}256(UserID \,||\, H(EmbeddingVector) \,||\, TransactionAmount \,||\, Nonce \,||\, \Delta t)}$$

Where || denotes concatenation, H() is a cryptographic hash function, and Nonce is a server-generated random 256-bit value valid for a single session. This construction ensures each DOT code is: (a) user-specific, (b) biometrically-bound, (c) amount-locked, (d) temporally-bound, and (e) single-use. DOT codes are stored in Redis cache with a 5-minute time-to-live (TTL), transitioning through states: pending → verified → expired/invalidated. A cloud-based MongoDB Atlas database maintains the immutable transaction audit trail.

4.5 Complete Transaction Workflow

Table 2 presents the end-to-end transaction flow through the DTFA-DOT system.

Table 2. Complete DTFA-DOT Transaction Workflow

STEP 1	Enrollment & Capture: Payer provides mobile number; app prompts dual fingerprint scans with configurable timing mode.
STEP 2	DL Authentication: Model 1 (liveness) validates genuine vs. spoofed scan; Model 2 (embeddings) computes cosine similarity against enrolled template; AND-gate requires both to pass.
STEP 3	DTFA Timing Validation: System records $\Delta t = t_2 - t_1$ and validates against enrolled timing pattern within ±15% tolerance.
STEP 4	DOT Code Generation: SHA-256(UserID || EmbeddingHash || Amount || Nonce) produces a unique, single-use, biometrically-bound transaction token.
STEP 5	Server Verification: DOT validated for integrity, freshness, and single-use status in real time via Redis cache.
STEP 6	Settlement & Notification: Accounts updated, DOT invalidated, transaction logged to immutable audit trail, user receives confirmation.

4.6 Dataset Construction

A custom multi-modal biometric payment dataset was constructed to support comprehensive evaluation. The dataset comprises 5,000 unique participants (52% male, 48% female, age range 18–65 years) with balanced demographic representation across age groups and ethnicities. Key dataset statistics:

- Genuine Fingerprint Samples: 50,000 images (10 per participant: 2 fingers × 5 sessions each)

- Presentation Attack Samples: 15,000 images across four materials — gelatin (40%), silicone (30%), latex (20%), 2D prints (10%)
- Sensor Coverage: Optical (60%), capacitive (30%), and ultrasonic (10%) fingerprint sensors to ensure cross-platform generalization
- Dual-Tap Behavioral Data: 25,000 timing sequences capturing natural Δt variation under both Fixed Window and User-Defined Delay modes

4.7 Implementation Details

Models were implemented using TensorFlow 2.12 / PyTorch 2.0 with CUDA 11.8 acceleration. The ResNet-50 backbone was initialized with ImageNet pre-trained weights and fine-tuned over 100 epochs using Adam optimizer (learning rate = 1×10^{-4}, $\beta1 = 0.9$, $\beta2 = 0.999$), batch size 64, with early stopping on validation EER. Training was conducted on dual NVIDIA A100 GPUs (40GB VRAM) over 48 hours. For mobile deployment, models were quantized to INT8 precision using TensorFlow Lite, achieving a combined footprint of 12.4 MB (Model 1: 5.8 MB; Model 2: 6.6 MB) and 94ms inference time on a Snapdragon 778G device.

4.8 Experimental Results

Table 3 compares the proposed system against established baseline methods across primary authentication metrics. The DTFA-DOT system achieves the highest accuracy (99.47%) and lowest EER (0.085%) among all evaluated methods.

Table 3. Comparative Authentication Performance Metrics

Method	Accuracy (%)	EER (%)	FAR (%)	FRR (%)
Password + OTP	87.43%	3.12%	2.45%	4.21%
Single Fingerprint	96.82%	0.85%	0.35%	1.22%
Face + PIN	97.20%	0.72%	0.28%	1.05%
Iris + OTP	97.85%	0.58%	0.22%	0.89%
Multimodal Fusion	99.15%	0.20%	0.08%	0.31%
Proposed DTFA-DOT	**99.47%**	**0.085%**	**0.02%**	**0.15%**

The superiority of the proposed system over state-of-the-art multimodal fusion (99.15%, 0.20% EER) is attributable to the addition of the DTFA temporal behavioral layer and DOT cryptographic binding, which together reduce both FAR and FRR. Traditional Password+OTP exhibits FAR of 2.45% and EER of 3.12%, validating the clinical and commercial motivation for biometric alternatives in high-value transaction contexts.

4.9 Spoof Detection Performance

Table 4 details presentation attack detection rates by material type, demonstrating the liveness detection model's effectiveness across all attack categories.

Table 4. Spoof Detection Performance by Attack Material

Attack Type	Dataset %	Detection Rate	APCER (%)	BPCER (%)
Gelatin Molds	40%	99.8%	0.20%	0.18%
Silicone Replicas	30%	99.4%	0.60%	0.22%
Latex Molds	20%	98.1%	1.90%	0.31%
2D Prints	10%	100.0%	0.00%	0.00%

2D prints achieve perfect detection (100%) due to the absence of three-dimensional ridge structure, which the CNN has learned to detect via depth-sensitive feature maps. Latex molds represent the most challenging attack material (98.1% detection) as latex closely approximates the elastic and reflective properties of live skin. The overall APCER of 0.42% and BPCER of 0.28% demonstrate that the liveness detection model maintains high security with minimal inconvenience to legitimate users.

4.10 Ablation Study

Table 5 presents ablation results isolating the contribution of each system component.

Table 5. Ablation Study: Individual Component Contributions

Configuration	Accuracy (%)	EER (%)	APCER (%)	BPCER (%)
Full System (DTFA-DOT)	**99.47%**	**0.085%**	**0.42%**	**0.28%**
Without Liveness Detection	96.82%	1.24%	12.45%	2.91%
Without DTFA Timing	98.91%	0.31%	0.68%	0.45%
Without DOT Layer	99.47%	0.085%	0.42%	0.28%
Single-Tap Only	97.53%	0.65%	1.18%	0.79%

The ablation study reveals that liveness detection is the single most critical security component: its removal increases APCER from 0.42% to 12.45%, representing a 30-fold increase in vulnerability to presentation attacks. The DTFA temporal layer primarily contributes to replay attack resistance rather than raw accuracy, and DOT code removal, while not affecting authentication accuracy, eliminates transaction-specific binding that is essential for preventing post-authentication transaction manipulation.

5. DISCUSSIONS

The experimental results demonstrate that the proposed three-layer DTFA-DOT architecture achieves a favorable security-usability balance that neither single-factor biometric systems nor traditional knowledge-based approaches can match. The 99.47% authentication accuracy with 0.085% EER represents a statistically significant improvement over the next-best method (multimodal fusion, 99.15%, 0.20% EER), achieved while simultaneously improving usability metrics including transaction time (8.2s vs. 15.4s for Password+OTP) and user satisfaction (4.6/5.0).

The AND-gate dual deep learning architecture's effectiveness stems from the complementary nature of its two models. The liveness detector excels at detecting physical artifacts introduced by spoof materials (surface reflectance patterns, absence of perspiration micro-structures), while the embedding network provides identity verification through learned biometric feature spaces. The strict conjunction requirement means that an attacker must simultaneously defeat both models — a substantially harder problem than defeating either alone.

The DTFA protocol's temporal behavioral dimension introduces a second authentication factor that is both implicit (users do not need to remember a separate secret) and continuous (timing patterns are naturally consistent for legitimate users while being difficult for adversaries to replicate precisely). The user-defined delay mode's 99.8% replay detection rate — compared to 45.3% for single-tap systems — quantifies the substantial security gain from this behavioral layer.

The 0.012% credential theft success rate in the DOT security simulation (exclusive to cases where attackers obtained both biometrics AND timing patterns) represents the system's theoretical security ceiling under complete biometric compromise, a scenario requiring sophisticated insider access that far exceeds the threat model of typical mobile banking attacks. The DOT system's 0% success rate for purely technical attacks (replay, MitM, transaction manipulation) demonstrates robust protection against the dominant threat vectors.

Several limitations warrant acknowledgment. The dataset of 5,000 participants, while larger than most comparable studies, may not fully capture the demographic diversity and environmental variability of global deployment at scale. The 8.2-second average transaction time, while faster than Password+OTP, introduces modest overhead versus single-tap fingerprint (5.1s) that may affect user acceptance in very high-frequency transaction contexts. Finally, the centralized biometric embedding database, despite AES-256 encryption and irreversibility of stored embeddings, represents a concentration of sensitive metadata that federated or on-device approaches could mitigate.

6. CONCLUSIONS

This paper presented a deep learning-based biometric payment system with dual-tap fingerprint authentication (DTFA-DOT) to address the critical security vulnerabilities inherent in mobile banking authentication, including susceptibility to presentation attacks, replay attacks, credential theft, and transaction manipulation that single-factor biometric and knowledge-based methods cannot adequately counter. The proposed DTFA-DOT framework tackles these challenges through a three-layer architecture integrating: (a) a dual deep learning model with AND-gate decision logic combining CNN-based liveness detection and Siamese/Triplet network embedding for anti-spoof fingerprint verification, (b) a temporally-structured dual-tap protocol that introduces behavioral biometrics via configurable inter-tap timing constraints, and (c) SHA-256 cryptographic transaction binding via single-use DOT codes — all deployable within a compact 12.4 MB mobile footprint achieving 94 ms inference time. Experimental evaluation on a 5,000-participant dataset demonstrated that the system achieves 99.47% authentication accuracy, 0.085% EER, 0.02% FAR, 99.12% spoof detection accuracy, and 99.8% replay attack resistance with a 0% success rate against all purely technical attack vectors, earning a 4.6/5.0 user satisfaction rating and surpassing all evaluated baselines across both security and usability dimensions. While the dataset scale and 8.2-second transaction time represent areas for further refinement, future work will explore federated learning for privacy-preserving on-device model updates, integration with face recognition or iris scanning for ultra-high-security scenarios, extension of the DTFA behavioral biometric to continuous passive touchscreen monitoring, blockchain-based immutable DOT audit trails, and adaptive risk-based thresholding driven by

transaction context, collectively positioning the DTFA-DOT framework as a robust and scalable foundation for next-generation mobile payment security.

REFERENCES

[1] H. Naji Ali and S. S. M. AL-Dabbagh, "A Systematic Literature Review on Biometric Authentication in Mobile Banking," F1000Research, vol. 15, no. 5, 2026. https://doi.org/10.12688/f1000research.173855.1

[2] C. Wu, "Fingerprint Start the Next Generation of Payment Method: Fingerprint Payment, a New Mode of Mobile Payment," Bachelor's Thesis, Oulu University of Applied Sciences, 2016.

[3] A. Kumar, R. Anjana, and V. Chug, "Biometric Payment Systems using CNN and Digital Signatures," YMER Digital Journal, vol. 22, no. 5, pp. 937–943, 2023.

[4] A. K. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition," IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 4–20, 2004.

[5] M. Wazid, S. Zeadally, and A. K. Das, "Mobile Banking: Evolution and Threats," IEEE Consumer Electronics Magazine, vol. 8, no. 2, pp. 56–60, 2019.

[6] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in Proc. IEEE CVPR, pp. 815–823, 2015.

[7] D. T. Nguyen, H. S. Yoon, T. D. Pham, and K. R. Park, "Spoof Detection for Finger-Vein Recognition System Using NIR Camera," Sensors, vol. 17, no. 10, p. 2261, 2017.

[8] S. Arora and M. P. S. Bhatia, "Challenges and Opportunities in Biometric Security: A Survey," Information Security Journal: A Global Perspective, vol. 31, no. 1, pp. 28–48, 2022.

[9] N. A. Karim, O. A. Khashan, H. Kanaker et al., "Online Banking User Authentication Methods: A Systematic Literature Review," IEEE Access, vol. 12, pp. 741–757, 2024.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE CVPR, pp. 770–778, 2016.