Received: 30 August 2025 | Revised: 18 March 2026 | Accepted: 18 June 2026 | Published online: 22 July 2026

RESEARCH ARTICLE

Artificial Intelligence and Applications
2026, Vol. 00(00) 1–14
DOI: 10.47852/bonviewAIA62027476

BON VIEW PUBLISHING

One-Class Anomaly Detection for Finger Vein Presentation Attack Detection

Mohd Shahrimie Mohd Asaari[1], Bakhtiar Affendi Rosdi[1,2,*], Andrew Tiong Hoe Pin[1], Zahid Ur Rahman[1] and Muhammad Firdaus Akbar[1,2]

[1] *School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Malaysia*

[2] *Visionlytics Sdn Bhd, Universiti Sains Malaysia, Malaysia*

Abstract: The reliability of finger vein biometric systems is increasingly threatened by sophisticated presentation attacks. Current presentation attack detection (PAD) methods, often relying on supervised learning, are vulnerable to novel, unseen attacks because they depend on comprehensive labeled spoof datasets that are impractical to collect. To address this zero-day threat, this research proposes a one-class anomaly detection framework trained solely on authentic finger vein samples. While utilizing a standard U-Net-inspired denoising autoencoder as the architectural backbone, this work introduces two key novel contributions to tailor the model for biometric security: (1) a Multi-Objective Anomaly Detection Loss Framework that uniquely integrates multi-scale reconstruction error, gradient preservation constraints, and deep Support Vector Data Description loss to strictly regularize the latent space and (2) a Spectral Error Optimization technique that applies adaptive frequency weighting to amplify subtle texture artifacts inherent in spoof mediums. This combination is significant because the multi-objective loss forces the model to learn fine-grained physiological vein patterns, while the spectral optimization captures high-frequency anomalies often missed by spatial reconstruction alone. Experimental results on the FVPAD-USM dataset demonstrate that this approach achieves an Attack Presentation Classification Error Rate below 2% and an Average Classification Error Rate below 10%. By outperforming supervised baselines like support vector machines and convolutional neural networks in generalizing to unseen digital and glossy photo attacks, this work establishes the effectiveness of unsupervised, frequency-enhanced anomaly detection for robust biometric security.

Keywords: finger vein recognition, presentation attack detection, one-class anomaly detection, denoising autoencoder, unsupervised learning

1. Introduction

Biometric identification systems are nowadays an integral part of the modern security forces, offering the highest degree of protection against unlawful infiltration. Yet, the systems remain vulnerable to sophisticated presentation attacks where attackers make artificial replications or spurious representations of biometric characteristics. Anti-spoofing measures have largely concentrated on the fingerprint [1, 2], face recognition [3], and speaker verification [4]; finger vein recognition, however, rarely finds interest regarding presentation attack detection (PAD), touted as highly secure because spoofs of subcutaneous vascular patterns are far more difficult [5]. The finger vein is not completely immune to spoofing; hence, the PAD systems must evolve new methods of keeping their reliability in real-world applications. The occluded nature of vein patterns and the need for specialized near-infrared imaging make finger vein recognition inherently resilient to spoofing, as also demonstrated by multimodal approaches that combine finger vein with other traits to boost the recognition accuracy and attack resistance [3, 6].

Legacy PAD techniques for finger vein authentication, such as those employing local binary patterns (LBP) and support vector machines (SVM), rely on handcrafted features to distinguish between genuine and spoofed samples. However, these conventional techniques are not good at generalizing across a wide variety of attacks, especially as spoofing methods become more prevalent and sophisticated. Attackers now use high-resolution printing, textured surfaces such as photo paper, and even 3D-printed replicas to convincingly fake vein patterns, posing significant difficulties for traditional algorithms [7]. These methods often fail to capture the latent artifacts or material properties of transient spoofing materials, failing in catching novel attacks. The fixed feature selection approach of these traditional methods prevents them from adjusting to new attack patterns, which results in systems that are unprotected against zero-day threats. For instance, research demonstrates that classic descriptors such as LBP cannot properly identify advanced spoof materials like dragon-skin, silicon, and glue-based overlays because these materials show minimal differences from authentic samples when observed through near-infrared imaging [8].

***Corresponding author:** Bakhtiar Affendi Rosdi, School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Malaysia. Email: eebakhtiar@usm.my

A fundamental challenge in developing finger vein PAD systems is the difficulty of acquiring diverse and comprehensive spoof samples. Collecting such data is impractical and time-consuming, as attack techniques evolve rapidly. This scarcity of spoofed data reduces the effectiveness of supervised learning models, which typically require balanced genuine and spoofed samples. Consequently, systems based on convolutional neural networks (CNNs) or ensemble networks often achieve high accuracy under known attack conditions but experience significant performance drops against unknown attacks [9–12]. Supervised PAD techniques rely on attack data to learn discriminative patterns, making them vulnerable to zero-day spoofing attacks and limiting their robustness in real-world deployment [13].

To address this issue, research has shifted toward open-set and anomaly-based approaches that do not require pre-collected spoof samples [2, 10]. These methods instead learn the intrinsic characteristics of genuine finger vein patterns and classify deviations as anomalies, offering spoof-agnostic detection capabilities. While supervised methods such as depthwise separable CNNs and SVM have shown strong performance in controlled settings [2, 9, 10, 13], they remain constrained by their reliance on labeled spoof data. As a result, they fail to generalize novel attacks involving high-resolution synthetic films, 3D-printed molds, or advanced texture-transfer techniques [1, 14], which are highly relevant in real-world scenarios where unseen attacks are more likely [15, 16].

Another critical limitation is the absence of comprehensive public finger vein spoofing datasets. Generating realistic attack samples requires sophisticated fabrication methods, specialized equipment, and domain expertise, which restricts dataset diversity and scale [11, 17, 18]. For example, a recent 2024 study [19] employed a hybrid gray-level co-occurrence matrix (GLCM) with a Light Gradient Boosting Machine (LightGBM) classifier and achieved encouraging Attack Presentation Classification Error Rate (APCER) results. However, its dataset contained only a limited range of spoof types, reducing generalizability. This lack of variety causes many supervised PAD systems to overfit specific spoof classes and struggle to sustain performance in diverse real-world environments [6, 20–22].

To overcome these challenges, recent studies have explored unsupervised and one-class anomaly detection methods that train exclusively on real samples [12, 23, 24]. These approaches learn the inherent distribution of authentic finger vein patterns and identify deviations caused by spoof artifacts, making them suitable for both known and unknown presentation attacks. One-class models, such as convolutional autoencoders, Clustered Deep One-Class Classification (CD-OCC), and Interpolated Gaussian Descriptors (IGDs), have shown promise in proactively detecting anomalies.

Building on this direction, the present work proposes a one-class anomaly detection model using a denoising autoencoder (DAE). The model learns from genuine finger vein samples and detects spoofing attempts across unknown attack types while maintaining user convenience by optimizing the Average Classification Error Rate (ACER). It also seeks to minimize the APCER while balancing the Bonafide Presentation Classification Error Rate (BPCER) to ensure robustness against novel spoofing materials. The proposed DAE framework integrates convolutional operations with residual learning and attention-inspired mechanisms to enhance spatial texture learning, offering a reliable PAD solution for practical deployment.

The primary contribution of this work is a novel one-class anomaly detection framework utilizing DAE trained exclusively on genuine finger vein samples, which eliminates the dependency on labeled spoof datasets and enables the detection of "zero-day" attacks. This is achieved through a specialized multi-objective loss framework integrating multi-scale reconstruction error, gradient preservation, and deep Support Vector Data Description (SVDD) to strictly regularize the latent space. Furthermore, a spectral error optimization technique is utilized, employing adaptive frequency weighting to amplify high-frequency artifacts, such as print textures that are often obscured in the spatial domain. The resulting method demonstrates superior generalization against unseen attacks on the FVPAD-USM dataset compared to supervised baselines, while maintaining a lightweight computational footprint suitable for real-time deployment.

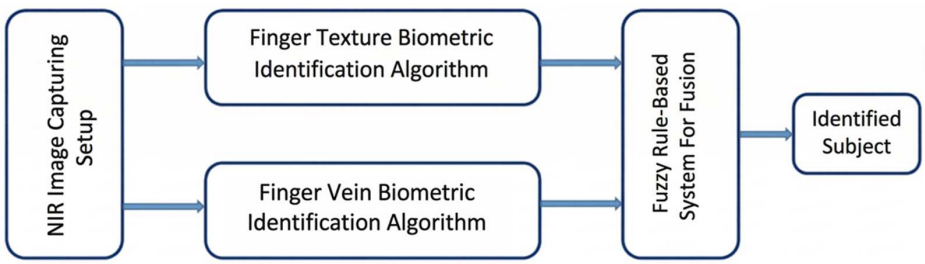
The remainder of this paper is organized as follows: Section 2 reviews related works, Section 3 describes the materials and methods, Section 4 presents results and discussion, and Section 5 concludes with implications and future work.

2. Related Work

Early research into PAD for finger vein systems relied on conventional machine learning techniques that distinguished authentic images from counterfeits through manually engineered feature descriptors. Among these, LBP became a widely used method due to its computational efficiency and robustness in capturing the subtle local texture variations that often differ between genuine and spoofed samples [6]. Similar approaches have been applied in fingerprint PAD, where orchestrated techniques combining multiple features have shown effectiveness [1].

A common approach combined these handcrafted features with powerful classifiers. For instance, Haider et al. [6] proposed a multimodal biometric system, illustrated in Figure 1 [6], that extracted LBP features and classified them using SVM with a score-level fuzzification mechanism. Their work demonstrated that combining texture patterns from multiple modalities

Figure 1
Block diagram of a multimodal biometric system



significantly improved PAD performance compared to single-modality systems. Similarly, other texture-based descriptors have proven effective. Shaheed et al. [8] utilized optimized GLCM features with a LightGBM classifier, achieving strong results against printed and display-based attacks. Schuiki et al. [11] conducted a thorough threat analysis and found that combining Gabor filters with LBP features also yielded reliable spoof detection across various attack media. These studies collectively affirm that well-designed texture descriptors are fundamental to conventional PAD.

Beyond individual feature-classifier pairs, researchers explored more advanced frameworks to enhance robustness. Foundational work in biometric recognition has demonstrated that fusing information from multiple sources significantly reduces error rates. For instance, multi-instance fusion strategies [25] have proven that combining feature sets from different fingers or instances can overcome the limitations of single-source biometrics. Similarly, sophisticated score-level fusion techniques [26] have been successfully employed to optimize decision boundaries in multimodal systems. Building on these fusion concepts, Kim et al. [10] developed a non-deep learning PAD method based on ensemble learning, which fused scores from multiple models, including SVMs, as shown in the architecture in Figure 2 [10]. This approach provided a stronger defense against various spoofing methods and performed well in cross-type evaluations. To further refine these pipelines, dimensionality reduction techniques such as principal component analysis (PCA) were employed. Al-Obaidy et al. [17] showed that applying PCA to handcrafted features before SVM classification improved reliability, particularly in low-data scenarios, by reducing noise and improving generalization. However, a critical limitation shared by these conventional methods is their reliance on known attack types. The study by Shaheed et al. [9], for example, was constrained by a limited number of spoof samples, which exposed the framework's potential inability to handle unknown attacks. This dependency on predefined spoof characteristics remains a significant weakness in traditional PAD systems.

Early finger vein PAD systems were built upon non-deep learning approaches, utilizing handcrafted feature descriptors like LBP, GLCM, and Gabor filters paired with classifiers such as SVM or LightGBM. The primary advantages of these methods are their simplicity, interpretability, and low computational cost, making them suitable for real-time operations in resource-limited environments. While advancements in score-level fusion and ensemble techniques can enhance their reliability, these traditional models struggle to keep pace with the increasing complexity of modern spoofing attacks. This limitation has driven the adoption of deep learning, which offers superior representational capabilities. CNNs, in particular, have become essential to modern PAD systems by enabling end-to-end learning. Unlike traditional methods, CNNs automatically learn the distinctive hierarchical features that separate authentic finger vein images from fakes, eliminating the need for manual feature engineering.

Recent research has also explored hybrid models that merge the strengths of both paradigms. For example, Kim et al. [14] introduced a novel approach that integrates fractal analysis with a CNN classifier, as depicted in Figure 3 [14]. By calculating the fractal dimensions of vein patterns, the model can identify subtle, non-obvious irregularities present in high-quality fakes. This hybrid method reported strong results across several spoof categories, demonstrating that augmenting CNNs with statistical features can significantly improve their resilience against sophisticated presentation attacks.

The evolution of PAD has increasingly incorporated advanced deep learning architectures to address sophisticated spoofing threats. Beyond standard CNNs, innovative models are being explored to capture more complex data patterns, such as in finger vein authentication systems that leverage deep learning for enhanced security [27]. One significant advancement is the application of attention mechanisms and Transformer-based networks. Chen et al. [18] proposed a lightweight Vision Transformer model combined with a Multi-Scale Spatial-Temporal map. By using self-attention modules to process both spatial and temporal texture dynamics, their model achieved exceptional APCER and BPCER performance across intricate spoof scenarios, proving its suitability for embedded PAD systems. Attention mechanisms have also been leveraged for more targeted feature enhancement. Sulaiman et al. [20] developed a deep regional learning model that dynamically focuses on local regions of interest, such as vein

Figure 2
Score fusion using ensemble network architecture

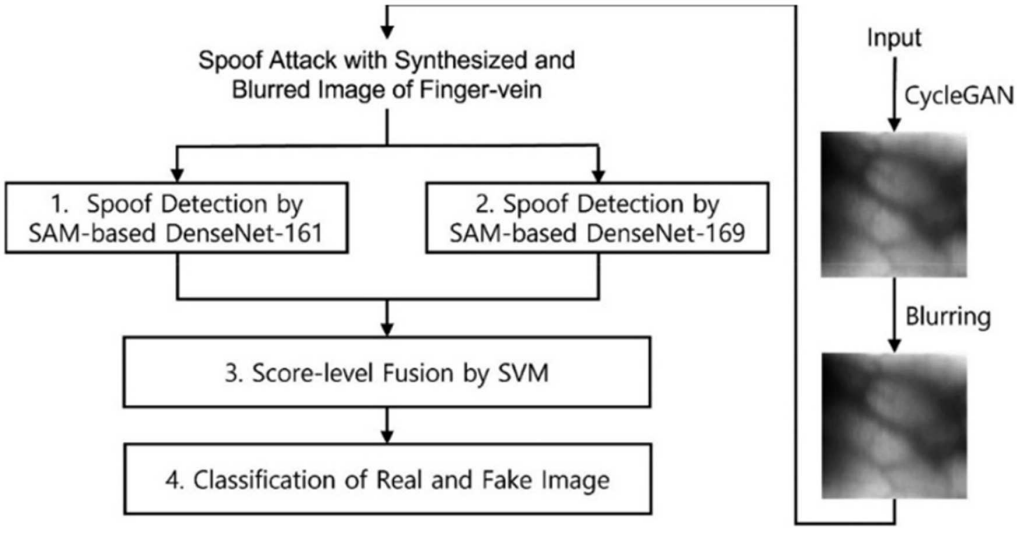
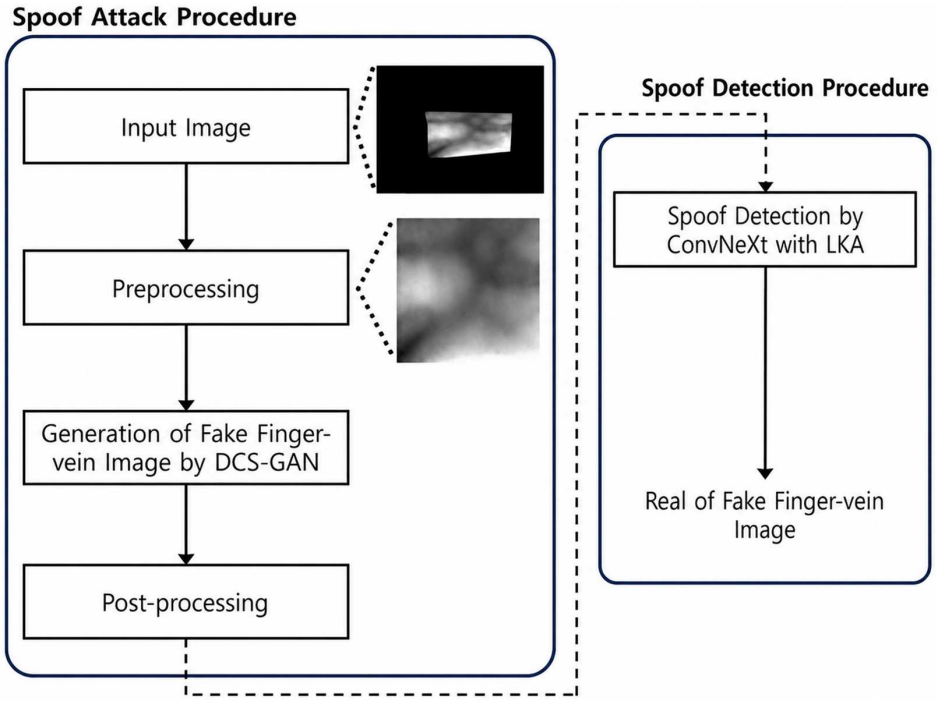


Figure 3
DCS-GAN architecture



bifurcations and junctions. This approach improves PAD accuracy in challenging conditions like motion blur or low light and reduces false positives by distinguishing genuine anatomical features from shallow surface artifacts.

Hybrid architecture that combines different model types is also proving effective. For instance, the integration of sequential models like Long Short-Term Memory (LSTM) networks with CNNs has shown promise. Abbas [22] proposed a hybrid model where a CNN extracts spatial features and an LSTM component models the structural or temporal dependencies between vein segments, making it particularly useful for video-based acquisition systems. Furthermore, specialized CNNs are being designed for niche applications. Tran et al. [21] developed an anti-aliasing CNN specifically for virtual and augmented reality environments. Their network is tailored to eliminate rendering artifacts and jagged lines from finger vein images captured on nonstandard displays, achieving 99.94% accuracy on specialized datasets and demonstrating how context-aware enhancements can enable robust PAD in real-time immersive systems. Collectively, these advanced deep learning approaches demonstrate superior performance over traditional handcrafted methods, especially against novel spoofs and in challenging acquisition environments. However, their ability to generalize against unknown attack types remains a critical challenge, requiring access to diverse training data and extensive optimization.

In addition to enhancing recognition robustness through hybrid strategies, understanding the evolving threat of template-based biometric reconstruction is critical. Recent studies have introduced advanced methodologies to restore or synthesize high-fidelity fake biometric samples from leaked templates. For instance, Sun et al. [28] utilized deep reinforcement strategies to iteratively modify initial images, achieving highly natural palmprint reconstructions. Similarly, Yan et al. [29] proposed a versatile black-box attack using a modified Progressive Generative Adversarial Network (ProGAN) to generate realistic images capable of deceiving both handcrafted and deep learning–based recognition algorithms.

A critical vulnerability in existing finger vein PAD systems is their poor generalization against unseen or "zero-day" attacks. This weakness stems from a fundamental reliance on supervised learning paradigms, such as CNNs or hybrid architectures, which require large, labeled datasets of both genuine and known spoof samples to build discriminative models [12]. However, collecting diverse and comprehensive spoof data is logistically challenging and costly, making it impossible to keep pace with rapidly evolving attack methods like advanced 3D-printed molds or deepfake-generated vein patterns. Consequently, these models tend to overfit to known attack signatures and fail when confronted with novel threats that exploit gaps in their training data.

Even advanced hybrid frameworks designed to be computationally efficient, such as those combining depthwise separable convolutions (DSC) with linear support vector machines (LSVM), remain constrained by this dependency on labeled spoof data, rendering them ineffective against zero-day exploits [9]. This vulnerability is further compounded by threats like backdoor attacks, where malicious actors can poison the training data to create hidden triggers, causing the system to misclassify specific spoofs as genuine during deployment [23].

While hardware-based countermeasures like multispectral imaging or haptic sensors attempt to address this by validating physiological traits like tissue density or blood perfusion, they are not a complete solution. These systems are often expensive and can be circumvented by sophisticated spoofs engineered to mimic biological properties, such as the thermal conductivity of human skin or the pulsation of blood flow [24]. Collectively, these weaknesses highlight that current PAD systems are inherently reactive and static. By relying on predefined attack signatures, they cannot

adapt to the rapid innovation of adversaries. This necessitates a paradigm shift toward unsupervised, adaptive techniques that can proactively identify anomalies without prior exposure to spoof samples.

One-class anomaly detection offers a promising solution to the vulnerabilities of finger vein PAD systems, particularly against unseen presentation attack instruments (PAIs). Unlike traditional methods, these approaches focus on learning the intrinsic characteristics of genuine data. Convolutional autoencoders achieve a 2.00% detection equal error rate against diverse spoofs [30], while CD-OCC enhances discrimination through unsupervised clustering [31]. IGD models ensure robustness despite sparse or noisy data [32], offering proactive, generalizable defense against evolving spoofing threats without relying on labeled attacks.

The existing literature on finger vein PAD reveals significant limitations in traditional and supervised deep learning approaches, particularly their dependence on comprehensive labeled spoof datasets and their limited generalizability to novel, unseen attacks. These shortcomings underscore the need for adaptive, spoof-agnostic methods capable of addressing zero-day threats in real-world biometric systems. To overcome these challenges, this study proposes a novel one-class anomaly detection framework based on a DAE, trained exclusively on genuine finger vein samples to ensure robust detection of diverse spoofing attempts. The following sections detail the methodology and implementation of this approach.

3. Methodology

3.1. Dataset preparation

The Finger Vein USM (FVPAD-USM) dataset [33] was employed in this study, comprising authentic finger vein images alongside two categories of spoof attacks: those presented via digital displays and glossy photo paper. As summarized in Table 1, the experimental dataset consisted of 719 authentic samples from live subjects and 1440 presentation attack samples, with the latter evenly divided into 720 digital display spoofs and 720 glossy photo paper attacks. To facilitate model development and evaluation, the dataset was partitioned as follows: the training set included 575 authentic samples (approximately 80% of the genuine data) in accordance with one-class learning principles; the validation set comprised 72 genuine samples and 144 attack samples (72 from each attack type) for hyperparameter optimization and early stopping determination; and the test set encompassed the remaining 72 genuine samples along with all 1440 attack samples, thereby enabling a thorough assessment across attack variants while maintaining the natural data distribution. Representative images of genuine samples and spoof attack types are illustrated in Table 2.

Table 1
Data composition and partitioning

Data category	Training	Validation	Test	Total
Genuine samples	575	72	72	719
Digital display	–	72	648	720
Glossy photo paper	–	72	648	720
Presentation attack	–	144	1296	1440
Overall total	575	216	1368	2159

Table 2
Example of genuine and spoof data

Genuine sample used in the train dataset	Digital display spoof used for PAD	Glossy photo paper spoof used for PAD
		

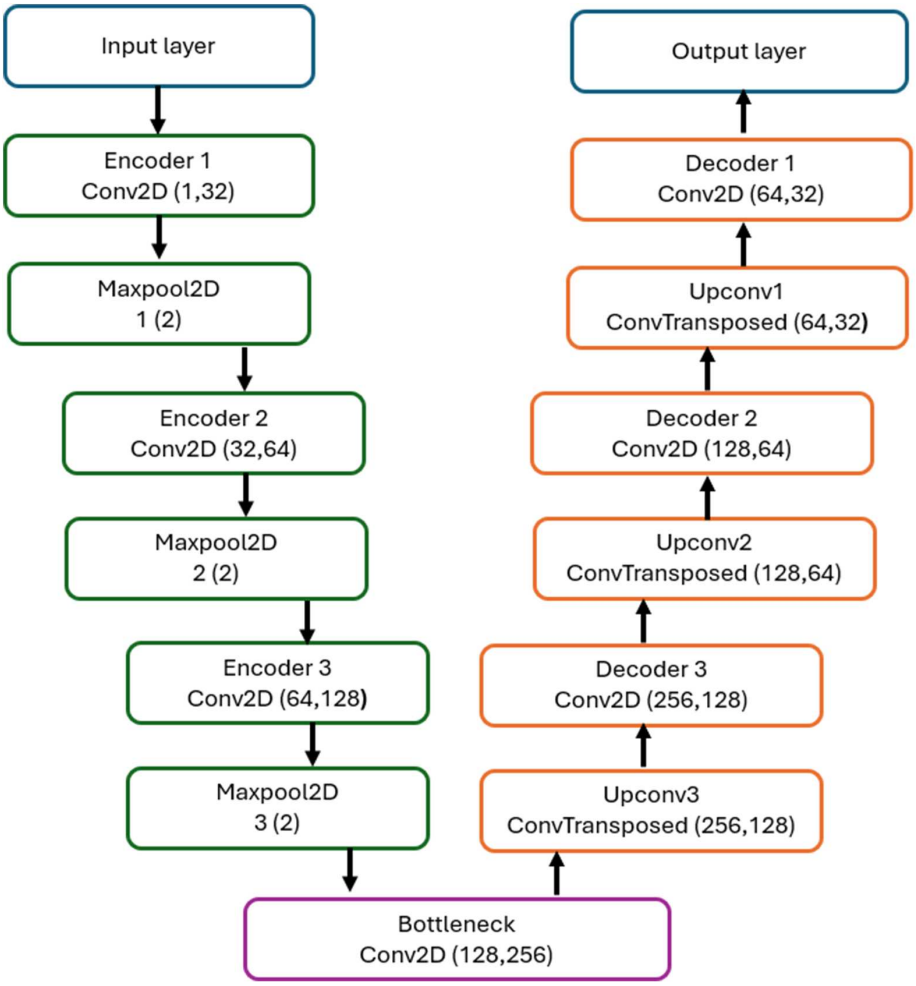
To emulate a realistic anomaly detection scenario, only genuine samples were utilized during the training phase, with spoof images reserved exclusively for validation and testing purposes, thereby enabling the assessment of the model's capability to detect unseen attack types. All images were preprocessed by resizing to 128 × 128 pixels, converting to single-channel grayscale, and normalizing pixel intensities to the range [−1, 1] to facilitate accelerated convergence during training.

To enhance model robustness and mitigate overfitting, a carefully designed data augmentation strategy was applied exclusively to the genuine training images. This approach was intended to handle natural variations in biometric acquisition while strictly avoiding any synthetic spoof artifacts that could complicate the one-class learning process. The augmentations encompassed limited geometric transformations, specifically random rotations within ±5 degrees and translations up to ±5% of the image dimensions, effectively mimicking natural variations in finger placement. Additionally, controlled Gaussian noise was introduced to emulate sensor variability. To further expand the diversity of the genuine class without relying on unrealistic spoofing traits, physiology-inspired enhancements were applied. These included random vein masking and localized contrast adjustments to help the model distinguish between venous and non-venous regions. Finally, to aid the model in identifying artifacts from electronic displays while adhering to one-class principles, conditional pixel-grid simulations were utilized by darkening specific rows and columns.

3.2. Model architecture

The core model used in this study is a DAE (Figure 4) based on a U-Net-inspired architecture. The input finger vein images were preprocessed into single-channel grayscale images sized at 128 × 128 pixels, normalizing them to a range of [−1, 1] to improve how well the model converges. The encoder pathway consists of three sequential blocks, each featuring two convolutional layers with 3 × 3 kernels, along with batch normalization and ReLU

Figure 4
Denoising autoencoder



activations. After that, 2 × 2 max pooling was applied, which gradually shrinks the spatial dimensions while boosting the feature depth from 32 to 128 channels. This compression helps the model to filter out unnecessary variations while preserving the essential vein topology.

The bottleneck, which has a spatial resolution of 16 × 16 pixels and uses 256 channels, is designed to create a focused latent representation dimension with the essential characteristics of veins. The decoder architecture then reconstructs the original resolution from the bottleneck utilizing symmetry through the transposed convolutions and through skip links that tap back to the parallel stages of the encoder with the upsampled outputs from the decoder, enabling precise spatial recovery of vein structures. Each block in the decoder processes these combined features through two convolutional layers before advancing to the next resolution level. Finally, the reconstruction head uses a 1 × 1 convolutional layer that activates with the tanh function, which generates denoised outputs of the input size. This design is very elegant in that it uses under-completeness, which forces the model to concentrate on the most important vein features, while allowing for reconstruction and detail through hierarchical feature fusion. This gives a strong starting point for identifying anomalies in finger vein presentation attacks.

3.3. Multi-objective anomaly detection loss framework

The training goal combines a multi-scale reconstruction error, constraints for preserving gradients, and a deep SVDD loss to enhance both denoising performance and anomaly detection at the same time. The main reconstruction loss uses a multi-scale mean squared error method that checks the output quality across three different resolution levels: the original (100%), downsampled (50%), and a further reduced scale (25%), with weights of 0.5, 0.3, and 0.2, respectively. This layered evaluation guarantees that both the overall structure and the finer textures are maintained. This is mathematically defined as:

$$L_{Recon} = \sum_{i=1}^{3} w_i \cdot \mathrm{MSE}\left(\hat{y}s_i, ys_i\right) \tag{1}$$

where $s = \{1.0,\ 0.5,\ 0.25\}$ represent the scale factors for bilinear interpolation, $w = \{0.5,\ 0.3,\ 0.2\}$ are the respective weights, $\hat{y}$ is the predicted output, and y is the target image.

Additionally, a gradient loss term helps maintain edge consistency by comparing Sobel-filtered gradients between the reconstructed and target images using L1-norm, which penalizes

any blurring of important vein boundaries. This is calculated as the mean absolute error of the horizontal and vertical gradients:

$$L_{Grad} = \|\nabla_x \hat{y} - \nabla_x y\|_1 + \|\nabla_y \hat{y} - \nabla_y y\|_1 \quad (2)$$

where ∇x and ∇y represent 2D convolutions using 3 × 3 Sobel filters.

The SVDD component works in the latent space to effectively group genuine features while pushing away any spoof representations. During the training process, features from genuine samples that are pulled from the bottleneck layer go through average pooling and are fine-tuned toward a hypersphere center that updates dynamically. The center and its adaptive radius are kept in check through momentum-based updates (with a momentum factor of 0.9). The loss penalizes genuine features that fall outside this hypersphere:

$$L_{SVDD} = \frac{1}{N_{gen}} \sum_{i=1}^{N_{gen}} \max(0, \|f_i - c\|^2 - R) \quad (3)$$

where N_{gen} is the number of genuine samples in the batch, f_i represents the pooled bottleneck features for genuine sample i, c is the dynamically updated hypersphere center, and R is the dynamically updated radius. Finally, the global loss function seamlessly combines these components. The overall loss function uses specific scaling hyperparameters to allow the model to focus on accomplishing the vein reconstruction well while developing a discriminative latent space:

$$L_{Total} = \gamma L_{Recon} + \alpha L_{Grad} + \beta L_{SVDD} \quad (4)$$

where the initialized weights are $\gamma = 0.8$ for reconstructions, $\alpha = 0.1$ for keeping gradients, and $\beta = 0.1$ for ensuring SVDD compactness. This strategy in weighting the loss function allows the model to focus on accomplishing the vein reconstruction well while developing a discriminative latent space to segregate the spoof anomalies from the true feature distributions.

3.4. Model training

Model training involved a strict optimization procedure that integrates adaptive learning rate scheduling, gradient stabilization, and early stopping to optimize for generalization and avoid overfitting. Model training was initialized with Adam optimization (*learningrate* = $1e^{-4}$, *weightdecay* = $1e^{-5}$) and continued for 100 epochs, average batch size of 8, and with gradually adjusted learning rates using a ReduceLROnPlateau learning rate scheduler that has a 50% decrease after 5 epochs of stagnant improvement (or worse) in validation loss. Each training iteration involved training with only genuine samples, where the input images were adapted through controlled augmentation of a limited geometric transformation (±5%) degree rotations and (±5%) of image size translations combined with reduced Gaussian noise ($\sigma = 0.05$) to imitate variability in acquisition, while avoiding changing the topology of the vein structures. During the forward propagation, reconstructed outputs and bottleneck features were computed using the multi-objective loss, and backpropagation was limited via gradient clipping at a max norm of 5.0 for stable convergence.

After each epoch, validation is performed using balanced sets of genuine and spoof samples. In the calculation of the loss metric, real-time adjustments are made to the SVDD center and radius, using only the genuine feature distributions. Training consisted of an early stopping criterion to ensure the training was stopped while still improving model performance. If validation does not improve for the last 30 epochs, training will stop, and the checkpoint will be restored from its best performance. Alongside capturing training and validation loss, using the Tensorboard logging functionality, how the training was on the convergence trajectory was visualized, as well as whether or not the model was experiencing stability. It became apparent that training loss is not the primary criterion, supporting an empirical methodology toward training rather than achieving the best pure reconstruction accuracy from the generative model. The goal of the training was to ensure the feature representations had discriminative ability, and it could be properly tested with new examples and varied presentation attacks. But the model was prevented from learning all the artifacts of the training data. The model's best performance was captured when validation loss reached its minimum, serving as the final output of the database model to be assessed for its anomaly detection performance.

3.5. Spectral error optimization and dynamic thresholding

The spectral error calculation method uses fast Fourier transforms to shift the image samples to the frequency domain and uses a complex weighting method that is based on the different types of spoof protection vulnerability. The start of the complex mask starts with a radial distance mask relative to the Direct Current (DC) zero-frequency component of the spectra and weights different frequency bands differently according to the attack characteristics of the attack: frequency bands between 25 and 50 (~ *radius*) are provided weighting of 1.2× for paper-based spoofing so that the texture artifacts found in print, and for digital display, weighting is set to 1.8× for high-frequency bands > 50 to amplify the pixel-grid distortion. The conditional weighting during inference is invoked in real time by checking for spoof type identifiers, ensuring spectral error calculation targets attack-specific spectral signatures during error computation. The final error metric consists of calculating the mean absolute difference between these adaptively weighted magnitude spectra that produce a highly discriminative signal that exposes subtle reconstruction discontinuities for each of the presentation vulnerabilities.

Calibrating spoof-specific thresholds involves an exhaustive search for each attack pattern over the distribution of the validation error to minimize the Average Classification Error Rate (ACER). The validation errors from genuine samples and the validation errors from the respective spoof samples are combined for each attack category, and an exhaustive search is conducted at every possible threshold. At each candidate threshold, the model will compute the APCER and the BPCER and select the thresholds that minimize their average ACER. This is important because the error distributions of different types of spoofs will be different in fundamentally different ways. For example, digital display attacks typically produce noticeably higher reconstruction errors than paper-based spoofs. Having a universal threshold will yield unacceptable trade-offs that either miss the smaller compromises involved in paper attacks or reject genuine samples when genuine samples are overtly spoofed using displays. Attack-specific decision boundaries allow the model to play with maximum sensitivity without having to be manually recalibrated.

Performance evaluation employs a fully stratified method, where each spoof type is then evaluated independently using bespoke detection parameters. The anti-spoofing study is conducted separately for each type of attack on a different dataset and for each digital display and glossy photo paper. Each of the respective datasets is then run through the same reconstruction

and intrinsic spectral analysis pipelines; however, when computing error, the frequency weighting is specific to the spoof type. The detection thresholds, which were optimized for each attack type during the validation phase, are then applied to the respective test datasets during evaluation. The assessment uses three pairs of standardized metrics for the assessment of biometric PAD systems, based on the ISO/IEC 30107-3 definition, and used in existing literature for finger vein PAD. The first is APCER, which is defined as:

$$APCER = \frac{N_{APG}}{N_{AP}} \times 100 \quad (5)$$

where N_{APG} represents the number of attack presentations classified as genuine and N_{AP} represents the number of attack presentations. APCER describes the proportion of spoof attacks made by an attacker that are not detected such as presentation attacks falsely classified as genuine. The next metric is the Bonafide Presentation Classification Error Rate (BPCER), which is defined as follows:

$$BPCER = \frac{N_{GPA}}{N_{GP}} \times 100 \quad (6)$$

where N_{GPA} represents the number of genuine presentations classified as attack and N_{GP} represents the number of genuine presentations. BPCER describes the rate of legitimate user rejections, for example, genuine samples incorrectly identified as attacks. The last one is ACER, which uses APCER and BPCER to assess the overall performance bias of the measurement systems as a whole. ACER is defined as:

$$ACER = \frac{APCER + BPCER}{2} \quad (7)$$

As stated in the study by J. Kolberg et al. [16] regarding finger vein-based PAD, these standardized metrics are critical for measuring detection performance against various potential PAIs. They also ensure compliance with international standards and allow for reproducible evaluations across multimodal finger vein recognition systems. This approach allows for serious verification of cross-attack robustness and, importantly, shows performance differences in attack modality, from which information can be gained for refinement of a better detection system, all while considering real-world situations where the types of proxies cannot be decided in advance.

4. Result and Discussion

4.1. Performance of the proposed denoising autoencoder

Table 3 shows the result of the proposed DAE. The proposed DAE model was evaluated on two types of presentation attacks, digital display spoof and glossy photo paper, using the FVPAD-USM dataset. The model was trained solely on genuine finger vein images, making this a true one-class anomaly detection task. The results demonstrate strong overall performance. Specifically, the model achieved an APCER of 0.56%, BPCER of 5.56%, and ACER of 3.06% for digital display spoof. Against glossy photo paper, it recorded an APCER of 1.81%, a BPCER of 16.67%, and an ACER of 9.24%. These outcomes suggest that the model is particularly effective in detecting spoofing attempts that introduce significant texture or feature inconsistencies.

Table 3
Performance of the denoising autoencoder across two types of PAIs

Spoof types	APCER (%)	BPCER (%)	ACER (%)
Digital display spoof	0.56	5.56	3.06
Glossy photo paper	1.81	16.67	9.24

Throughout the training process, the model's loss curves for both training and validation sets demonstrated consistent convergence behavior as shown in Figure 5. There was no significant gap between the training and validation losses, indicating that the model was neither underfitting nor overfitting the genuine data distribution. This stability can be attributed to the application of early stopping, which halted training once the validation loss ceased to improve over several epochs. By doing so, the model preserved its generalization capability and avoided memorizing noise or overemphasizing fine details in the training data.

4.2. Comparative analysis with baseline methods

To validate the effectiveness of the proposed method, a comparative analysis was conducted against four baseline approaches:

Figure 5
Training loss and validation loss during training

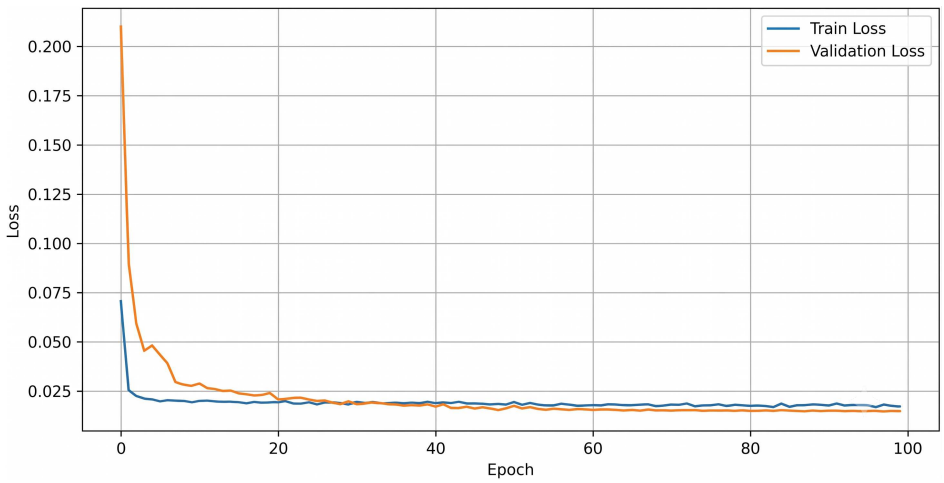


Table 4
Performance of the baseline methods across two types of PAIs

Method	Training PAI	Test: Digital display APCER (%)	Test: Glossy photo paper APCER (%)	BPCER (%)
SVM [16]	Digital display spoof	0.42	83.19	1.39
	Glossy photo paper	97.78	0	0
DSC-LSVM [9]	Digital display spoof	0	98.33	0
	Glossy photo paper	99.72	0	0
CNN [34]	Digital display spoof	0.14	99.03	0
	Glossy photo paper	0	0	100
CNN + LSTM [19]	Digital display spoof	34.78	100	0
	Glossy photo paper	48.88	11.31	75

SVM [16], Depthwise Separable Convolution with Linear SVM (DSC-LSVM) [9], CNN [34], and CNN combined with LSTM (CNN + LSTM) [19]. Each method was evaluated using the same test conditions and dataset, ensuring consistency across attack types and metrics.

According to the results reported in Table 4, the SVM model achieved varied performance depending on the training spoof type. For instance, when trained on digital display spoof, it recorded a low APCER of 0.42% for digital display but a significantly high APCER of 83.19% for glossy paper. When trained on glossy paper, SVM completely failed to detect digital display, reaching APCERs of 97.78%. In all configurations, BPCER remained low at 0% or 1.39%.

The DSC-LSVM model also displayed poor generalization. When trained on digital display spoof, it misclassified glossy paper samples with an APCER of 98.33%. Likewise, when trained on glossy paper, it failed to detect digital display attacks, with APCERs of 99.72%. As with SVM, the BPCER for DSC-LSVM remained consistently low across all training configurations. In contrast, CNN [34] exhibited inconsistent behavior. For example, when trained on digital display spoof, it achieved a very low APCER of 0.14% for that spoof type but failed almost completely for glossy paper with an APCER of 99.03%. When trained on glossy paper, the model failed to detect any spoofs, but wrongly rejected all genuine samples, as indicated by a 100% BPCER.

The CNN + LSTM model offered more balanced, though still imperfect, results. For example, when trained on digital display spoof, it achieved APCERs of 34.78% (digital) and 100% (glossy), with a BPCER of 0%. When trained on glossy paper, it achieved improved APCERs of 48.88% for digital, but at the cost of a high BPCER of 75%. The high variance in spoof detection and genuine rejection highlights its instability under cross-type evaluation. These results clearly demonstrate that most existing supervised methods, which include both traditional and deep learning models, struggle to generalize to unseen attack types. In contrast, the proposed DAE, trained solely on genuine samples, successfully detected a wide range of spoofing attacks while maintaining low error rates across both APCER and BPCER, as detailed in Table 3.

From a computational perspective, the DAE maintains a practical balance with 1,927,841 parameters, 7.37 MB model size, and 6.06 ms inference time. This efficiency is comparable to lightweight CNNs developed for finger vein recognition [35], while the CNN + LSTM model, by contrast, is resource-intensive with over 34 million parameters, 131.18 MB in memory, and 23.72 ms inference time. The SVM and DSC-LSVM, while lightweight (2.01 MB and 0.0451 MB, respectively), offer limited generalization. These comparisons confirm that the proposed autoencoder provides the best trade-off between detection accuracy, model size, and computational cost, making it well-suited for real-time deployment in biometric security systems. These results are shown in Tables 5 and 6.

Table 5
Comparison of model summary between denoising autoencoder and existing methods

Model	Denoising Autoencoder	SVM [16]	DSC-LSVM [9]	CNN [34]	CNN-LSTM [19]
Total parameters	1927841	528164	8993	170660827	34386722
Trainable parameters	1927841	528164	8993	170660827	34386722
Non-trainable Parameters	0	0	0	0	0
Parameter memory	7.3541 MB	2.0148 MB	0.0343 MB	65.1019 MB	131.1749 MB
Buffer memory	0.0108 MB	0 MB	0.0108 MB	0.0035 MB	0.0035 MB
Estimated total size	7.3650 MB	2.0148 MB	0.0451 MB	65.1054 MB	131.1784 MB

Table 6
Comparison of total inference time between denoising autoencoder and existing methods

Model	Denoising autoencoder	SVM [16]	DSC-LSVM [9]	CNN [34]	CNN-LSTM [19]
Total Inference Time (ms)	6.0564	9.0941	0.805	2.0279	23.7157

4.3. Contribution of the proposed denoising autoencoder method

The proposed DAE framework offers some game-changing benefits for finger vein PAD, especially when it comes to its resilience against new, unseen attacks. This feature is crucial for real-world biometric security systems, where attackers are always coming up with innovative spoofing materials, like synthetic films and advanced displays, that were not part of the training process. Traditional supervised models tend to fail dramatically when faced with these zero-day attacks. The baseline result shows the APCER, which can soar above 83%. In contrast, our approach keeps error rates impressively low ($ACER \leq 9.24\%$) even against untrained digital and glossy paper PAIs. This adaptability comes from its one-class paradigm, which focuses on learning the genuine physiological features of real veins instead of just memorizing attack patterns. This means our system is well-equipped to handle the evolving threats found in high-stakes environments like banking or border control, where security needs to be both robust and flexible.

Besides, one of the key benefits of this proposed method is that it removes the need to depend on spoof data during training. Traditional PAD methods often involve a lengthy process of gathering and labeling a wide range of attack samples, which can really take a toll on time, resources, and budget. Just think about the costs involved in creating spoofing materials and gathering data over several sessions. Our method completely sidesteps this issue by focusing solely on genuine samples, which not only shortens development timelines by several months but also significantly cuts down on resource costs. Even more importantly, it tackles the unrealistic challenge of trying to obtain every possible spoof variant. This is because attackers in real-world situations use all sorts of unpredictable materials, making methods that depend on spoof data outdated. The autoencoder's design, which is agnostic to spoofing, allows for scalable and cost-effective implementation, even in settings with limited resources.

Finally, the framework ensures consistently high detection accuracy while maintaining operational integrity. With impressive sub-2% APCER and sub-17% BPCER rates even in challenging PAIs, it achieves an ideal balance between security and usability. This means there are fewer chances of mistakenly accepting unauthorized users, which can lead to security issues, and fewer instances of wrongly rejecting legitimate users, which can be frustrating. Having this level of reliability is essential for fostering trust in biometric systems. To keep unauthorized access at bay from clever spoofing attempts, it is crucial to maintain a low APCER. At the same time, a well-managed BPCER is key to ensuring that genuine users are not mistakenly turned away. However, achieving this kind of operational stability can be quite a challenge for traditional methods, like CNNs, which have hit 100% BPCER. This is where the autoencoder really shines as a practical solution, effectively balancing security needs with the real-world demands of usability.

4.4. Limitation of the proposed method

While the proposed DAE achieved high overall accuracy and generalization across spoof types, several limitations were observed that may impact its performance in more diverse real-world scenarios. The most notable issue arises when dealing with digital display and glossy photo paper spoof attacks. As illustrated in Figures 6, 7, 8, and 9, the reconstructed images of

Figure 6
Comparison of the reconstructed genuine image with the reconstructed digital display spoof

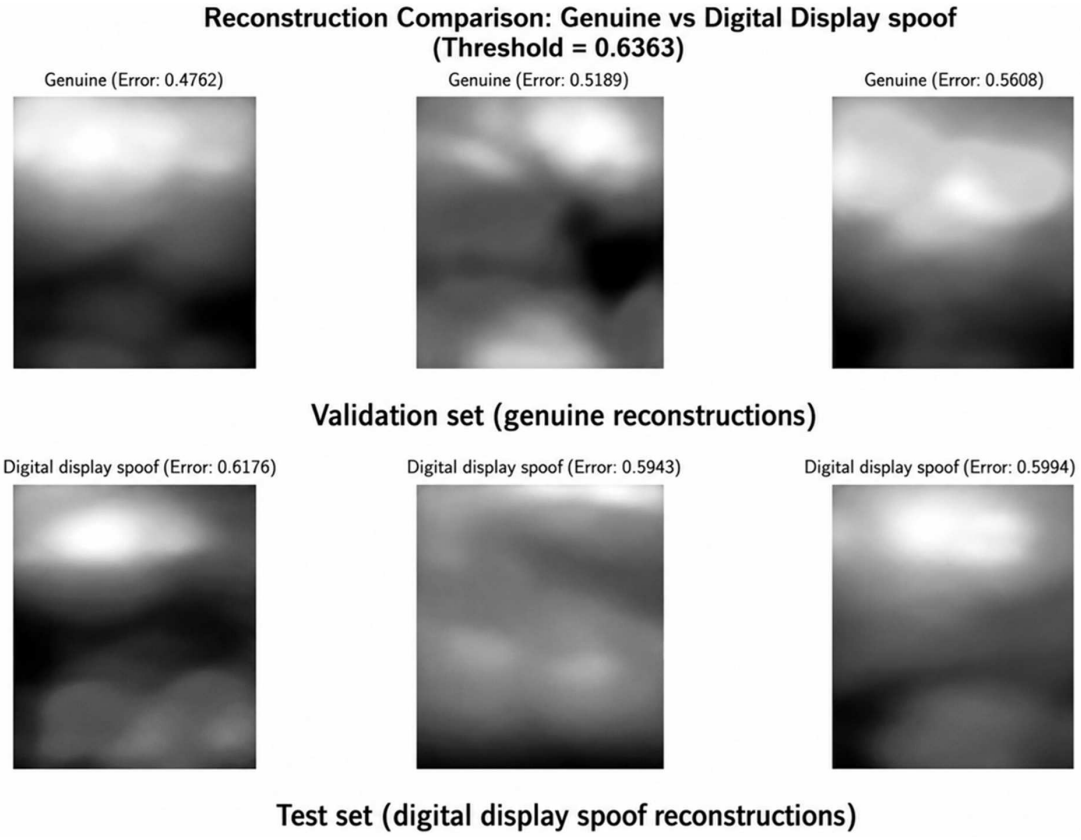


Figure 7
Comparison of reconstruction error of genuine vs digital display spoof

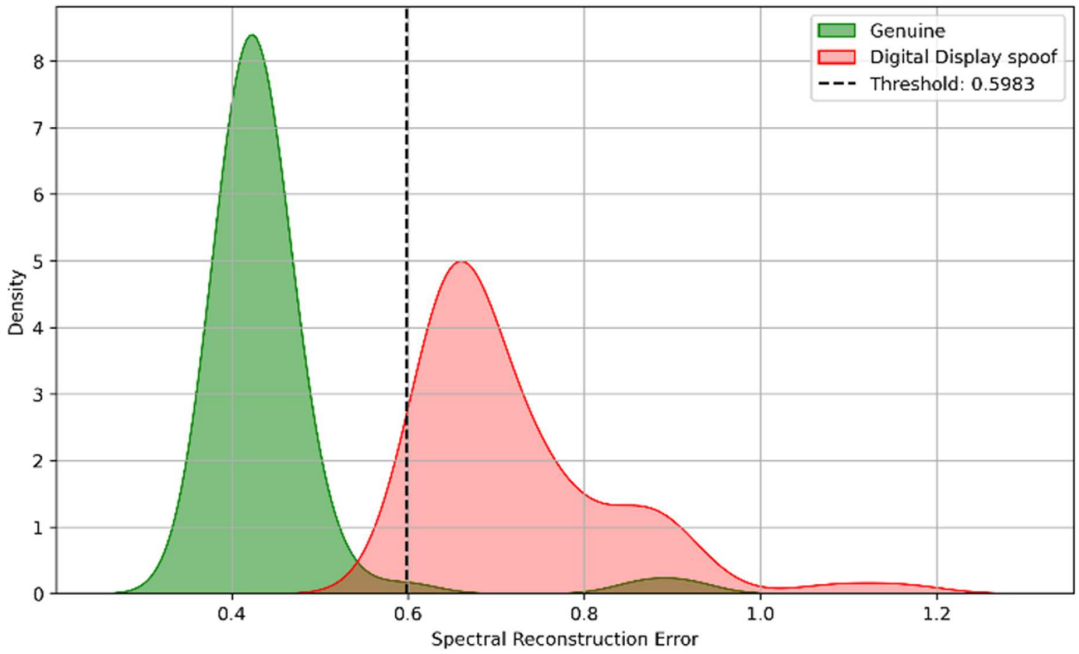
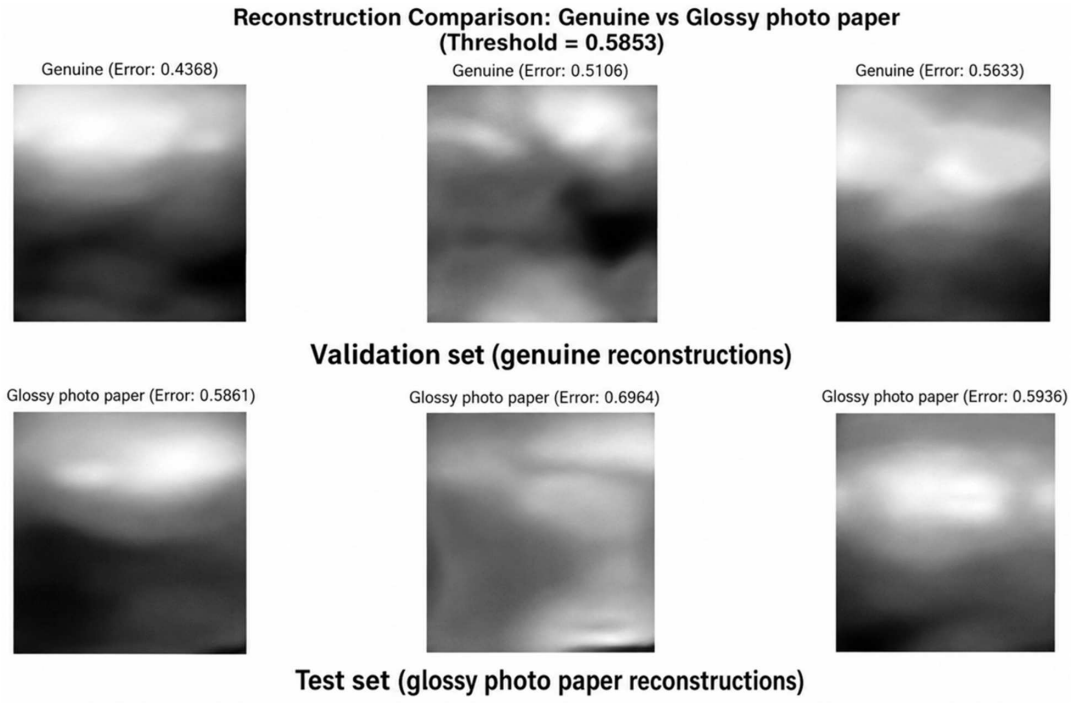


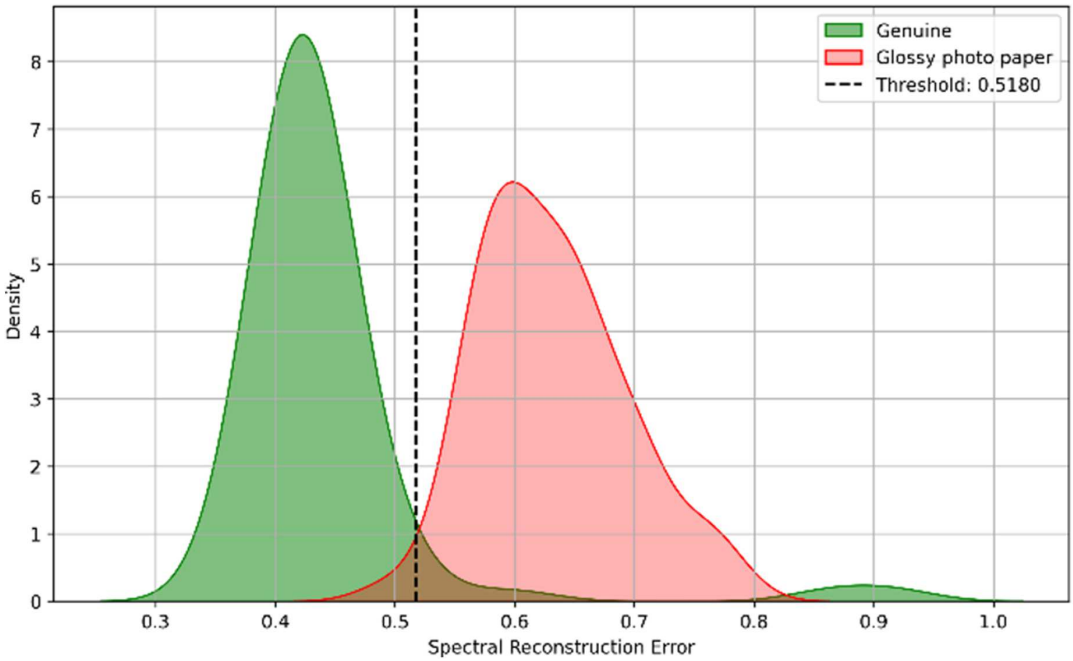
Figure 8
Comparison of the reconstructed genuine image with the reconstructed glossy photo paper



these spoof types are visually very similar to the reconstructed genuine images. This high similarity results in overlapping reconstruction error distributions, which reduces the model's ability to cleanly separate spoof and genuine samples based on thresholding.

In particular, the reflective surface and visual sharpness of glossy and screen-based spoofs closely mimic the texture and structure of real finger vein patterns. As a result, their spectral and spatial features are sometimes preserved through the autoencoder, causing lower reconstruction errors that fall within the acceptable range of genuine samples. This ambiguity contributes to higher APCER values for these spoof types, despite the model being trained solely on genuine data. Additionally, fixed thresholding, although optimal per spoof type, may still lack adaptability for spoof types with subtle feature deviations, especially under varied lighting or sensor conditions.

Figure 9
Comparison of reconstruction error of genuine vs glossy photo paper



These challenges highlight the need for further refinement. Future work could integrate dynamic or spoof-aware threshold calibration, adversarial training techniques, or enhanced frequency-domain attention mechanisms to increase sensitivity to reflective and texture-rich spoof materials. Moreover, expanding the diversity of genuine training samples through augmentation or synthetic generation could improve the model's boundary sensitivity to anomalous inputs.

5. Conclusion

This study proposed a one-class anomaly detection model based on a DAE architecture for finger vein PAD. Unlike traditional supervised approaches, the proposed method is trained exclusively on genuine finger vein images, enabling it to detect spoofing attempts without prior exposure to spoof data. By leveraging an undercomplete U-Net-inspired architecture with multi-scale reconstruction loss, gradient loss, and SVDD-based feature regularization, the model effectively learns the intrinsic distribution of genuine vein patterns and flags anomalies introduced by PAIs.

Experimental evaluation using the FVPAD-USM dataset demonstrated that the proposed model performs particularly well against unseen attacks. It achieved competitive performance on digital display and glossy photo paper attacks, although the reconstruction similarity in those spoof types highlighted areas for further improvement. Comparisons with baseline methods, such as SVM, DSC-LSVM, CNN, and CNN + LSTM, revealed that the DAE achieved a superior trade-off between accuracy, computational efficiency, and generalization to unseen attacks. The model's relatively small size and fast inference time make it well-suited for deployment in practical biometric systems.

Despite its effectiveness, the model exhibits limitations when spoof samples closely mimic the structure and contrast of genuine images. Future enhancements could include adaptive thresholding, more advanced frequency-domain error modeling, or integration with adversarial robustness techniques. Overall, this research contributes a lightweight, unsupervised, and generalizable solution to the evolving challenge of PAD in finger vein biometrics, addressing real-world constraints on spoof data availability while preserving strong security and user convenience.

6. Future Recommendation

The current framework lays a solid groundwork for detecting finger vein presentation attacks, but there are still several strategic paths that need to be explored to boost its effectiveness in real-world scenarios. For starters, fine-tuning the multi-scale spectral error metric by using adaptive material-aware frequency band weighting could really help lower the high BPCER seen with glossy photo paper attacks (16.67%). By adjusting frequency sensitivity based on the optical characteristics of spoof materials, like enhancing high-frequency components for shiny surfaces or focusing on mid-range bands for textured papers, the model could be more effectively separated from sensor noise and actual physiological signals, which would help reduce false rejections.

Additionally, looking into constrained multimodal fusion with other biometric traits like finger geometry or thermal patterns or extra sensors such as near-infrared depth imaging could strengthen our defenses against advanced attacks where the differences between real and fake samples blur. This fusion should focus on lightweight, attention-based feature gating to prevent the performance drops that have been observed in fingerprint PAD studies when combining modalities without care.

While the current study operates on individual images, multi-instance fusion, such as aggregating PAD scores from multiple fingers using the index and middle fingers, offers a pathway to mitigate the risk of a single false rejection due to local noise or partial occlusion. Similarly, employing score-level fusion to intelligently combine our unsupervised anomaly scores with auxiliary supervised classifiers or quality metrics could create a mixture-of-experts defense. This approach would leverage the

generalization of our one-class model, along with the discriminative power of supervised methods, potentially reducing the ACER in high-security deployments.

To truly enhance our models, rigorously validating them using larger datasets that reflect a wide range of demographics, environmental factors like humidity and temperature, and new types of attack methods such as 3D-printed vein replicas and synthetic films was essential. Teaming up with industry leaders could open the door to valuable real-world data from fields like banking or border control. This would enable us to tackle challenging situations, such as dealing with aged skin or scar tissue. Apart from that, by integrating optical physics-based enhancements like mimicking how light scatters, absorbs, and refracts in finger vein imaging, the model will be strengthened against emerging spoofing techniques. For instance, by simulating how various wavelengths interact with skin layers and synthetic materials, reliance on gathering empirical data could be reduced while enhancing the model's robustness.

Finally, bringing this system into federated learning frameworks would really boost privacy-focused improvements across a range of devices, like ATMs and smartphones. This approach would let the system continuously adapt to lock attack patterns without the need to centralize sensitive biometric data. It is definitely a positive move, especially considering the urgent demand for scalable and flexible PAD solutions in our interconnected authentication systems, where zero-day attacks call for swift and effective action. All in all, these strategies could transform the prototype into a dependable shield against the presentation threats that might be faced in the future.

Acknowledgment

The authors express their gratitude to Universiti Sains Malaysia for providing the facilities and support necessary to carry out this research.

Funding Support

This research was supported by Universiti Sains Malaysia through the PRIME Grant (No. R502-KRARP003-0000000897 K134) and Visionlytics Sdn. Bhd. via the Industry-DTD Grant (No. R504-LR-GAL007-0000000896-V106).

Ethical Statement

The authors declare that this study did not require formal ethical approval because it exclusively utilized a publicly available, open-access dataset (the FVPAD-USM database). No new data collection involving human or animal subjects was performed by any of the authors. The data utilized were previously anonymized by the original creators and used solely for research purposes.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in the FVPAD-USM database at http://drfendi.com/fvpad_usm_database/, reference number [33].

Author Contribution Statement

Mohd Shahrimie Mohd Asaari: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Bakhtiar Affendi Rosdi:** Validation, Investigation, Writing review & editing, Supervision. **Andrew Tiong Hoe Pin:** Methodology, Software, Validation, Investigation, Data curation, Visualization. **Zahid Ur Rahman:** Investigation, Resources, Data curation. **Muhammad Firdaus Akbar:** Validation, Funding acquisition.

References

[1] Lee, Y. K., Jeong, J., & Kang, D. (2022). An effective orchestration for fingerprint presentation attack detection. *Electronics*, *11*(16), 2515. https://doi.org/10.3390/electronics11162515

[2] Gupta, A., & Verma, A. (2021). Fingerprint presentation attack detection approaches in open-set and closed-set scenario. *Journal of Physics: Conference Series*, *1964*(4), 042050. https://doi.org/10.1088/1742-6596/1964/4/042050

[3] Abdullakutty, F., Elyan, E., & Johnston, P. (2021). A review of state-of-the-art in Face Presentation Attack Detection: From early development to advanced deep learning and multi-modal fusion methods. *Information Fusion*, *75*, 55–69. https://doi.org/10.1016/j.inffus.2021.04.015

[4] Tan, C. B., Hijazi, M. H. A., Khamis, N., Nohuddin, P. N. E. B., Zainol, Z., Coenen, F., & Gani, A. (2021). A survey on presentation attack detection for automatic speaker verification systems: State-of-the-art, taxonomy, issues and future direction. *Multimedia Tools and Applications*, *80*(21-23), 32725–32762. https://doi.org/10.1007/s11042-021-11235-x

[5] Bai, H., Tan, Y., & Li, Y.-J. (2024). Mask-guided network for finger vein feature extraction and biometric identification. *Biomedical Optics Express*, *15*(12), 6845. https://doi.org/10.1364/BOE.535390

[6] Haider, S. A., Ashraf, S., Larik, R. M., Husain, N., Muqeet, H. A., Humayun, U., . . . , & Khan, M. F. (2023). An improved multimodal biometric identification system employing score-level fuzzification of finger texture and finger vein biometrics. *Sensors*, *23*(24), 9706. https://doi.org/10.3390/s23249706

[7] Jenkins, J., & Roy, K. (2024). Exploring deep convolutional generative adversarial networks (DCGAN) in biometric systems: A survey study. *Discover Artificial Intelligence*, *4*(1), 42. https://doi.org/10.1007/s44163-024-00138-z

[8] Shaheed, K., Szczuko, P., Ullah, I., Mojeed, H. A., Balogun, A. O., & Capretz, L. F. (2024). Finger vein presentation attack detection method using a hybridized gray-level co-occurrence matrix feature with light-gradient boosting machine model. In *32nd International Conference on Information Systems Development*, 1–13. https://doi.org/10.62036/ISD.2024.54

[9] Shaheed, K., Mao, A., Qureshi, I., Abbas, Q., Kumar, M., & Zhang, X. (2022). Finger-vein presentation attack detection using depthwise separable convolution neural network. *Expert Systems with Applications*, *198*, 116786. https://doi.org/10.1016/j.eswa.2022.116786

[10] Kim, S. G., Choi, J., Hong, J. S., & Park, K. R. (2022). Spoof detection based on score fusion using ensemble networks robust against adversarial attacks of fake finger-vein images. *Journal of King Saud University - Computer and Information Sciences*, *34*(10), 9343–9362. https://doi.org/10.1016/j.jksuci.2022.09.012

[11] Schuiki, J., Linortner, M., Wimmer, G., & Uhl, A. (2023). Extensive threat analysis of vein attack databases and attack detection by fusion of comparison scores. In S. Marcel, J. Fierrez, & N. Evans (Eds.), *Handbook of biometric anti-spoofing: Presentation attack detection and vulnerability assessment* (3rd ed., pp. 467–487). Springer. https://doi.org/10.1007/978-981-19-5288-3_17

[12] Afonso Pereira, J., Sequeira, A. F., Pernes, D., & Cardoso, J. S. (2020). A robust fingerprint presentation attack detection method against unseen attacks through adversarial learning. In *2020 International Conference of the Biometrics Special Interest Group*, 1–5.

[13] Singh, J. M., Madhun, A., Li, G., & Ramachandra, R. (2021). A survey on unknown presentation attack detection for fingerprint. In *Intelligent Technologies and Applications: Third International Conference*, 189–202. https://doi.org/10.1007/978-3-030-71711-7_16

[14] Kim, S. G., Hong, J. S., Kim, J. S., & Park, K. R. (2024). Estimation of fractal dimension and detection of fake finger-vein images for finger-vein recognition. *Fractal and Fractional*, *8*(11), 646. https://doi.org/10.3390/fractalfract8110646

[15] Kolberg, J., Gomez-Barrero, M., & Busch, C. (2021). On the generalisation capabilities of fingerprint presentation attack detection methods in the short wave infrared domain. *IET Biometrics*, *10*(4), 359–373. https://doi.org/10.1049/bme2.12020

[16] Kolberg, J., Gomez-Barrero, M., Venkatesh, S., Ramachandra, R., & Busch, C. (2020). Presentation attack detection for finger recognition. In A. Uhl, C. Busch, S. Marcel, & R. Veldhuis (Eds.), *Handbook of vascular biometrics* (pp. 435–463). Springer International Publishing. https://doi.org/10.1007/978-3-030-27731-4_14

[17] Al-Obaidy, N. A. I., Mahmood, B. S., & Alkababji, A. M. F. (2022). Finger veins verification by exploiting the deep learning technique. *Ingénierie des Systèmes d'Information*, *27*(6), 923–931. https://doi.org/10.18280/isi.270608

[18] Chen, L., Guo, T., Li, L., Jiang, H., Luo, W., & Li, Z. (2023). A finger vein liveness detection system based on multi-scale spatial-temporal map and Light-ViT model. *Sensors*, *23*(24), 9637. https://doi.org/10.3390/s23249637

[19] Sameer, N. A., & Nema, B. M. (2025). LSTM and CNN hybrid model for enhanced fingerprint recognition. *Emerging Trends in Drugs, Addictions, and Health*, *5*, 100174. https://doi.org/10.1016/j.etdah.2025.100174

[20] Sulaiman, D. M., Abdulazeez, A. M., Zeba, D. A., Zeebaree, D. Q., Mostafa, S. A., & Sadiq, S. S. (2022). An attention-based deep regional learning model for enhanced finger vein identification. *Traitement du Signal*, *39*(6), 1991–2002. https://doi.org/10.18280/ts.390611

[21] Tran, N. C., Wang, J., Vu, T. H., Tai, T.-C., & Wang, J.-C. (2023). Anti-aliasing convolution neural network of finger vein recognition for virtual reality (VR) human–robot equipment of metaverse. *The Journal of Supercomputing*, *79*(3), 2767–2782. https://doi.org/10.1007/s11227-022-04680-4

[22] Abbas, T. (2023). Finger vein recognition with hybrid deep learning approach. *Journal La Multiapp*, *4*(1), 23–33. https://doi.org/10.37899/journallamultiapp.v4i1.788

[23] Zhang, H., Sun, W., & Lv, L. (2024). A frequency-injection backdoor attack against DNN-Based finger vein verification. *Computers and Security*, *144*, 103956. https://doi.org/10.1016/j.cose.2024.103956

[24] Rathore, A. S., Shen, Y., Xu, C., Snyderman, J., Han, J., Zhang, F., . . . , & Ren, K. (2022). FakeGuard: Exploring haptic response to mitigate the vulnerability in commercial fingerprint anti-spoofing. In *Network and Distributed System Security Symposium* (pp. 1–17). https://doi.org/10.14722/ndss.2022.24082

[25] Leng, L., Li, M., Kim, C., & Bi, X. (2017). Dual-source discrimination power analysis for multi-instance contactless palmprint recognition. *Multimedia Tools and Applications*, *76*(1), 333–354. https://doi.org/10.1007/s11042-015-3058-7

[26] Leng, L., & Zhang, J. (2013). PalmHash code vs PalmPhasor code. *Neurocomputing*, *108*, 1–12. https://doi.org/10.1016/j.neucom.2012.08.028

[27] Padmapriya, P., Saminathan, J., Priya, I. L., & Dayanidhy, M. (2024). Deep learning-based biometric finger vein authentication system for enhanced security. *International Journal For Multidisciplinary Research*, *6*(5), 27652. https://doi.org/10.36948/ijfmr.2024.v06i05.27652

[28] Sun, Y., Leng, L., Jin, Z., & Kim, B.-G. (2022). Reinforced palmprint reconstruction attacks in biometric systems. *Sensors*, *22*(2), 591. https://doi.org/10.3390/s22020591

[29] Yan, L., Wang, F., Leng, L., & Teoh, A. B. J. (2024). Toward comprehensive and effective palmprint reconstruction attack. *Pattern Recognition*, *155*, 110655. https://doi.org/10.1016/j.patcog.2024.110655

[30] Kolberg, J., Grimmer, M., Gomez-Barrero, M., & Busch, C. (2021). Anomaly detection with convolutional autoencoders for fingerprint presentation attack detection. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, *3*(2), 190–202. https://doi.org/10.1109/TBIOM.2021.3050036

[31] Kim, Y., & Kim, H. K. (2021). Cluster-based deep one-class classification model for anomaly detection. *Journal of Internet Technology*, *22*(4), 903–911. https://doi.org/10.53106/160792642021072204017

[32] Chen, Y., Tian, Y., Pang, G., & Carneiro, G. (2022). Deep one-class classification via interpolated Gaussian descriptor. *Proceedings of the AAAI Conference on Artificial Intelligence*, *36*(1), 383–392. https://doi.org/10.1609/aaai.v36i1.19915

[33] Rosdi, B. A. (n.d). *Presentation attack detection of finger vein USM (FVPAD-USM) database*. Drfendi. http://drfendi.com/fvpad_usm_database/

[34] Wang, Y., Shi, D., & Zhou, W. (2022). Convolutional neural network approach based on multimodal biometric system with fusion of face and finger vein features. *Sensors*, *22*(16), 6039. https://doi.org/10.3390/s22166039

[35] Hsia, C.-H., Ke, L.-Y., & Chen, S.-T. (2023). Improved lightweight convolutional neural network for finger vein recognition system. *Bioengineering*, *10*(8), 919. https://doi.org/10.3390/bioengineering10080919

How to Cite: Asaari, M. S. M., Rosdi, B. A., Pin, A. T. H., Rahman, Z. U., & Akbar, M. F. (2026). One-Class Anomaly Detection for Finger Vein Presentation Attack Detection. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA62027476