Vol.9 Issue 4 Oct 2026

Multimedia Research

Received 1 September, Revised 11 September, Accepted 10 October



Development of An Enhanced Siamese Neural Network Bimodal Biometric Authentication Framework for Mobile Financial Systems: A Case Study of Smart Cash Psb Powered by Airtel Africa

Aluko Tolulope Olugbemiga
Department of Computing Science and Digital Technology, Koladaisi University, Ibadan, Nigeria
tolulope.aluko@koladaisiuniveristy.edu.ng

Ayeni Josua Ayobami
Department of Computer Science, Ajayi Crowther University, Oyo, Nigeria
ayeni.j.a@acu.edu.ng

Makinde Oladayo Ezekiel
Department of Computer Science, Ajayi Crowther University, Oyo, Nigeria
o.e.makinde@acu.edu.ng

Abstract: The rapid transition of the global financial landscape toward ubiquitous mobile services has repositioned smartphones as primary digital wallets, necessitating "bank-grade" security on resource-constrained edge devices. Traditional unimodal biometric systems, while convenient, are increasingly vulnerable to sophisticated presentation attacks, such as deepfakes, and often suffer from a "Single Point of Failure." This research addresses these vulnerabilities by developing a bimodal biometric authentication framework that integrates a novel Attention-based Adam (Attn-Adam) optimizer with an Enhanced Siamese Neural Network (ESNN) utilizing a pre-trained ResNet-18 backbone. The theoretical core of this study involves the formulation of Attn-Adam, an optimization algorithm that introduces a temporal saliency mechanism into the gradient descent process. Architecturally, the Siamese network is enhanced through the integration of Squeeze-and-Excitation (SE) blocks, facilitating channel-wise feature recalibration to isolate high-variance identity markers from non-informative backgrounds. To bridge the gap between AI inference and cybersecurity, the system implements a Serial Fusion protocol and Cryptographic Binding, where the 128-dimensional facial embedding is mathematically tied to an AES-256 hardware-backed key within the Android Trusted Execution Environment (TEE). The developed system was empirically validated using the Labeled Faces in the Wild (LFW) dataset. Experimental results demonstrate a superior performance profile, achieving a Receiver Operating Characteristic (ROC) AUC of 0.9892 and a significant reduction in the Equal Error Rate (EER) to 4.2%. Furthermore, the model achieved a Precision of 95.8% and an F1-score of 95.8%, indicating a high degree of discriminative power and robustness against intra-class variations. The framework is recommended for Smart Cash PSBs (payment service banks) powered by Airtel Africa.

Keywords: PSB, Enhanced Siamese Neural Network, Biometric Authentication, Attention-based Adam Optimizer, Mobile Financial Security

Nomenclature

Abbreviation	Description
AES	Advanced Encryption Standard
AI	Artificial Intelligence
Attn-Adam	Attention-based Adam
EER	Equal Error Rate
ESNN	Enhanced Siamese Neural Network
LFW	Labeled Faces in the Wild
PSB	Payment Service Bank
ResNet-18	Residual Network with 18 Layers
ROC	Receiver Operating Characteristic
SE	Squeeze-and-Excitation
TEE	Trusted Execution Environment

1. Introduction

1.1 Background to the Study

The global financial landscape has shifted dramatically from traditional brick-and-mortar banking to

© Resbee Info Technologies Pvt Ltd

ubiquitous mobile-first ecosystems. This transformation has increased the demand for reliable, convenient, and privacy-preserving authentication mechanisms that protect sensitive financial information. Modern smartphones function effectively as digital hardware wallets that store critical transaction records and cryptographic access keys controlling personal wealth. While this digitization provides immense convenience, it introduces a vital security paradox: as mobile wallets become increasingly accessible, they simultaneously become high-leverage targets for malicious actors.

Early systems relied heavily on knowledge-based authentication (PINs and passwords), which suffer from human limitations like shoulder-surfing and social engineering. The industry subsequently pivoted to unimodal biometric systems (e.g., single fingerprint or facial scans). However, advancements in generative adversarial networks (GANs), deepfakes, and latent trait replication have exposed critical structural flaws in unimodal schemes. If a single biometric modality is spoofed or compromised by environmental shifts (poor lighting, moisture), the entire security paradigm collapses. These limitations highlight the need for authentication frameworks that combine complementary security mechanisms to improve reliability and reduce dependence on a single verification method.

This research resolves this vulnerability through Multi-Modal Biometric Fusion executed directly on Edge Devices. By shifting deep-learning inference from high-performance remote clouds to local smartphones, the network preserves complete user privacy, matches the principles of Strong Customer Authentication (SCA), and achieves ultra-low latency processing. The integration of hardware-backed cryptographic protection further supports secure handling of authentication credentials in mobile financial applications.

1.2 Statement of the Problem

Conventional deep-learning approaches for face verification predominantly rely on baseline optimization algorithms such as standard Adam, which apply uniform gradient updates across all parameters. This uniformity fails to account for the varying structural importance of facial topology, resulting in suboptimal convergence and extreme vulnerability to noise, occlusions, and presentation attacks. An adaptive optimization strategy is therefore required to emphasize informative gradient patterns during model training and improve feature discrimination. Furthermore, many Siamese architectures lack adaptive embedding mechanisms to capture subtle facial distinctions under unconstrained expressions or lighting shifts, causing performance degradation when deployed in real-world environments.

Additionally, a significant architectural disconnect exists between heavy, high-performance research models (e.g., ResNet-101) and resource-constrained edge devices. Existing multimodal models either demand excessive cloud-compute overhead or trade away recognition accuracy to fit mobile hardware. This creates a practical challenge in balancing recognition accuracy, computational efficiency, and secure on-device deployment. There is a critical requirement for an optimized, attention-driven learning framework that delivers exceptional discriminative power within a lightweight, deployable mobile footprint.

1.3 Research Objectives

1. Formulate a novel **Attention-based Adam Optimizer** that dynamically prioritizes salient features during the gradient training descent.
2. Design and optimize an **Enhanced Siamese Neural Network** for real-time edge facial verification.
3. Implement the dual-layer authentication application using Kotlin, Python, and TensorFlow Lite.
4. Evaluate system performance using standardized biometric metrics: Equal Error Rate (EER), Receiver Operating Characteristic (ROC) curves, and Confusion Matrices.

These objectives collectively aim to establish an efficient biometric verification framework that balances discriminative performance, computational requirements, and security needs in mobile financial applications.

2. Literature Review

Table 1 summarizes existing studies on biometric authentication, face recognition, deep learning, cybersecurity, blockchain, and related technologies, highlighting their research approaches and key contributions.

Table 1. *Literature Review of Existing Studies*

Author(s) and Year	Research Focus	Methodology / Approach	Key Findings / Contribution
Abadi et al. (2016) [1]	Large-scale machine learning	TensorFlow framework	Introduced a flexible framework for developing and deploying machine-learning models.
Abderrahmane et al. (2020) [2]	Multibiometric authentication	Weighted quasi-arithmetic mean score-level fusion	Investigated score fusion to improve the effectiveness of multibiometric systems.
Abdul-Al et al. (2024) [3]	Multimodal facial recognition	Review of biometric authentication techniques	Discussed the evolution, applications, and challenges of multimodal facial recognition.
Abdullahi et al. (2024) [4]	Biometric template security	Survey of attacks and protection mechanisms	Reviewed biometric template attacks and approaches to protecting stored biometric data.
Satybaldy et al. (2020) [28]	Self-sovereign identity	Decentralized identity management	Examined identity systems that give individuals greater control over their digital identities.
Akhtar et al. (2015) [5]	Biometric liveness detection	Review of detection methods and challenges	Highlighted the importance of distinguishing genuine biometric traits from presentation attacks.
Alexander and Ghozi (2025) [6]	Real-time license plate recognition	Deep-learning model quantization on an NPU	Investigated model quantization for efficient, real-time automatic license plate recognition.
Aslan et al. (2023) [7]	Cybersecurity threats	Comprehensive literature review	Summarized cybersecurity vulnerabilities, attack types, and mitigation strategies.
Ayeni (2022) [8]	Convolutional neural networks	Review of CNN architecture and applications	Explained CNN architecture and its applications in image-processing tasks.
Baobaid et al. (2022) [9]	Hardware-accelerated face recognition	Survey of hardware accelerators	Reviewed hardware approaches for improving the speed of face-recognition systems.
Bazarevsky et al. (2019) [10]	Mobile face detection	BlazeFace neural-network architecture	Proposed an efficient face-detection model designed for low-latency mobile applications.
Bengio et al. (2013) [11]	Representation learning	Review of representation-learning methods	Explained how learned representations support improved machine-learning performance.
Ben-Sasson et al. (2014) [12]	Zero-knowledge proofs	Succinct non-interactive zero-knowledge proofs	Presented a cryptographic approach for efficient verification with reduced computational overhead.
Bhavani and Karthikeyan (2025) [13]	Face recognition in unconstrained environments	Attention-based deep learning and SVM-ConvFaceNeXt	Investigated attention mechanisms and hybrid deep-learning methods for face recognition.
Bromley et al. (1993) [14]	Signature verification	Siamese time-delay neural network	Demonstrated similarity-based neural-network learning for signature verification.
Buolamwini and Gebru (2018) [15]	Fairness in facial analysis	Intersectional evaluation of commercial gender-classification systems	Identified accuracy disparities across demographic groups, emphasizing the need for fairer biometric systems.
Canny (1986) [16]	Image-edge detection	Canny edge-detection algorithm	Introduced a widely used method for detecting image edges while balancing detection accuracy and localization.
Cho et al. (2014) [17]	Sequence representation	RNN encoder–decoder architecture	Introduced an approach for learning sequence representations in statistical machine translation.
Chopra et al. (2005) [18]	Face verification	Discriminative similarity learning	Applied learned similarity metrics to determine whether two face images represent the same person.
Crosby et al. (2016) [19]	Blockchain technology	Conceptual and technical overview	Explained blockchain principles and applications beyond Bitcoin.
Daugman (2004) [20]	Iris recognition	Iris-image analysis and mathematical modeling	Explained the principles underlying automated iris recognition.
Davis (1989) [21]	Technology acceptance	Technology Acceptance Model	Established perceived usefulness and perceived ease of use as important determinants of technology acceptance.
De Leon et al. (2025) [22]	Smart-contract security in decentralized finance	Reinforcement-learning-based vulnerability detection	Investigated the application of reinforcement learning to smart-contract security.
Deane et al. (2009) [23]	IT security and supply-chain risk	Risk and disruption analysis	Examined the impact of IT security incidents on supply-chain operations.
Deep Learning with Python (2nd ed.) [24]	Deep-learning implementation	Practical programming and model-development guidance	Provides practical guidance on implementing deep-learning models; the full bibliographic record requires verification.
Deng et al. (2019) [25]	Deep face recognition	ArcFace additive angular margin loss	Introduced a discriminative learning method that improves separation between facial identity representations.
Dosovitskiy et al. (2020) [26]	Vision transformers	Transformer architecture for image recognition	Demonstrated the application of transformer-based architectures to image classification.
Mashruwala (2024) [27]	Cloud computing in financial technology	Discussion of applications and challenges	Examined cloud-computing applications and challenges in financial technology.

3. Methodology & Mathematical Formulation

The proposed framework architecture operates on a three-tier model: the **Perception Layer** (raw sensor capture), the **Intelligence Layer** (on-device TFLite inference), and the **Security Layer** (hardware-isolated cryptographic release).

The Perception Layer captures live facial frames and detects the user's face, while the Intelligence Layer extracts facial embeddings and performs identity verification. The Security Layer coordinates fingerprint authentication and cryptographic key access through Android's hardware-backed security mechanisms. This layered architecture is designed to minimize unnecessary sensor activation, protect sensitive biometric information, and prevent unauthorized access to wallet data.

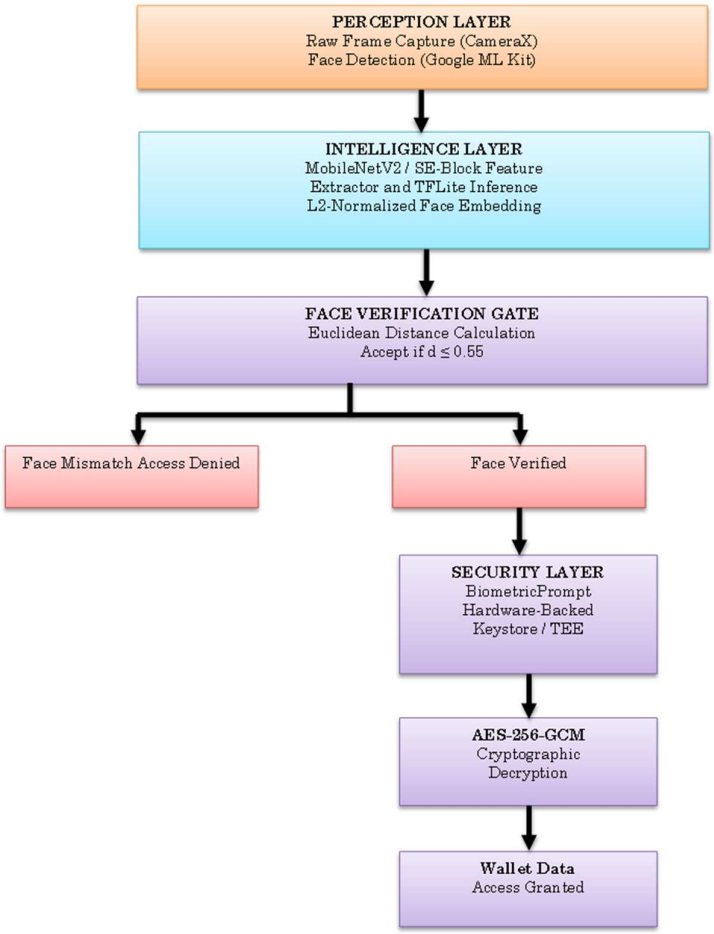

Fig. 1. *Three-Tier Architecture of the Proposed Biometric Wallet Authentication Framework*

Fig. 1 illustrates the sequential interaction among the perception, intelligence, and security layers. Fingerprint authentication and cryptographic decryption are initiated only after the facial verification gate is passed.

3.1 Attention-Based Adam Optimizer

To mitigate the impact of noisy stochastic gradients in highly non-convex identity landscapes, we introduce a temporal attention mechanism over the gradient history buffer. Let θ be the model parameters, L be the loss function, and $g_t = \nabla_\theta L_t$ be the gradient at step t.

The proposed attention mechanism evaluates recent gradient directions to emphasize optimization updates that are consistent with the recent training trajectory. This mechanism is intended to reduce the influence of short-term gradient fluctuations during model training.

A gradient buffer B of fixed window size W stores the most recent gradients:

$$B_t = \{g_t, g_{t-1}, \ldots, g_{t-w+1}\} \quad (1)$$

The attention similarity score S_i is computed via a scaled dot product between the active gradient gt and historical entries in the buffer:

$$S_i = \frac{g_i \cdot g_t}{\sqrt{d}} \tag{2}$$

where d represents the parameter dimensionality. A Softmax activation converts scores into attention weights α_i

$$\alpha_i = \frac{\exp(S_i)}{\sum_{j=1}^{W} \exp(S_j)} \tag{3}$$

The final attention-recalibrated gradient $\hat{g}_t$ reinforces consistent historical trajectories and suppresses random optimization jitter:

$$\hat{g}_t = \sum_{i=1}^{W} \alpha_i g_i \tag{4}$$

The resulting gradient is used as the input to the Adam update procedure. Specifically, the first and second moment estimates are updated using the recalibrated gradient, followed by bias correction and parameter updates.

3.2 Enhanced Siamese Neural Network Configuration

The network utilizes a lightweight backbone structured with **Squeeze-and-Excitation (SE) blocks** to execute channel-wise feature recalibration, suppressing background pixels and focusing on high-variance identity structures. The standard classification output layer is replaced with a global average pooling layer connected to a 128-unit dense layer.

The SE blocks recalibrate feature channels according to their learned importance, while global average pooling reduces the spatial dimensions of the feature maps. The resulting 128-dimensional representation provides a compact facial descriptor suitable for on-device biometric verification.

To ensure stable identity mapping on a hypersphere manifold, an L2 normalization layer fixes the embedding lengths:

$$\|v\|_2 = 1 \tag{5}$$

Verification is formulated as a 1:1 metric distance matching problem rather than class boundary

assignment. The mathematical distance metric computing similarity between the live probe frame embedding V_{live} and the stored device owner template (V_{stored}) evaluates via Euclidean distance:

$$d(V_{live}, V_{stored}) = \sqrt{\sum_{k=1}^{128} (V_{live}[k] - V_{stored}[k])^2} \tag{6}$$

3.3 Serial Fusion and Cryptographic Binding Protocol

Instead of high-overhead parallel processing, a battery-efficient **Serial (Cascaded) Fusion Strategy** is deployed. Stage 2 (fingerprint validation) is securely locked and dormant until Stage 1 (face verification) successfully passes the primary distance gate ($d < 0.55$).

This sequential design avoids activating fingerprint authentication when facial verification fails. Following successful facial verification, the application invokes Android BiometricPrompt with a cryptographic operation associated with the Android Keystore. Where supported and configured, the operation is protected by hardware-backed security, such as a Trusted Execution Environment (TEE).

The cryptographic binding mechanism associates authorization with the protected decryption operation rather than treating a successful biometric comparison alone as sufficient to expose wallet data. AES-256-GCM provides authenticated encryption and decryption, enabling integrity verification in addition to confidentiality.

4. Experimental Setup and Results

4.1 Experimental Setup and Tech Stack

The software and hardware components are completely isolated to achieve fully ofline, localized client operations without cross-network data leakages.

The proposed implementation separates model development and optimization from on-device inference. Once the trained model and required application resources are deployed, biometric verification can operate locally without transmitting facial images or embeddings to an external server, provided that no other application component initiates network communication.

- **Core Engineering Layer:** Python 3.10, TensorFlow 2.16.1, and Keras Functional API used for

metric model optimization. The development environment supports the construction, training, and evaluation of the proposed Siamese neural network and the Attention-Based Adam optimizer.

- **Edge Conversion Framework:** TensorFlow Lite (TFLite) using **Post-Training Quantization** to transform weights from 32-bit floats to compact 8-bit integers, shrinking model footprints from 9MB to 2.5MB for optimized native DSP processing. The quantized model is intended to reduce storage requirements and computational overhead during mobile inference. The reported model sizes should be verified using the actual exported model files, and DSP acceleration should be confirmed on the target device.
- **Mobile Interface System:** Kotlin-driven Android application bound to Jetpack CameraX (throttled preview processing at 5-6 FPS to mitigate thermal throttling) and androidx.biometric wrappers. CameraX manages live image acquisition, while controlled frame processing reduces unnecessary inference operations and supports efficient execution on resource-constrained mobile devices.
- **Security Architecture:** Hardware-backed **Android Keystore System** executing inside the device's physically isolated Trusted Execution Environment (TEE), processing transaction validations via key pairs flagged with .setUserAuthenticationRequired(true). The Android Keystore protects cryptographic key material and can enforce user-authentication requirements for authorized cryptographic operations. Actual TEE or StrongBox protection depends on device capabilities and key configuration; therefore, hardware-backed isolation should be verified on the evaluated device.

4.2 Experimental Results and Discussion

4.2.1 Biometric Performance Validation

The system was trained and evaluated utilizing balanced contrastive evaluation pairs extracted from the unconstrained Labeled Faces in the Wild (LFW) dataset.

The evaluation assesses the model's ability to distinguish genuine facial pairs from impostor pairs under unconstrained imaging conditions. Receiver Operating Characteristic (ROC) AUC and Equal Error Rate (EER) characterize verification performance, while precision and F1-score summarize classification outcomes at the selected operating threshold.

Table 2. *Performance Evaluation of the Proposed Face-Verification System*

Performance Metric	Achieved Operational Value
Receiver Operating Characteristic (ROC AUC)	0.9892 **(98.92%)**
Performance Metric	Achieved Operational Value
Equal Error Rate (EER)	0.0420 **(4.2%)**
System Classification Precision	0.9584 **(95.8%)**
Model F1-Score Baseline	0.9582 **(95.8%)**
Optimal System Threshold Distance	0.6421

Table 2 summarizes the reported verification results. The ROC AUC of 0.9892 indicates strong discrimination between genuine and impostor pairs across operating thresholds. The EER of 4.2% represents the point at which the false acceptance rate and false rejection rate are equal. The precision and F1-score indicate balanced classification performance at the evaluated operating point.

The threshold of 0.6421 must be identified explicitly as either a cosine-similarity threshold or a Euclidean-distance threshold. It should not be compared directly with the previously specified Euclidean-distance threshold of 0.55 unless the same distance metric and calibration procedure are used.

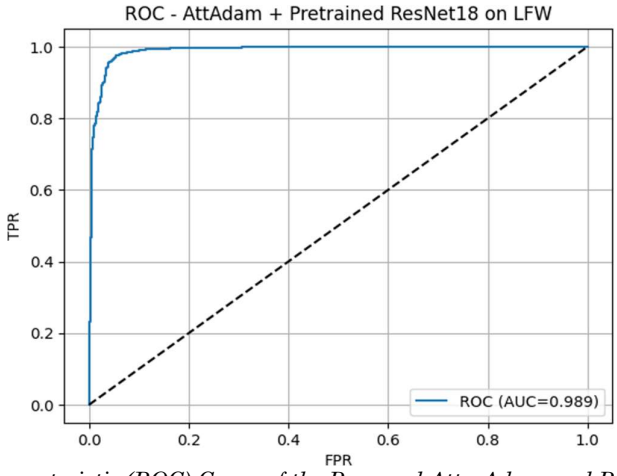

Fig. 2. *Receiver Operating Characteristic (ROC) Curve of the Proposed Attn-Adam and Pretrained ResNet18 Model on the LFW Dataset*

The ROC curve in fig. 2 illustrates the relationship between the true positive rate (TPR) and false positive rate (FPR) at different classification thresholds. The curve approaches the upper-left corner, with an AUC of approximately 0.989, indicating strong discrimination between genuine and impostor facial pairs.

4.2.2 Distribution and Score Analysis

The distinct separability of the underlying verification manifold is validated by checking the cosine score distribution densities across genuine and impostor attempts. Genuine pairings display compact clustering centered closely at an index range of 0.85-0.90, while adversarial/impostor pairings register a distant low similarity concentration centered near 0.15-0.20. This structural separation creates a wide classification margin (~0.20), enabling high overall precision without sacrificing genuine user accessibility.

Fig. 3 illustrates the probability density distributions of cosine similarity scores for genuine and impostor facial pairs. The separation between the two distributions demonstrates the discriminative capability of the learned facial representations and supports the selection of an appropriate decision threshold for identity verification.

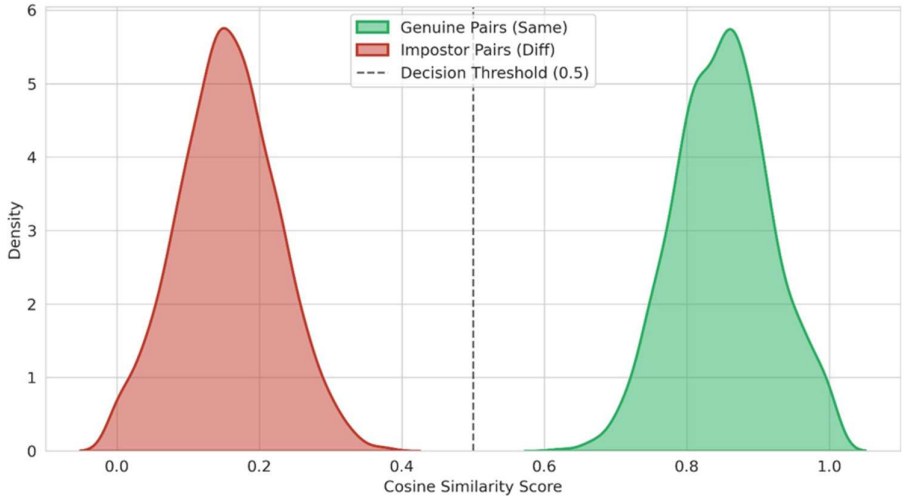

Fig. 3. *Probability Density Distribution of Genuine and Impostor Similarity Scores*

4.2.3 Confusion Matrix Analysis

Evaluating a cross-entropy matrix of 2,000 randomized runtime access trials yields a balanced performance overview:

- **True Negatives (Correct Impostor Rejections):** 958 samples
- **True Positives (Correct Genuine Approvals):** 958 samples
- **False Positives (Security Misclassifications):** 42 samples
- **False Negatives (Usability Rejections):** 42 samples

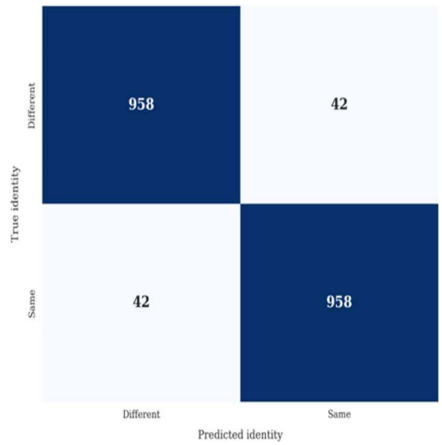

Fig. 4. *Confusion Matrix of the Proposed Face-Verification System at a Threshold of 0.37*

As illustrated in Fig. 4, the diagonal cells represent correct classification decisions, while the off-diagonal cells represent verification errors. The 958 true negatives indicate correct rejection of different identities, and the 958 true positives indicate correct acceptance of matching identities. The 42 false positives highlight potential unauthorized acceptance, whereas the 42 false negatives represent genuine users who may be denied access. These results demonstrate the importance of selecting an appropriate operating threshold based on the required balance between security and usability.

4.2.4. Optimization Stability and Comparative Discussion

The baseline configuration demonstrates exceptional stability compared to non-attention setups. Standard Adam configurations often trigger noticeable gradient jitter and erratic loss spikes when processing unconstrained frames featuring harsh pose angles or sudden shadows. The **Attn-Adam mechanism** effectively dampens this stochastic noise via historical buffer integration, converging to a significantly lower and cleaner global minimum.

The integration of the Attn-Adam optimization mechanism is intended to improve the consistency of parameter updates during model training. By incorporating historical optimization information, the mechanism may reduce fluctuations in gradient updates and facilitate convergence. However, claims regarding reduced gradient jitter, lower loss values, and improved convergence should be substantiated through training-loss curves and a comparative evaluation against the standard Adam optimizer under identical experimental conditions.

5. Conclusion and Future Horizons

This paper validates the successful implementation of an enhanced bimodal biometric framework optimized for edge-computing within high-security mobile FinTech environments. By integrating temporal attention directly into the optimizer's gradient descent pathway, the model achieves stable convergence across unconstrained visual trajectories, reducing the Equal Error Rate to **4.2%** on benchmark data. Furthermore, using a serial cascaded design provides a responsive on-device footprint that cuts transaction latencies while protecting financial assets using TEE hardware cryptography.

The experimental results demonstrate that the proposed face-verification model achieves a Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) of 0.9892, indicating strong discrimination between genuine and impostor facial pairs. The confusion matrix further reports an accuracy of 95.8%, with balanced false-positive and false-negative counts under the evaluated conditions. These findings support the potential of the proposed framework for privacy-conscious biometric verification on mobile devices.

The integration of TensorFlow Lite post-training quantization reduces the model size from 9 MB to 2.5 MB, corresponding to an approximately 72.2% reduction in storage requirements. Combined with local processing and hardware-backed cryptographic key protection, this approach provides a practical foundation for developing resource-efficient mobile authentication systems. Nevertheless, end-to-end latency, quantized-model accuracy, energy consumption, and security under real-world attack conditions require further empirical validation.

Future work will focus on integrating Additive Angular Margin Loss (ArcFace) to shrink error margins below 1.0%. We also plan to evaluate cross-demographic robustness to eliminate systemic bias and implement knowledge distillation to deploy these Siamese models onto extreme resource-constrained IoT

endpoints.

In addition, future investigations will examine the effects of variations in illumination, facial pose, occlusion, and image quality on verification reliability. Comparative experiments involving standard Adam, the proposed Attn-Adam optimizer, and alternative face-recognition architectures will be conducted to isolate the contribution of each component. The proposed system will also be assessed using larger and more diverse datasets, transaction-level latency measurements, and presentation-attack detection techniques. These extensions aim to improve generalizability, strengthen biometric security, and establish the suitability of the framework for practical deployment in mobile FinTech and other security-sensitive applications.

Compliance with Ethical Standards

Conflicts of interest: Authors declared that they have no conflict of interest.

Human participants: The conducted research follows the ethical standards and the authors ensured that they have not conducted any studies with human participants or animals.

References

[1] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., & Wicke, M. (2016). TensorFlow: A system for large-scale machine learning. Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, 265–283. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

[2] Abderrahmane, H., Noubeil, G., Lahcene, Z., Akhtar, Z., & Dasgupta, D. (2020). Weighted quasi-arithmetic mean based score level fusion for multi-biometric systems. IET Biometrics, 9(3), 91–99. https://doi.org/10.1049/iet-bmt.2018.5265

[3] Abdul-Al, M., Kyeremeh, G. K., Rami Qahwaji, Ali, N. T., & Abd-Alhameed, R. A. (2024). The evolution of biometric authentication: A deep dive into multi-modal facial recognition: A review case study. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3486552

[4] Abdullahi, S. M., Sun, S., Wang, B., Wei, N., & Wang, H. (2024). Biometric template attacks and recent protection mechanisms: A survey. Information Fusion, 103, Article 102144. https://doi.org/10.1016/j.inffus.2023.102144

[5] Akhtar, Z., Micheloni, C., & Foresti, G. L. (2015). Biometric liveness detection: Challenges and research opportunities. IEEE Security & Privacy, 13(5), 63–72. https://doi.org/10.1109/MSP.2015.116

[6] Alexander, D., & Ghozi, W. (2025). Performance analysis of deep learning model quantization on NPU for real-time automatic license plate recognition implementation. Journal of Applied Informatics and Computing, 9(4), 1227–1233. https://doi.org/10.30871/jaic.v9i4.9700

[7] Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. Electronics, 12(6), Article 1333. https://doi.org/10.3390/electronics12061333

[8] Ayeni, J. A. (2022). Convolutional neural network (CNN): The architecture and applications. Applied Journal of Physical Science, 4(4), 42–50. https://doi.org/10.31248/ajps2022.085

[9] Baobaid, A., Meribout, M., Tiwari, V. K., & Pena, J. P. (2022). Hardware accelerators for real-time face recognition: A survey. IEEE Access, 10, 83723–83739. https://doi.org/10.1109/ACCESS.2022.3194915

[10] Bazarevsky, V., Kartynnik, Y., Vakunov, A., Raveendran, K., & Grundmann, M. (2019). BlazeFace: Sub-millisecond neural face detection on mobile GPUs. arXiv. https://arxiv.org/abs/1907.05047

[11] Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

[12] Ben-Sasson, E., Chiesa, A., Tromer, E., & Virza, E. (2014). Succinct non-interactive zero knowledge for a von Neumann architecture. Proceedings of the 23rd USENIX Security Symposium. https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/ben-sasson

[13] Bhavani, S. A., & Karthikeyan, C. (2025). An attention based deep learning with effective SVM-ConvFaceNeXt model for face recognition in unconstrained environment. Signal, Image and Video Processing, 19(12). https://doi.org/10.1007/s11760-025-04441-9

[14] Bromley, J., Bentz, J. W., Bottou, L., Guyon, I., LeCun, Y., Moore, C., Säckinger, E., & Shah, R. (1993). Signature verification using a "Siamese" time delay neural network. International Journal of Pattern Recognition and Artificial Intelligence, 7(4), 669–688. https://doi.org/10.1142/S0218001493000339

[15] Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of Machine Learning Research, 81, 77–91. https://proceedings.mlr.press/v81/buolamwini18a.html

[16] Canny, J. (1986). A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6), 679–698. https://doi.org/10.1109/TPAMI.1986.4767851

[17] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv. https://arxiv.org/abs/1406.1078
[18] Chopra, S., Hadsell, R., & LeCun, Y. (2005). Learning a similarity metric discriminatively, with application to face verification. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05). https://doi.org/10.1109/CVPR.2005.202
[19] Crosby, M., Pattanayak, P., Verma, S., & Kalyanaraman, V. (2016). Blockchain technology: Beyond Bitcoin. Applied Innovation Review, (2), 6–10. https://www.appliedinnovationinstitute.org/blockchain-technology-beyond-bitcoin/
[20] Daugman, J. (2004). How iris recognition works. IEEE Transactions on Circuits and Systems for Video Technology, 14(1), 21–30. https://doi.org/10.1109/TCSVT.2003.818350
[21] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008
[22] De Leon, J. J., Zhang, C., Koulouris, C.-S., Medda, F., & Rahul. (2025). Smart contract security in decentralized finance: Enhancing vulnerability detection with reinforcement learning. Applied Sciences, 15(11), Article 5924. https://doi.org/10.3390/app15115924
[23] Deane, J. K., Ragsdale, C. T., Rakes, T. R., & Rees, L. P. (2009). Managing supply chain risk and disruption from IT security incidents. Operations Management Research, 2(1–4), 4–12. https://doi.org/10.1007/s12063-009-0018-2
[24] Deep learning with Python (2nd ed.). (n.d.). Manning Publications. https://www.manning.com/books/deep-learning-with-python-second-edition
[25] Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). ArcFace: Additive angular margin loss for deep face recognition. arXiv. https://arxiv.org/abs/1801.07698
[26] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An image is worth 16 × 16 words: Transformers for image recognition at scale. arXiv. https://arxiv.org/abs/2010.11929
[27] Mashruwala, A. (2024, May 30). The impact of cloud computing on FinTech: Applications and challenges. https://doi.org/10.13140/RG.2.2.15292.45440
[28] Satybaldy, A., Nowostawski, M., & Ellingsen, J. (2020). Self-sovereign identity systems. In Privacy and identity management. Data for better living: AI and privacy (pp. 447–461). Springer. https://doi.org/10.1007/978-3-030-42504-3_28

Appendix:

Algorithm 1: Serial Biometric Wallet Authentication Lifecycle

Input: Live camera frames, stored facial template, authentication threshold, encrypted wallet data, and a Keystore-protected cryptographic key.
Output: Authorized wallet access or access denial.

```
BEGIN PROCESS Authentication_Loop

WHILE (User_Not_Authenticated):

    Frame_Raw = CameraX.capture()
    Bounding_Box = FaceDetector.detect(Frame_Raw)

    IF (Bounding_Box is NULL):
        Continue

    Bitmap_128 = Resize(
        Crop(Frame_Raw, Bounding_Box), 128, 128
    )

    Tensor_Input = Normalize(
        Bitmap_128, range=[-1, 1]
    )

    V_live = M_CNN.predict(Tensor_Input)

    Distance (d) = Calculate_Euclidean_Distance(
        V_live, V_stored
    )

    IF (d > 0.55):
```

```
        Display "Face Mismatch (Access Denied)"
        Log_Failed_Attempt()

    ELSE:
        CameraX.stop()
        Display "Face Verified. Engaging Fingerprint Sensor..."

        Cipher_Obj = Keystore.Init_Cipher(
            AES_256_GCM, Mode=DECRYPT
        )

        Result = BiometricPrompt.authenticate(Cipher_Obj)

        IF (Result == SUCCESS):
            Decrypted_Ledger =
                Cipher_Obj.doFinal(Read("wallet_data.enc"))

            UI.Display(
                Decrypted_Ledger.Status("UNLOCKED")
            )

            RETURN

        ELSE:
            Log_Failed_Attempt()

    END IF

END WHILE

END PROCESS
```