Indonesian Journal of Electrical Engineering and Computer Science
Vol. 41, No. 2, February 2026, pp. 680~689
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v41.i2.pp680-689

Robust palmprint biometric solution for secure mobile authentication

Son Nguyen, Arthorn Luangsodsai, Pattarasinee Bhattarakosol
Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

Article Info

Article history:

Received Sep 4, 2025
Revised Jan 3, 2026
Accepted Jan 11, 2026

Keywords:

Biometric security
Cloud computing
Contrastive learning
Mobile security
Palmprint authentication
Self-supervised learning
Vector database

ABSTRACT

Smartphones increasingly rely on biometric authentication for access to financial and personal services, creating a need for palmprint recognition that is accurate, fast, and deployable on device. This paper proposes an end-to-end smartphone palmprint authentication framework that integrates guided mobile image capture, landmark-based region-of-interest (ROI) extraction, and compact embedding inference. A ResNet-18 teacher is first trained with self-supervised contrastive learning to reduce dependence on labeled biometric data, then distilled into a lightweight MobileNetV3 student for efficient mobile deployment. The learned embeddings support both on-device verification and large-scale identification using an approximate nearest neighbor index (FAISS). Experiments on a public Kaggle palm dataset achieve 99.2% accuracy with a 0.15% equal error rate (EER). On an iPhone 13, the end-to-end pipeline runs in 87.0 ms with a 12.4 MB student model. For a 1 million-entry gallery, FAISS provides 32 ms query latency while maintaining 99.5% Recall@1. Limitations include evaluation under mostly controlled capture conditions and the absence of an explicit liveness or presentation attack detection (PAD) module; future work will address unconstrained testing and anti-spoofing integration.

Corresponding Author:

Pattarasinee Bhattarakosol
Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University
Bangkok, Thailand
Email: pattarasinee.b@chula.ac.th

1. INTRODUCTION

Smartphones have become the primary gateway to digital services such as mobile banking, electronic payments, telemedicine, and remote work platforms. As a result, authentication mechanisms must achieve both strong security and acceptable user experience under tight on-device constraints. Biometric authentication is attractive because it can reduce reliance on passwords and mitigate risks associated with credential theft; however, practical deployment also raises concerns about privacy, spoofing, and secure template storage. Within this context, palmprint recognition is a compelling modality because the palm contains dense and discriminative texture patterns (principal lines, wrinkles, and fine creases) that are relatively stable over time, and it can be captured in a contactless manner using commodity cameras. Recent surveys and reviews consistently report that modern palmprint systems especially deep-learning-based approaches have achieved substantial gains in recognition accuracy and robustness compared with earlier handcrafted pipelines, while also highlighting that realistic deployment remains challenging due to acquisition variability and operational constraints [1], [2].

From a technical perspective, two aspects strongly influence palmprint performance in real applications: reliable region-of-interest (ROI) extraction and discriminative feature learning. In unconstrained contactless capture, changes in hand pose, distance to camera, illumination, and partial occlusion can degrade

ROI consistency and downstream recognition. Recent research has therefore emphasized complete and stable ROI strategies for unconstrained palmprint recognition and demonstrated measurable robustness improvements under variable acquisition conditions [3]. In parallel, deep models with feature fusion or gating mechanisms have been introduced to enhance discriminative representation learning, particularly when intra-class variation is non-trivial [4]. Nonetheless, many high-performing academic models still assume well-controlled capture and do not explicitly report end-to-end system behavior for smartphone deployment. For mobile settings, efficiency-oriented palmprint designs have been explored, such as compact coding or hashing-based representations that reduce matching cost and storage overhead while maintaining competitive verification performance [5]. Classical coding approaches remain relevant as lightweight baselines and as a reference point for understanding what performance is achievable with low computational complexity [6]. In addition, interpretability is increasingly important for biometric systems used in security-sensitive applications; explainable palmprint recognition pipelines have been studied to provide human-understandable evidence about which regions or patterns contribute most to the decision [7].

Beyond recognition accuracy, real-world palmprint authentication must also address emerging security threats and privacy constraints. Generative modeling has been actively studied for palm imagery, including the use of GAN-based pipelines in palmprint recognition research and analysis [8], as well as the explicit problem of detecting synthetic or manipulated palm images [9]. Recent work has also shown that diffusion models can generate realistic contactless palmprints, which further motivates the need for systematic anti-spoofing evaluation when deploying palmprint authentication outside controlled environments [10]. More broadly, biometrics are known to be vulnerable to adversarial manipulation and presentation attacks, and recent studies on combined attacks against recognition and presentation attack detection (PAD) emphasize that strong verification accuracy alone is insufficient to guarantee security [11]. At the template level, once biometric templates are compromised, they cannot be "reset" like passwords; cancellable template generation is therefore an important direction to reduce the impact of template leakage and enable revocation and re-issuance [12]. In addition, since biometric data is sensitive, privacy-preserving training and updating paradigms such as federated learning have been proposed for palm verification to reduce the need to centralize raw data during model improvement cycles [13].

These considerations become more critical on mobile and edge platforms. Edge intelligence research emphasizes that the operational constraints of on-device AI latency, memory footprint, energy usage, and reliability under heterogeneous hardware often dominate system feasibility and user acceptance [14]. Therefore, mobile palmprint authentication requires a complete workflow that includes stable acquisition, fast preprocessing, and efficient inference, rather than an accuracy-only evaluation. Practical mobile vision deployment has benefited from real-time perception frameworks that support fast landmark detection and streamlined on-device pipelines, enabling robust ROI processing under interactive camera capture conditions [15]. At the same time, collecting large-scale labeled biometric datasets is expensive and raises ethical and legal challenges. Self-supervised learning (SSL) offers a promising approach to reduce label dependence by learning view-invariant representations from unlabeled data; contrastive learning frameworks such as SimCLR provide a widely adopted foundation for this paradigm [16]. SSL has also demonstrated value in biometric-adjacent domains such as face representation learning, motivating its use as a data-efficient pretraining strategy when labels are limited or costly [17]. Similar SSL principles have been applied in other structured recognition problems, supporting the broader claim that representation quality can be improved by learning from context and invariances rather than explicit labels alone [18].

In addition to data efficiency, mobile feasibility requires compact neural architectures and compression techniques. Mobile-optimized backbones (e.g., MobileNet-family models) have been designed explicitly to improve the accuracy–latency trade-off on smartphones and embedded devices [19]. Lightweight network design strategies that generate more features from inexpensive operations further strengthen this direction, providing an additional option for improving efficiency without prohibitive accuracy loss [20]. Knowledge distillation is a practical compression mechanism for transferring representational capability from a larger teacher to a smaller student model, often preserving verification performance while reducing model size and inference time [21]. Quantization and integer-only inference methods can further reduce runtime and memory costs, but must be applied carefully to avoid unacceptable accuracy degradation in biometric verification settings [22].

Finally, mobile palmprint systems must support both verification (1:1) and identification (1:N). While many academic studies focus on verification, real deployments may require searching a large enrollment database (e.g., campus access, enterprise identity, or multi-user device scenarios). At scale, naïve exhaustive matching becomes impractical, motivating approximate nearest neighbor (ANN) indexing. FAISS provides a practical and widely used toolkit for similarity search and indexing that supports high-throughput retrieval pipelines [23]. GPU-accelerated similarity search has demonstrated that large-scale matching can be achieved efficiently, enabling billion-scale retrieval under suitable indexing and hardware configurations [24]. Recent ANN research continues to study the trade-offs between recall, latency, and memory, which are

critical to system design when the gallery size grows significantly [25]. Related work in other smartphone biometrics, such as visible-light iris recognition, further confirms that high-accuracy biometric authentication can be feasible on consumer devices when system design is tightly aligned with mobile constraints [26].

2. THE PROPOSED METHOD

This section presents the proposed end-to-end palmprint authentication framework, designed to jointly address three practical deployment requirements: (i) robust palm representation learning with reduced reliance on labeled biometric data, (ii) efficient on-device inference for real-time smartphone use, and (iii) scalable matching that supports both 1:1 verification and 1:N identification. The overall workflow is summarized in Figure 1, and the mobile acquisition interfaces and ROI localization outputs are shown in Figure 2 and Figure 3, respectively.

2.1. System workflow overview

The system follows a mobile-first pipeline, as depicted in Figure 1. During enrollment, the user captures palm images using the smartphone camera with in-app guidance; the captured frames are preprocessed on-device to extract a standardized palm ROI, and a compact CNN generates a fixed-length embedding that is stored as the user template. During authentication, the same preprocessing and embedding steps are applied to the query sample. The query embedding is then matched either locally (1:1 verification) or sent to a cloud backend for large-scale retrieval (1:N identification). Offline, the feature extractor is trained in two stages: a self-supervised teacher model is first learned from unlabeled palm ROIs and then distilled into a lightweight student model suitable for deployment on resource-constrained devices [14].

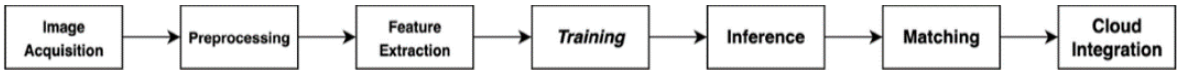


Figure 1. Workflow pipeline

2.2. Mobile image acquisition and real-time user guidance

The proposed system acquires contactless palm images using standard smartphone cameras. To standardize capture quality and reduce user-induced variation, the mobile application provides real-time feedback before allowing capture. Figure 2 illustrates three interface states used in this work: Figure 2(a) indicates that the palm is stable and correctly positioned for capture; Figure 2(b) warns that the palm ROI is too small (typically when the user is too far from the camera), prompting the user to move closer; and Figure 2(c) indicates that no palm has been detected in the current view. This guidance mechanism improves ROI consistency and reduces low-quality samples caused by motion blur, incorrect distance, or missing hands.

2.3. Preprocessing and ROI extraction

After capture, each RGB frame is standardized on device to produce a consistent palm ROI for feature extraction. The preprocessing sequence is designed to be lightweight while preserving discriminative palm line structures. First, the input is converted to grayscale and contrast is normalized using contrast-limited adaptive histogram equalization (CLAHE). A median filter is then applied to reduce sensor noise while maintaining edge and line details. ROI localization is performed using a lightweight hand landmark detector implemented in MediaPipe [15]. The detector returns 21 hand keypoints; a polygonal region is constructed from a subset of stable palm landmarks to isolate the central palm area. The cropped ROI is resized to 224×224 and intensity-normalized before being passed to the neural feature extractor. Figure 3 shows an example of the landmark output used to compute the ROI.

2.4. Stage 1: self-supervised representation learning (teacher model)

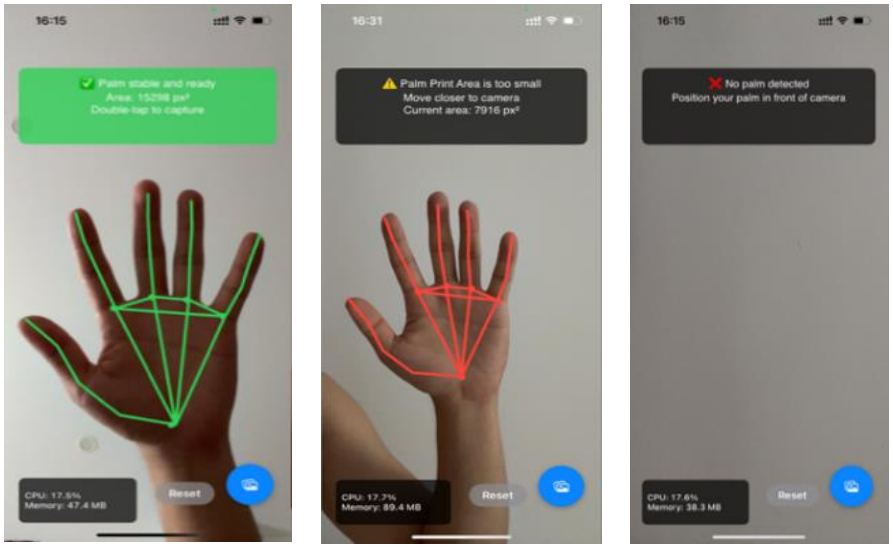
To reduce dependency on labeled biometric data, the framework trains a high-capacity teacher network using self-supervised contrastive learning. A ResNet-18 backbone is adopted as the teacher feature extractor, and each ROI sample is transformed into two augmented views (e.g., random crop, mild rotation, blur/noise) to form a positive pair. The objective encourages view-invariant representations by maximizing agreement between the two views while separating them from other samples in the batch. The training follows a SimCLR-style contrastive formulation using the normalized temperature-scaled cross-entropy (NT-Xent) loss [16]. To improve representation quality for contrastive learning, the loss is applied in a

projection space produced by a small MLP head attached to the backbone, consistent with standard SimCLR practice.

The training is guided by the NT-Xent loss function (1), which is central to the SimCLR framework [16]. The objective of this loss is to maximize the similarity between the embeddings of the positive pair while simultaneously minimizing their similarity to all negative pairs. The NT-Xent loss for a positive pair (i, j) is formulated as:

$$L_{i,j} = -\log \frac{\exp(sim(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} 1_{[k \neq i]} \exp(sim(z_i, z_j)/\tau)} \quad (1)$$

where z_i and z_j are the embedding vectors of the positive pair, $sim(\cdot)$ is the cosine similarity function, τ is a temperature parameter that controls the separation of negative samples, and N is the batch size. A lower temperature increases the penalty for hard-to-distinguish negative samples, forcing the model to learn more discriminative features. During this stage, a two-layer MLP projection head is added to the ResNet-18 encoder.



(a) (b) (c)

Figure 2. Examples of real-time user guidance in the mobile application. From left to right (a) Successful detection with a green overlay indicating the palm is stable and ready (b) A warning that the palm area is too small, prompting the user to move closer and (c) A notification that no palm has been detected in the camera's view

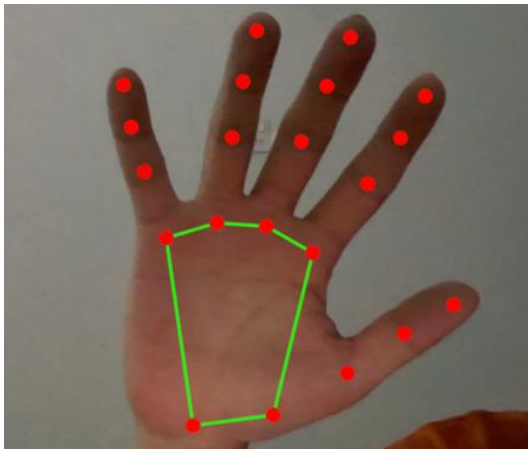


Figure 3. Palm landmark detection for ROI extraction

2.5. Stage 2: knowledge distillation for mobile deployment (student model)

After training the teacher network, its representation is transferred to a compact student model to enable low-latency inference on smartphones. A MobileNetV3 backbone is adopted as the student due to its favorable accuracy–latency trade-off on mobile hardware [19]. The student is trained to mimic the teacher embedding for the same ROI input using a distillation objective that aligns the student's embedding space with the teacher's embedding space. This approach is consistent with common deployment practice where distillation is used to preserve recognition capability while reducing model size and inference cost [21]. In addition, lightweight architectural principles (e.g., generating more features using inexpensive operations) can be incorporated to further improve mobile efficiency when needed [20]. If additional compression is required, integer-only quantization can be applied after training; however, quantization should be validated carefully because biometric verification accuracy can be sensitive to embedding perturbations [22].

2.6. Hybrid matching for verification and scalable identification

The deployed student model outputs a fixed-dimensional embedding (256-D in this work) for each ROI. This embedding size is selected as a balance between discriminative power and on-device efficiency. For local authentication, the system performs 1:1 verification by computing cosine similarity between the query embedding and the enrolled template, followed by a threshold decision. For applications requiring identification among many enrolled users, the query embedding is transmitted to a cloud backend that performs ANN search over the embedding database. Because exhaustive search becomes inefficient for large galleries, the backend uses FAISS indexing to support fast similarity search at scale [23]. The design follows established ANN practice for high-dimensional retrieval, where recall–latency trade-offs are tuned to meet application constraints [25]. For very large deployments, GPU-accelerated similarity search can be used to improve throughput when appropriate infrastructure is available [24].

3. METHOD

This section describes the experimental design used to evaluate the proposed framework in a reproducible manner, including the dataset usage constraint, training configuration, evaluation metrics, and the measurement protocol for mobile latency and large-scale identification.

3.1. Dataset and experimental protocol

All experiments in this study were conducted using only the public Kaggle dataset "Palm Recognition Dataset for Authentication System" (Palm Dataset). The dataset contains approximately 12,000 palm images and is distributed primarily in TIFF format. Because public biometric datasets often lack a universally enforced protocol, we adopt a consistent enrollment–probe evaluation strategy that matches real authentication usage. Each identity is represented by multiple samples. One sample per identity is selected as the enrollment template, while the remaining samples are treated as probe attempts. Genuine comparisons are formed by matching each probe to its enrolled template of the same identity, and impostor comparisons are formed by matching probes against enrolled templates from other identities. This produces score distributions required for FAR/FRR analysis and EER computation. The same identity grouping is also used for reporting identification-style "accuracy" (top-1 match of a probe against the enrolled gallery) to keep the reported accuracy and EER logically consistent within a single protocol.

3.2. Implementation details

All palm images were processed using the same on-device preprocessing and ROI extraction pipeline described in Section 2. Each input frame was converted to grayscale, enhanced using CLAHE, and denoised using a median filter. A lightweight MediaPipe hand landmark detector was then applied to obtain keypoints for constructing a polygonal palm region, which was cropped and resized to 224×224 before feature extraction [15]. For representation learning, a ResNet-18 teacher model was trained using a SimCLR-style contrastive objective with NT-Xent loss, where two augmented views were generated per ROI and optimized to maximize agreement for positive pairs while separating negatives within the batch [16]. For mobile deployment, the learned teacher representation was transferred to a lightweight MobileNetV3 student network [19] using knowledge distillation by aligning student embeddings to teacher embeddings for the same ROI input (embedding-alignment loss) [21]. Unless otherwise stated, the student embedding dimension was fixed at 256 for verification and identification experiments.

3.3. Evaluation metrics

Verification performance is quantified using (i) authentication accuracy under the enrollment–probe identification protocol and (ii) EER. EER is the operating point where the false acceptance rate (FAR) equals

the false rejection rate (FRR), and it is widely used in biometric verification studies [1]. For completeness, FAR and FRR are computed by sweeping the similarity threshold over the genuine and impostor score distributions; EER is the threshold where FAR and FRR intersect.

For scalability, 1:N identification performance is evaluated using Recall@1, defined as the fraction of probes whose correct enrolled identity appears as the top-ranked retrieval result. This aligns with standard identification evaluation practice and supports a practical interpretation of system behavior at scale [23], [25].

3.4. Mobile latency and large-scale identification measurement

To evaluate deployability, on-device performance is measured on an Apple iPhone 13 (A15 Bionic chipset, 4 GB RAM), consistent with the deployment target described in the Results section. Latency is reported as end-to-end runtime (preprocessing + ROI extraction + model inference + similarity matching), and component-level timing is also recorded to attribute the total cost to each stage.

To evaluate large-scale identification, we construct a simulated gallery of one million embeddings and compare exhaustive search against an ANN index. FAISS is used as the vector search backend, and the HNSW index is selected due to its practical speed–recall trade-off in medium-to-large scale settings [23]. The reported recall/latency trade-off is interpreted as an operational design choice, where a small loss in retrieval accuracy can yield large improvements in query time, consistent with established ANN behavior [25].

4. RESULTS AND DISCUSSION

This section evaluates the proposed end-to-end palmprint authentication framework in terms of (i) verification performance (accuracy and EER), (ii) mobile deployability (end-to-end latency and model footprint), (iii) robustness under practical capture variations, and (iv) scalability for large-gallery identification.

4.1. Authentication accuracy and error rates

Table 1 reports verification performance on the Kaggle palm dataset using the same enrollment–probe protocol for both the supervised MobileNetV3 baseline and the proposed framework. The baseline achieves 97.5% accuracy with 1.80% EER, while the proposed self-supervised pretraining plus distilled MobileNetV3 improves accuracy to 99.2% and reduces EER to 0.15%. This corresponds to a +1.7 percentage-point gain in accuracy and a 1.65 percentage-point absolute reduction in EER, indicating that the learned embedding becomes more discriminative and stable while remaining suitable for mobile deployment.

For broader context, Table 1 also includes representative results reported in recent palmprint literature GLGANet [4] and a CCNet result summarized in the palmprint deep-learning survey [1]. These literature values are listed together with their original datasets/protocols (e.g., Tongji/PolyU-style benchmarks) and are provided for positioning rather than direct numerical comparison, because palmprint performance is strongly affected by dataset characteristics, capture conditions, and evaluation protocols [1], [2]. Within this framing, the contribution of this work is not limited to recognition accuracy; it is the demonstration of an end-to-end, smartphone-oriented pipeline that couples label-efficient representation learning and a compact deployed model with measurable on-device latency and a scalable identification backend, while maintaining verification performance within the high-accuracy regime reported by recent studies [1], [4].

Table 1. Comparison of authentication accuracy and EER with state-of-the-art methods

Method	Training	Dataset	Accuracy (%)	EER (%)
Supervised MobileNetV3 (Baseline)	Supervised	Kaggle	97.5	1.80
Proposed SSL + Distilled MobileNetV3	SSL + distillation	Kaggle	99.2	0.15
GLGANet [4]	Supervised	Tongji/PolyU	98.5/99.5	Not reported
CCNet [1]	Supervised	Tongji	100.0	0.00004

4.2. On-device performance and deployment relevance

A mobile biometric system must be accurate and responsive under smartphone constraints. Table 2 reports a full end-to-end latency of 87.0 ms on iPhone 13, including preprocessing/ROI extraction (48.0 ms), student inference (21.5 ms), and cosine matching (17.5 ms). Importantly, this measurement reflects the complete pipeline rather than model inference alone, which better represents real user experience in mobile authentication.

The distillation stage is the key factor enabling deployability: the student model reduces inference time by approximately 3.5× relative to the teacher (75.2 ms to 21.5 ms) and reduces the model footprint from

45.1 MB to 12.4 MB. This is consistent with the broader edge intelligence perspective that model compression and system-level latency are primary feasibility constraints for real-world edge AI, not only accuracy [14]. While some palmprint studies emphasize recognition performance [4] or interpretability [7], many do not report complete on-device timing and memory measurements, making it difficult to assess deployment readiness. By reporting end-to-end runtime with explicit module breakdown, this work contributes operational evidence that the proposed framework can be integrated into a practical mobile application.

Table 2. On-device performance evaluation on iPhone 13

Component	Latency (ms)	Model Size (MB)
Preprocessing Pipeline (incl. ROI Extraction)	48.0	-
ResNet-18 Teacher (Theoretical Inference)	75.2	45.1
MobileNetV3 Student (On-Device Inference)	21.5	12.4
Cosine Similarity Matching	17.5	-
Full End-to-End Authentication	87.0	-

4.3. Robustness to real-world conditions

Table 3 evaluates robustness under variations that frequently occur in smartphone capture: low light, partial occlusion, and mild rotation. The proposed SSL model consistently outperforms the supervised baseline under all stress conditions, with particularly notable gains under partial occlusion and low light. This behavior aligns with the motivation of contrastive learning, which learns invariances through augmentation and encourages representations that remain stable under perturbations that resemble real capture noise [16].

From a literature perspective, ROI quality is widely recognized as a dominant factor in unconstrained palmprint recognition, and recent ROI research explicitly targets robustness under less controlled acquisition [3]. In our system, robustness improvements arise from two coupled design choices: guided acquisition (Figure 2) reduces severe capture errors before inference, and landmark-based ROI extraction (Figure 3) provides consistent palm alignment at low runtime cost. While our ROI method is designed to be lightweight for mobile deployment, the results in Table 3 suggest that combining stable ROI extraction with SSL is a practical path to improved robustness without increasing on-device computation.

Table 3. Robustness evaluation under challenging conditions accuracy

Condition	Supervised MobileNetV3 (%)	Proposed SSL model (%)
Normal lighting	97.5	99.2
Low light	94.1	97.8
Partial occlusion	92.7	97.8
Rotation (±15 degrees)	95.2	98.5

4.4. Scalability for large-scale identification

Table 4 evaluates the system in a large-scale identification setting using a one million embedding gallery. Exhaustive search requires approximately 3540 ms per query, which is not feasible for interactive identification. In contrast, FAISS with an HNSW index reduces latency to approximately 32 ms per query while maintaining 99.5% Recall@1. This is consistent with the fundamental ANN trade-off: a small and tunable reduction in exactness can yield orders-of-magnitude speedups for high-dimensional retrieval [23], [25]. Prior work also shows that similarity search throughput can be further improved with GPU acceleration when infrastructure permits, reinforcing that scalable identification is a tractable component of real deployments [24].

This scalability result is a key part of the contribution because many palmprint papers focus primarily on verification performance without demonstrating system behavior when the gallery size increases substantially. By integrating ANN-based retrieval into the architecture and quantifying latency/recall at a realistic scale, the proposed framework moves beyond accuracy-only evaluation toward deployment-grade identification capability.

Table 4. Scalability performance for 1:N Identification (1 million gallery size)

Search method	Search latency (ms per query)	Recall@1 (%)
Brute-force (exhaustive)	~3540	100.0
FAISS (HNSW Index)	~32	99.5

4.5. Discussion, limitations, and future work

Taken together, Tables 1-4 demonstrate that the framework achieves a balanced profile across accuracy, robustness, mobile efficiency, and scalability. The primary contribution relative to representative palmprint studies is not merely a marginal metric gain, but an end-to-end design that integrates (i) label-efficient training via SSL [16], (ii) mobile-friendly inference via distillation (Table 2), and (iii) scalable identification via ANN indexing [23], [25]. This integration is rarely reported as a single validated pipeline in prior palmprint work, which tends to emphasize either recognition modeling [4], ROI robustness [3], or efficiency-oriented representations [5].

Nevertheless, limitations remain. First, the Kaggle dataset reflects mostly controlled or semi-controlled acquisition, so generalization to fully unconstrained outdoor settings and cross-device variability is not yet established; ROI studies indicate this remains a core challenge in unconstrained palmprint recognition [3]. Second, the current system does not include an explicit liveness or PAD component. This is important because recent works demonstrate both the feasibility of detecting deepfake palm imagery [9] and the increasing realism of generated palmprints using diffusion models [10], while broader biometric security research shows that combined attacks can target both recognition and PAD modules [11]. Third, template protection is not yet implemented; cancellable templates provide a concrete direction for revocation and re-issuance if templates are compromised [12]. Finally, privacy-preserving model updates are not yet explored in this implementation; federated learning is a relevant pathway for reducing centralized raw-biometric exposure during updates [13].

5. CONCLUSION

This paper presented a smartphone-oriented palmprint authentication framework that integrates guided mobile acquisition, landmark-based ROI extraction, compact embedding inference, and hybrid matching for both on-device verification and large-scale identification. The results show that combining self-supervised contrastive pretraining with knowledge distillation can yield a strong palm representation while meeting practical mobile constraints, and that ANN-based retrieval using FAISS provides a feasible path to low-latency identification when the gallery size becomes large.

Several limitations should be acknowledged. First, the experimental evaluation relies on a public dataset that reflects mostly controlled or semi-controlled capture conditions; therefore, the reported performance may not fully represent more challenging real-world scenarios such as outdoor lighting, severe motion blur, large pose variation, and cross-device camera differences. Second, the current system does not include an explicit liveness or PAD module, which is important for security because synthetic or manipulated palm imagery is becoming increasingly realistic. Third, template protection and revocability are not yet implemented, which is a relevant consideration for any system that stores embeddings for authentication or identification.

Future work will therefore focus on (1) broader validation under unconstrained and cross-device capture settings with a consistent protocol, (2) integrating a lightweight anti-spoofing component and quantifying its security–latency trade-off, and (3) strengthening template security through cancellable templates and privacy-preserving model update strategies, while also exploring additional compression (e.g., quantization) under careful verification stability checks.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. Bello Musa Yakubu for his valuable time, insightful comments, and constructive feedback in reviewing this journal. His guidance greatly contributed to improving the clarity and quality of this work.

FUNDING INFORMATION

This research was supported by a scholarship from the International Students Scholarship, Faculty of Science, Chulalongkorn University, Bangkok, Thailand.

AUTHOR CONTRIBUTIONS STATEMENT

The specific contributions of each author to this research are outlined in the following table, following the CRediT (Contributor Roles Taxonomy) framework.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Son Nguyen	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓			
Arthorn Luangsodsai				✓		✓	✓					✓	✓	✓
Pattarasinee Bhattarakosol	✓	✓		✓			✓			✓	✓	✓	✓	✓

C : Conceptualization
M : Methodology
So : Software
Va : Validation
Fo : Formal analysis
I : Investigation
R : Resources
D : Data Curation
O : Writing - Original Draft
E : Writing - Review & Editing
Vi : Visualization
Su : Supervision
P : Project administration
Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Palm Recognition Dataset for Authentication System at https://www.kaggle.com/datasets/saqibshoaibdz/palm-dataset/data.

REFERENCES

[1] C. Gao, Z. Yang, W. Jia, L. Leng, B. Zhang, and A. B. J. Teoh, "Deep learning in palmprint recognition-a comprehensive survey," Oct. 2025, [Online]. Available: http://arxiv.org/abs/2501.01166

[2] S. Zhao, L. Fei, and J. Wen, "Multiview-learning-based generic palmprint recognition: a literature review," *Mathematics*, vol. 11, no. 5, p. 1261, Mar. 2023, doi: 10.3390/math11051261.

[3] L. Su, L. Fei, B. Zhang, S. Zhao, J. Wen, and Y. Xu, "Complete region of interest for unconstrained palmprint recognition," *IEEE Transactions on Image Processing*, vol. 33, pp. 3662–3675, 2024, doi: 10.1109/TIP.2024.3407666.

[4] K. Zhang, G. Xu, Y. K. Jin, G. Qi, X. Yang, and L. Bai, "Palmprint recognition based on gating mechanism and adaptive feature fusion," *Frontiers in Neurorobotics*, vol. 17, May 2023, doi: 10.3389/fnbot.2023.1203962.

[5] H. Shao, D. Zhong, and X. Du, "Efficient deep palmprint recognition via distilled hashing coding," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, IEEE, Jun. 2019, pp. 714–723, doi: 10.1109/CVPRW.2019.00098.

[6] H. Wang and X. Cao, "Palmprint recognition using bifurcation line direction coding," *IEEE Access*, vol. 13, pp. 70366–70377, 2025, doi: 10.1109/ACCESS.2025.3562648.

[7] A. Bhandary, C. Bhatt, and R. Jain, "An interpretable palmprint recognition approach using a CNN and explainable-AI frameworks," in *Proceedings of the 2024 International Conference on Artificial Intelligence and Emerging Technology, Global AI Summit 2024*, IEEE, Sep. 2024, pp. 329–334, doi: 10.1109/GlobalAISummit62156.2024.10947987.

[8] S. Sinha and S. B. Verma, "Harnessing deep learning and GAN technologies for palmprint recognition," in *2024 International Conference on Cybernation and Computation, CYBERCOM 2024*, IEEE, Nov. 2024, pp. 68–76, doi: 10.1109/CYBERCOM63683.2024.10803245.

[9] T. Min-Jen and C. Cheng-Tao, "Convolutional neural network for detecting deepfake palmprint images," *IEEE Access*, vol. 12, pp. 103405–103418, 2024, doi: 10.1109/ACCESS.2024.3433497.

[10] S. A. Grosz and A. K. Jain, "GenPalm: contactless palmprint generation with diffusion models," in *Proceedings - 2024 IEEE International Joint Conference on Biometrics, IJCB 2024*, IEEE, Sep. 2024, pp. 1–9, doi: 10.1109/IJCB62174.2024.10744493.

[11] S. Zedda, S. M. La Cava, S. Carta, R. Casula, and G. L. Marcialis, "Combined adversarial attacks for breaking fingerprint recognition and presentation attack detection," in *Proc. 2025 33rd European Signal Processing Conf. (EUSIPCO)*, 2025, pp. 1–5.

[12] M. Imran, M. S. Umar, and S. Malhotra, "Privacy preserving cancellable template generation for crypto-biometric authentication system," *IEEE Access*, vol. 13, pp. 158322–158339, 2025, doi: 10.1109/ACCESS.2025.3602795.

[13] Z. Yang, A. B. J. Teoh, B. Zhang, L. Leng, and Y. Zhang, "Physics-driven spectrum-consistent federated learning for palmprint verification," *International Journal of Computer Vision*, vol. 132, no. 10, pp. 4253–4268, Oct. 2024, doi: 10.1007/s11263-024-02077-9.

[14] S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar, and A. Y. Zomaya, "Edge intelligence: the confluence of edge computing and artificial intelligence," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7457–7469, Aug. 2020, doi: 10.1109/JIOT.2020.2984887.

[15] C. Lugaresi *et al.*, "MediaPipe: a framework for building perception pipelines," Jun. 2019, [Online]. Available: http://arxiv.org/abs/1906.08172

[16] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Proceedings of Machine Learning Research*, 2020, pp. 1597–1607.

[17] Z. Gao and I. Patras, "Self-supervised facial representation learning with facial region awareness," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2024, pp. 2081–2092, doi: 10.1109/CVPR52733.2024.00203.

[18] W. Zhou, L. Kong, Y. Han, J. Qin, and Z. Sun, "Contextualized relation predictive model for self-supervised group activity representation learning," *IEEE Transactions on Multimedia*, vol. 26, pp. 353–366, 2024, doi: 10.1109/TMM.2023.3265280.

[19] A. Howard *et al.*, "Searching for mobileNetV3," in *Proceedings of the IEEE International Conference on Computer Vision*, IEEE, Oct. 2019, pp. 1314–1324, doi: 10.1109/ICCV.2019.00140.

[20] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: more features from cheap operations," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2020, pp. 1577–1586, doi: 10.1109/CVPR42600.2020.00165.
[21] N. İbrahimoğu, M. C. Aytekin, and F. Yıldız, "Knowledge distillation from resnet to mobilenet for accurate on-device face recognition," *AIPA's International Journal on Artificial Intelligence: Bridging Technology, Society and Policy*, vol. 1, no. 2, pp. 1–15, 2025, doi: 10.5281/zenodo.15502091.
[22] B. Jacob *et al.*, "Quantization and training of neural networks for efficient integer-arithmetic-only inference," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 2704–2713, doi: 10.1109/CVPR.2018.00286.
[23] M. Douze *et al.*, "The faiss library," *IEEE Transactions on Big Data*, Oct. 2025, doi: 10.1109/TBDATA.2025.3618474.
[24] J. Johnson, M. Douze, and H. Jegou, "Billion-scale similarity search with GPUs," *IEEE Transactions on Big Data*, vol. 7, no. 3, pp. 535–547, Jul. 2021, doi: 10.1109/TBDATA.2019.2921572.
[25] H. Zhibing, "Quick and efficient large scale approximate nearest neighbor search on high-dimensional data," in *International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP*, IEEE, Dec. 2024, pp. 1–5, doi: 10.1109/ICCWAMTIP64812.2024.10873693.
[26] K. B. Raja, R. Raghavendra, V. K. Vemuri, and C. Busch, "Smartphone based visible iris recognition using deep sparse filtering," *Pattern Recognition Letters*, vol. 57, pp. 33–42, May 2015, doi: 10.1016/j.patrec.2014.09.006.

BIOGRAPHIES OF AUTHORS

Son Nguyen obtained B.Sc. in Computer Science from Texas A&M University, Corpus Christi, Texas, USA. He is pursing master's degree in computer science at Chulalongkorn University, Bangkok,Thailand. His research focues on security and software development. He can be contacted at email: 6778014123@student.chula.ac.th

Professor Dr. Arthorn Luangsodsai obtained B.Sc. in Computer Science from Thammasat Univeristy,Thailand; M.Sc. Analysis, Design, and Management of Information Systems from London School of Economics, U.K.; M.Sc. and Ph.D. in Computer Science from University of Essex, U.K. He holds the position of Assistant Professor in the Department of Mathematics and Computer Science at the Faculty of Science, Chulalongkorn University, Thailand. His research focuses on software engineering, outlier detection and information systems management. He can be contacted at email: arthorn.l@chula.ac.th

Professor Dr. Pattarasinee Bhattarakosol obtained a B.Sc. in mathematics from Chulalongkorn University, Bangkok, Thailand; an M.Sc. in applied statistics from the National Institute of Development Administration, Bangkok, Thailand; and a Ph.D. in computer science from Wollongong University, Australia. She holds the position of associate professor in the Department of Mathematics and Computer Science at the Faculty of Science, Chulalongkorn University, Thailand. Her research focuses on computer networks, security, and software architecture. She can be contacted at email: pattarasinee.b@chula.ac.th