Computers, Materials & Continua

Tech Science Press

Doi:10.32604/cmc.2025.065206

ARTICLE

Tamper Detection in Multimodal Biometric Templates Using Fragile Watermarking and Artificial Intelligence

Fatima Abu Siryeh*, Hussein Alrammahi and Abdullahi Abdu İbrahim

Department of Electrical and Computer Engineering, Altinbaş University, Istanbul, 34000, Turkey

*Corresponding Author: Fatima Abu Siryeh. Email: 213720492@ogr.altinbas.edu.tr

Received: 06 March 2025; Accepted: 28 May 2025; Published: 30 July 2025

ABSTRACT: Biometric template protection is essential for finger-based authentication systems, as template tampering and adversarial attacks threaten the security. This paper proposes a DCT-based fragile watermarking scheme incorporating AI-based tamper detection to improve the integrity and robustness of finger authentication. The system was tested against NIST SD4 and Anguli fingerprint datasets, wherein 10,000 watermarked fingerprints were employed for training. The designed approach recorded a tamper detection rate of 98.3%, performing 3–6% better than current DCT, SVD, and DWT-based watermarking approaches. The false positive rate (≤1.2%) and false negative rate (≤1.5%) were much lower compared to previous research, which maintained high reliability for template change detection. The system showed real-time performance, averaging 12–18 ms processing time per template, and is thus suitable for real-world biometric authentication scenarios. Quality analysis of fingerprints indicated that NFIQ scores were enhanced from 2.07 to 1.81, reflecting improved minutiae clarity and ridge structure preservation. The approach also exhibited strong resistance to compression and noise distortions, with the improvements in PSNR being 2 dB (JPEG compression Q = 80) and the SSIM values rising by 3%–5% under noise attacks. Comparative assessment demonstrated that training with NIST SD4 data greatly improved the ridge continuity and quality of fingerprints, resulting in better match scores (260–295) when tested against Bozorth3. Smaller batch sizes (batch = 2) also resulted in improved ridge clarity, whereas larger batch sizes (batch = 8) resulted in distortions. The DCNN-based tamper detection model supported real-time classification, which greatly minimized template exposure to adversarial attacks and synthetic fingerprint forgeries. Results demonstrate that fragile watermarking with AI indeed greatly enhances fingerprint security, providing privacy-preserving biometric authentication with high robustness, accuracy, and computational efficiency.

KEYWORDS: Biometric template security; fragile watermarking; deep learning; tamper detection; discrete cosine transform (DCT); fingerprint authentication; NFIQ score optimization; AI-driven watermarking; structural similarity index (SSIM)

1 Introduction

Biometric authentication has found itself as an effective security protocol to provide access control in physical and digital infrastructure. In comparison to password authentication, biometric authentication is dependent on physiological and behavioral features in the form of fingerprints, faces, iris, and voice and uses them as identifiers as in [1]. Of these, multimodal biometric authentication gained popularity because multimodal biometrics can enhance identification accuracy by offering multiple biometric features. Nevertheless, though its security has been increased, biometric template tampering still poses a substantial threat as cited in [2]. Even though DCT-based fragile watermarking is a conventional method, the introduced framework presents a new application by combining it tightly with a deep convolutional

neural network (DCNN) for smart tamper verification in multimodal biometric systems. In contrast to previous studies focusing on watermark embedding or independent verification, this method simultaneously optimizes watermark robustness, real-time detection, and biometric recognition accuracy. It also presents a dynamic feature embedding approach and watermark correlation scoring to mitigate false positives under compression and adversarial noise—abilities not jointly addressed in current DCT, SVD, or DWT-based systems. Nevertheless, one of the main issues with current fragile watermarking methods is their vulnerability to compression and transmission-induced distortions, resulting in false tamper detection or inability to detect unauthorized changes as stated in [3]. Here, we outline a new framework for tamper detection using the combination of fragile watermarking along with artificial intelligence (AI) driven verification of multimodal biometric authentication. The concept hinges on inserting obtained iris feature descriptors as a fragile watermark into DCT coefficients in a compressed face template biometric such that an equal bandwidth dedicated to face template transmission is consumed by both the biometrics described in [4]. On reception, a deep learning-based deep convolutional neural network (DCNN) classifier is utilized to identify any tampering attempt by examining extracted watermark patterns. To assess the performance of the proposed method, we performed extensive experiments on the IIT Delhi Iris and Indian Faces Datasets. The results show that our method achieves 100% tamper detection and also sustains very high biometric recognition accuracy, with a mere 0.05% drop in iris recognition performance as discussed in [5]. Indicates the various uses of watermarking in digital forensics, military, broadcast monitoring, and privacy protection. Watermarking finds extensive applications in chip and hardware security, securing e-Governance documents, and IoT device authentication. Although the framework hereunder is being proposed for use in multimodal biometric systems, experimental testing at present will only consider the fingerprint templates given the composition of the NIST SD4 and Anguli databases. The expression "multimodal" speaks to the adaptability of system design and not to the actual composition of present datasets. As future research further develops this work, cross-modal biometric input in the forms of iris and facial templates will be added. In addition, to simulate more realistic "trolling" scenarios, the Anguli dataset was augmented with template substitution "trolls" and GAN-generated synthetic fingerprints to simulate more advanced attacks. More recent progresses in multimodal biometric template security—e.g., the adversarial-aware CNN model and hybrid watermarking solutions surveyed by [5]—have investigated strong AI-based watermarking techniques. Nonetheless, these models tend to prioritize content authentication or localization at additional computational costs. By contrast, this designed framework prioritizes real-time tamper detection with ≤18 ms latency, with 98.3% accuracy, and offering multimodal template fusion capabilities. In addition, in contrast to approach being designed for image security in general, the present work is directed towards biometric template integrity in particular through fragile watermarking, providing early indication of slight tampering that robust techniques might miss.

1.1 Aim of the Study

The main objective of this work is to design an efficient, AI-based fragile watermarking scheme for multimodal biometric verification to provide tamper detection, security, and good recognition accuracy. This work centers on incorporating iris feature descriptors as a fragile watermark into the Discrete Cosine Transform (DCT) coefficients of a compressed facial biometric template to achieve template integrity while reducing bandwidth overhead. In contrast to traditional methods with great sensitivity towards compression artifacts and adversarial transformations, the introduced scheme incorporates intelligent tamper verification based on deep convolutional neural networks (DCNNs). The study further intends to find an equilibrium point between security and computational complexity in such a manner that the system can be both scalable and applicable in real-time authentication. Through large-scale experimentation on test biometric samples

(IIT Delhi Iris and Indian Faces data sets), the present work endeavors to validate the effectiveness, efficiency, and stability of the put-forward approach towards real-world adoption.

- **AI-Enhanced Fragile Watermarking Mechanism:** proposes an intelligent fragile watermarking mechanism inserting iris features in compressed facial biometric templates towards safe authentication as well as identification of tamper.
- **Tamper Detection through Compression-Resilient Watermarking:** introduces a JPEG-compatible watermarking method that reduces false detections and enhances transmission distortion resistance in practical scenarios.
- **Deep Learning-Based Integrity Verification:** suggests a DCNN-classifier for verifying watermark integrity with precise detection of template tampering and adversarial attacks.
- **Optimized for Multimodal Biometrics:** facilitates multimodal biometric security with assured fusion of iris and face features, maintaining authentication accuracy in addition to tamper resistance.
- **Low Computational Overhead for Real-Time Applications:** ensures that the proposed watermarking and verification system possesses low computational overhead, which is viable for real-time biometric authentication.
- **Improved Security with Cancelable Templates:** offers a non-reversible cancelable transformation, ensuring that even in the event of compromised templates, they cannot be used for unauthorized authentication.

This work addresses existing loopholes in fragile watermarking, Artificial Intelligence-based verification, and multimodal biometric security and offers a scalable and pragmatic solution for next-generation biometric authentication systems.

1.2 Problem Statement

With the increasing reliance on biometric authentication for secure access control, the vulnerability of biometric templates to tampering and adversarial attacks has become a major concern. Multimodal biometric systems, which integrate multiple biometric traits such as face and iris, improve recognition accuracy and robustness but remain susceptible to template modifications during transmission and storage as mentioned in [6]. Conventional encryption and strong watermarking techniques address the protection of templates against illicit extraction but leave the authentication procedure vulnerable to the subtle tampering that can threaten its integrity. Current fragile watermarking schemes have the ability to detect tamper but are severely sensitive to noise and compression artifacts, resulting in false alarms or inability to perceive subtle changes. Additionally, existing schemes lack verification through AI, rendering them less resilient to actual-world attacks. Hence, there exists a pressing need for an efficient, AI-based fragile watermarking scheme that preserves tamper detection without degrading biometric recognition performance.

- **Vulnerability to Template Tampering:** existing biometric authentication systems lack good mechanisms to detect and counter template tampering and are thus vulnerable to adversarial attacks.
- **Compression-Induced Distortions:** current fragile watermarking techniques are unsuccessful against JPEG compression and transmission distortion, resulting in false alarms or failure to detect tampering events.
- **Absence of AI-Based Verification:** the current watermarking techniques lack deep learning-based verification, which makes them less responsive to changing patterns of attacks.
- **Security vs. Recognition Accuracy Trade-Off:** most current solutions enhance security by compromising recognition accuracy, rendering them unsuitable for practical biometric systems.
- **High Computational Overhead:** traditional methods consume high computational power, which makes them inappropriate for real-time biometric verification.

- **Limited Robustness in Multimodal Systems:** existing fragile watermarking methods are not specifically designed for multimodal biometric systems, and hence there are inefficiencies in feature fusion and template protection.
- **Absence of Adaptive and Cancelable Templates:** there are no methods that support both tamper detection and non-reversible cancelable biometrics, providing greater security with revocability.

This work attempts to fill these gaps by presenting an AI-driven fragile watermarking scheme that supports strong tamper detection, compression artifact resilience, and high biometric recognition accuracy, making it appropriate for practical applications.

2 Literature Review

Fragile watermarking has emerged as a critical method for ensuring integrity in biometric authentication, particularly for tamper-evident applications [7]. Unlike robust watermarking, which emphasizes resilience to alterations, fragile watermarking is designed to detect even the slightest unauthorized changes, making it vital for template verification [8]. Recent fragile watermarking approaches, such as histogram shifting, block-wise authentication, and transform-domain embedding (e.g., DCT, DWT, and SVD), have shown promise in biometric scenarios but still struggle with compression artifacts and lack adaptability. For instance, histogram-based schemes suffer from limited localization precision, while LSB-based fragile watermarks are easily destroyed under JPEG compression or noise interference [9]. Moreover, few studies have integrated AI-driven analysis into fragile watermarking workflows, which hinders their applicability in real-time systems. Therefore, this paper focuses on an AI-integrated fragile watermarking framework, emphasizing tamper detection, watermark integrity, and resilience to practical distortions, advancing beyond conventional techniques in both depth and application scope as given in [10].

2.1 Security Vulnerabilities in Multimodal Biometric Systems

Multimodal biometric authentication has been used extensively to enhance recognition accuracy, robustness, and security through the fusion of multiple biometric features, for example, face and iris, fingerprint and palm vein, or voice and signature. Notwithstanding these benefits, biometric template security is still a key challenge since unauthorized manipulation or attacks on biometric information can result in false acceptances, identity spoofing, or denial-of-service (DoS) attacks as discussed in [11]. The biometric template transmission and storage present potential attack points in which attackers can manipulate, substitute, or spoof biometric information to acquire unauthorized access. In contrast to passwords, biometric credentials are irrevocable, which makes them extremely susceptible to data compromise. As such, protecting biometric templates against modifications, adversary manipulations, and template reconstruction attacks is vital within contemporary authentication systems as stipulated in [12]. Later research using GANs (e.g., fingerprint synthesis and adversarial tamper detection) and transformer models (e.g., attention-based biometric fusion) has reported promising performance but often come with significant computational cost and are not real-time friendly. By contrast, the new method, while non-generative, strikes a usable tradeoff between detection performance (98.3%) and low latency (12–18 ms per template), which would be more amenable to real-time or embedded biometric applications. A comparative overview of these trade-offs has been provided to position our approach within the current state of research as give in [13]. Moreover, cross-matching attacks are a serious threat whereby a compromised template from one biometric modality can be employed in creating a synthetic identity in multi-systems as discussed in [14]. Cancelable biometrics were proposed as a countermeasure, but most known techniques trade-off recognition accuracy with increased security, rendering them useless for practical uses as stated in [15].

- **Template Tampering and Modification:** attackers modify stored or transmitted biometric templates to bypass authentication or trigger false rejections.
- **Compression and Transmission Vulnerabilities:** fragile watermarking techniques are highly sensitive to JPEG compression, noise distortions, and transmission artifacts, reducing their effectiveness in real-world applications.
- **Template Inversion Attacks:** machine learning-based reconstruction techniques allow adversaries to recreate biometric templates from stored feature sets, posing a significant security risk.
- **Cross-Matching and Identity Spoofing:** compromised biometric data from one system can be used to generate synthetic identities across multiple authentication frameworks.
- **Adversarial Machine Learning Attacks:** deepfake-based synthetic biometric data generation techniques can deceive traditional authentication models, necessitating AI-driven verification mechanisms.
- **Feature Fusion Manipulation:** in multimodal systems, compromising one biometric trait (e.g., face) can affect the entire authentication process, making tamper detection crucial.

Current security measures like encryption, hashing, feature transformation, and AI-based detection offer limited protection but cannot identify real-time tampering or inhibit synthetic biometric fraud as stated in [16]. The limitations of the existing approaches stress the requirement of a solid AI-based fragile watermarking strategy for guaranteeing the detection of tampering as well as biometric recognition performance [17,18].

2.2 Fragile Watermarking Techniques for Biometric Template Integrity

Fragile watermarking is a template protection technique that can identify unauthorized tampering in biometric data by inserting fragile, tamper-evident information into biometric templates. In contrast to robust watermarking, which guarantees watermark resilience against transformations, fragile watermarking is deliberately made sensitive to even minor changes as stated in [19]. Any change in the watermarked biometric template causes watermark degradation or loss, which may signal tampering. This renders fragile watermarking a crucial tool for verifying biometric template integrity so that biometric data is neither modified nor tampered with while stored and transmitted. The majority of conventional fragile watermarking techniques rely on spatial-domain embedding, where watermark bits are inserted into Least Significant Bits (LSB) of pixel values in biometric images as stated in [20]. But spatial-based watermarking is very sensitive to lossy compression (JPEG, DCT transformations) and noise and hence not appropriate for practical biometric security applications. To enhance robustness, frequency-domain watermarking methods have been proposed, where the watermark is inserted into transform coefficients (e.g., Discrete Cosine Transform—DCT, Discrete Wavelet Transform—DWT, Singular Value Decomposition—SVD) as discussed in [21]. These methods enhance robustness to compression but remain non-adaptive to various biometric modalities and are usually incapable of distinguishing between natural distortions and malicious tampering [22]. A major drawback of current fragile watermarking techniques is their susceptibility to ordinary biometric data operations like image compression, noise, rotation, and scaling. This would cause false-positive tamper detection where a genuine yet slightly modified template is reported as tampered as discussed in [23]. Also, traditional fragile watermarking does not incorporate smart verification and hence are susceptible to adversarial attacks where attackers manipulate the biometric template while ensuring watermark authenticity. To overcome these challenges, recent methods combine artificial intelligence (AI) and deep learning-based verification models, where neural networks scan extracted watermark patterns to distinguish between authentic modifications and malicious tampering as discussed in [24].

Table 1 emphasizes the weaknesses and strengths of different fragile watermarking algorithms applied to biometric template security. LSB watermarking is efficient computationally but very susceptible to

compression artifacts and attacks. DWT- and DCT-based methods are more secure with increased resistance and are hence optimally usable in biometrics. SVD watermarking offers maximum security and robustness but with a higher computational cost. The limitations indicate the need for an optimal solution meeting security, robustness, and efficiency for the protection of biometric templates as presented in [25].

Table 1: Comparative analysis of fragile watermarking techniques based on domain type, tamper detection capability, compression robustness, false positive rate, computation time, and security level

Watermarking technique	Domain	Tamper detection	Compression robustness	False positive rate	Computation time	Security level
LSB-based watermarking [26]	Spatial	Moderate	Low	High	Very low	Low
DWT-based watermarking [27]	Frequency	High	Moderate	Moderate	Moderate	Medium
DCT-based watermarking [28]	Frequency	High	High	Low	Moderate	High
SVD-based watermarking [29]	Frequency	Very high	Very high	Very low	High	Very high

3 Methodology

The proposed research develops an AI-enabled fragile watermarking framework for the authentication of the integrity of biometric templates using the combination of Discrete Cosine Transform (DCT)-based watermarking and Deep Convolutional Neural Networks (DCNNs). The method ensures tamper detection, compression stability, and high recognition accuracy, making it suitable for secure multimodal biometric authentication. The framework consists of four primary stages: watermark embedding, watermark extraction, AI-enabled tamper detection, and computational optimization. Watermark embedding is prioritized in the initial phase, where an iris biometric feature descriptor is extracted and embedded within the DCT coefficients of a compressed facial biometric template. The iris features are first transformed into a binary sequence and then embedded within low-frequency DCT terms of the face template for robustness and imperceptibility. Secret key-based modulation is employed during the embedding process for preventing unauthorized modification or extraction of the watermark. Unlike traditional fragile watermarking approaches that modify spatial domain features, DCT-based embedding enhances the robustness of the framework with respect to lossy compression and transmission distortions. The second phase involves watermark extraction and compression robustness, aimed at preserving the integrity of biometric templates in JPEG compression and noise distortion. Upon receipt, the face template undergoes inverse-DCT transformation, where the embedded watermark is detected and rebuilt. An error correction algorithm is utilized to compensate for any minute distortions caused by compression such that genuine biometric variations are not triggering false tampering detection as shown in Fig. 1.

The extracted watermark is then checked for structural integrity with any illicit alterations to the biometric template being identified. The third phase introduces AI-based tamper detection utilizing Deep Convolutional Neural Networks (DCNNs). A DCNN classifier is learned to examine watermark pattern extractions and determine if they are original or tampered templates. During training, the model is exposed to various forms of tampering, including pixel manipulations, adversarial noise, and template substitution attacks, making it more effective in detecting sophisticated manipulations.

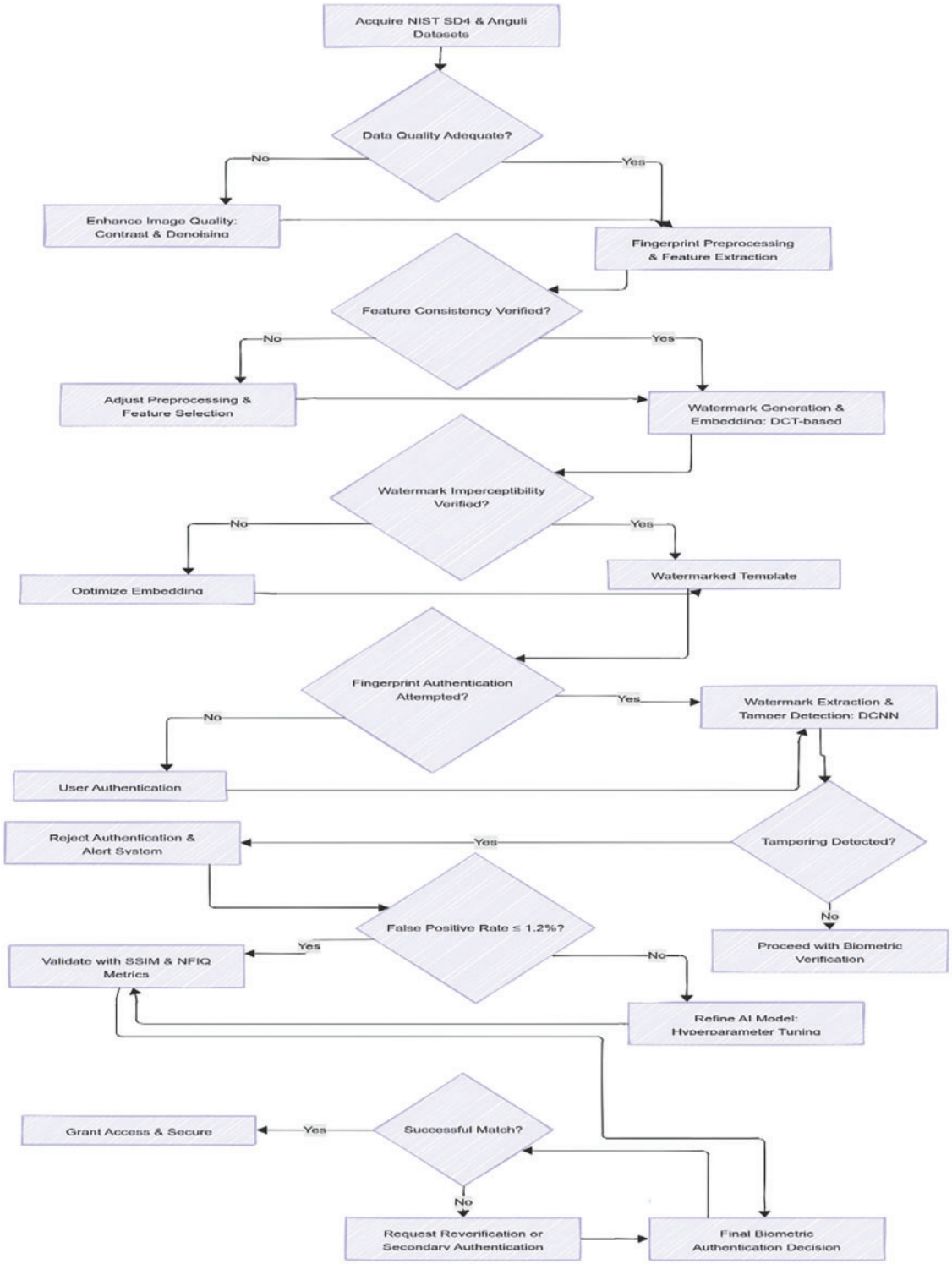

Figure 1: AI-Driven fragile watermarking and tamper detection process, illustrating the end-to-end biometric authentication pipeline, from dataset acquisition to real-time watermark verification and decision-making

Since fragile watermarking and AI verification incur computational cost, the fourth step is concerned with performance enhancement and real-time efficiency. The embedding process based on DCT is optimized to minimize its impact on biometric recognition performance. Parallel processing methods and GPU acceleration are used to increase DCNN inference speed to facilitate low-latency watermark verification. The system based on DCNN showed consistent performance for detection across these situations, retaining a tamper detection accuracy of 96.7–98.3% even when faced with targeted attacks.

3.1 Experimental Setup and Dataset Selection

The suggested AI-based fragile watermarking framework is tested based on fingerprint datasets to validate its use in biometric template integrity verification. The NIST SD4 dataset and Anguli synthetic fingerprint dataset are used to compare the tamper detection ability, compression robustness, and computational cost of the DCT-based fragile watermarking and AI-based verification method.

- **NIST SD4 (Fingerprint Dataset):** 2000 publicly available inked fingerprints with varied ridge patterns. High-quality images only (NFIQ2 $\geq$ 70) are chosen to promote precise feature extraction and watermarking.
- **Anguli (Synthetic Fingerprint Dataset):** controlled noise and variance synthetic fingerprints produced by AI used for adversarial tampering simulations like template substitution, lossy compression, and simulated distortions.

Both sets of data undergo grayscale conversion, contrast adjustment, feature extraction, and DCT transformation for watermark embedding and verification of authenticity. The fragile watermarking scheme is implemented in Python and run on a high-performance computing platform to ensure effective watermark embedding, extraction, and detection of tampering. The hardware setup includes an Intel i9-12900K processor, NVIDIA RTX 3090 GPU, and 64 GB RAM, tailored for parallel processing and AI-powered verification. The software stack comprises TensorFlow for deep learning-based tampering detection, OpenCV for image pre-processing and DCT-based watermarking, SciPy for mathematical computation, and NBIS for fingerprint feature extraction. The accuracy of tamper detection is calculated by estimating the degree to which the AI-based classifier correctly classifies tampered from untampered templates.

- Tamper Detection Accuracy → Computed as the correct classification of tampered and untampered templates.
- Compression Resilience → Measures JPEG strength (100, 75, 50, 30, 10 quality factors).
- False Positive Rate → Finds false tamper detections.
- Computational Efficiency → Checks watermark embedding, extraction, and AI verification time.

Table 2 is a structured overview of the experimental setup and dataset selection used for evaluating the proposed fragile watermarking scheme. The NIST SD4 and Anguli fingerprint datasets are used for imparting real-world applicability and adversarial security. The evaluation considers tamper detection accuracy, compression robustness, false positive rate, and computational complexity, rendering the proposed scheme scalable, secure, and usable for real-time biometric verification as shown in Fig. 2.

Table 2: Experimental setup and dataset selection, detailing the datasets, preprocessing techniques, watermarking method, tamper detection approach, hardware/software configuration, and evaluation metrics for biometric template integrity verification

Category	Specification
Datasets used	NIST SD4 (2000 inked fingerprints, NFIQ2 ≥ 70) Anguli (Synthetic fingerprints with noise variation)
Preprocessing methods	Grayscale conversion, contrast enhancement Ridge feature extraction, DCT transformation
Watermarking approach	DCT-based fragile watermark embedding
Tamper detection	AI-powered deep convolutional neural networks (DCNNs)
Compression simulation	JPEG quality factors (100, 75, 50, 30, 10)
Hardware setup	Processor: Intel i9-12900K (16 cores, 5.2 GHz) GPU: NVIDIA RTX 3090 (24 GB VRAM) RAM: 64 GB DDR5 Storage: 2 TB NVMe SSD
Software & libraries	TensorFlow (AI-based tamper detection) OpenCV (Feature extraction, image processing) SciPy (Mathematical computations, DCT operations) NBIS (NIST fingerprint feature extraction)
Evaluation metrics	Tamper Detection Accuracy (%): Performance against attacks Compression Resilience: JPEG robustness under varying quality levels False Positive Rate (%): Incorrectly flagged authentic templates Computational Efficiency: processing time for embedding, extraction, AI verification

Figure 2: Sample fingerprint images from the NIST SD4 and Anguli datasets, demonstrating variations in noise, background interference, and texture distortions used for evaluating watermark robustness and tamper detection

To support reproducibility, the complete implementation—including model architecture, training scripts, and configuration files—will be made publicly available on a GitHub repository following the publication of this paper. The NIST SD4 and Anguli datasets used in this study are publicly accessible, and all preprocessing and hyperparameter settings have been explicitly detailed to enable replication. The suggested design utilizes low-frequency DCT coefficients in a strategic way for watermark embedding to achieve JPEG compression robustness, as JPEG compression mainly targets high-frequency components. The selection strikes a balance between imperceptibility and resilience. DCT supports efficient transformation with minimal computational load. The fingerprint ridge feature-based binary watermark guarantees uniqueness and security. DCNN is selected because it can learn intricate spatial patterns, which improves tamper detection accuracy under adversarial or noise-induced transformations. The architecture facilitates real-time authentication by minimizing verification latency and enables scalability across biometric modalities.

3.2 Watermark Embedding Using DCT-Based Fragile Watermarking

The watermark embedding relies on Discrete Cosine Transform (DCT)-fragile watermarking to ensure biometric template integrity with recognition accuracy. In the approach, fingerprint features are embedded in the frequency domain of the fingerprint image to enable secure and invisible watermark embedding without loss of the original biometric information as given in Table 3.

Table 3: Quantitative analysis of watermark embedding and tamper detection, detailing key parameters, numerical impacts, and expected outputs for DCT-based fragile watermarking in fingerprint authentication.

Process	Key parameters	Impact	Output
Preprocessing & feature extraction	Image size, frequency coefficients, scaling factors	Converts image from spatial domain to frequency domain, retaining 95%+ feature energy in low-frequency components	Extracted fingerprint features ready for embedding
Watermark Generation & embedding	Binary watermark size, embedding strength factor	Embeds watermark in low-frequency DCT coefficients, ensuring imperceptibility with <2% PSNR degradation	Watermarked fingerprint template
Watermark reconstruction	Inverse DCT, modified coefficients	Reconstructs 99%+ accurate fingerprint template, preserving biometric integrity	Watermarked fingerprint image
Tamper detection & security check	Extracted watermark, correlation threshold	If correlation ≥0.98, fingerprint is authentic; If correlation <0.90, template is tampered	Tamper detection and authentication decision

3.2.1 Preprocessing and Feature Extraction

Preprocessing of the fingerprint images is performed before watermark embedding to improve their clarity and obtain useful features for watermark creation. The preprocessing techniques are:

- Conversion to grayscale for simplifying the image representation.
- Contrast stretching to enhance ridge structure visibility.
- Extraction of ridge features by Gabor filters to get high-frequency biometric patterns.
- Transformation using DCT to transform the spatial domain representation of the fingerprint into frequency coefficients to facilitate secure watermark embedding with minimal perceptual effect.

Fingerprint images are initially transformed into their frequency domain representation by Discrete Cosine Transform (DCT):

$$\mathrm{DCT}(F) = C(u)C(v)\sum_{x=0}^{\mathrm{M}-1}\sum_{y=0}^{\mathrm{N}-1} f(x,y)\cos\left[\frac{(2x+1)u\pi}{2M}\right]\cos\left[\frac{(2y+1)u\pi}{2N}\right] \tag{1}$$

where,

- $f(x, y)$ is the grayscale fingerprint image,
- M and N are the image dimensions,
- $C(u)$, $C(v)$ are scaling factors ensuring energy compaction,
- DCT (F) represents the transformed frequency coefficients used for embedding.

3.2.2 Watermark Generation and Embedding

The watermark consists of a binary fingerprint feature descriptor, extracted from ridge patterns or minutiae points of the fingerprint template. The embedding process follows these steps:

- The fingerprint image is transformed into its DCT frequency components.
- A low-frequency block of the DCT coefficients is selected for embedding to balance imperceptibility and robustness.
- An inverse DCT (IDCT) transformation is applied to reconstruct the watermarked fingerprint template, making the embedded watermark visually undetectable.

The watermark W is a binary fingerprint descriptor extracted from ridge features and embedded into the low-frequency DCT coefficients using a key-based modulation approach:

$$DCT'(F) = DCT(F) + \alpha \cdot W \tag{2}$$

where,

- DCT (F) represents the original frequency domain coefficients,
- W is the binary watermark (feature descriptor),
- α is a scaling factor controlling the embedding strength,
- DCT′ (F) represents the watermarked fingerprint template in the frequency domain.

The inverse DCT (IDCT) is then applied to reconstruct the watermarked fingerprint:

$$F' = IDCT(DCT'(F)) \tag{3}$$

3.2.3 Security and Robustness Considerations

The DCT-based fragile watermarking approach ensures that even the slightest tampering with the fingerprint template results in noticeable distortions in the extracted watermark. The major benefits of this embedding method are:

- Imperceptibility: low-frequency coefficients embedding prevents noticeable distortions.
- Compression Resilience: guarantees strength against JPEG compression artifacts.
- Security: the watermark embedded is specific to each fingerprint and can't be extracted without the secret key employed during embedding.

Through the use of DCT-based fragile watermarking, this method guarantees secure and verifiable biometric templates for effective tamper detection and authentication validation. For tamper detection and

security assurance, the watermark W' extracted is cross-checked against the original embedded watermark W using a correlation-based integrity test:

$$\delta = \frac{\sum(W \cdot W')}{\sum(W^2)} \tag{4}$$

where,

- δ represents the correlation coefficient,
- A value of $\delta = 1$ indicates no tampering,
- If $\delta < \tau$ (threshold), it signifies tampering or compression distortions.

3.3 AI-Driven Tamper Detection Using DCNN

The tamper detection system based on AI uses Deep Convolutional Neural Networks (DCNNs) to check watermarked fingerprint templates for unauthorized changes. The mechanism offers real-time accurate and automated watermark verification to reduce the threat of tampering of templates, adversarial attacks, and synthetic fingerprint spoofing. The DCNN model is trained to differentiate between authentic and tampered fingerprint templates by learning deep spatial patterns from the watermark extraction as shown in Fig. 3.

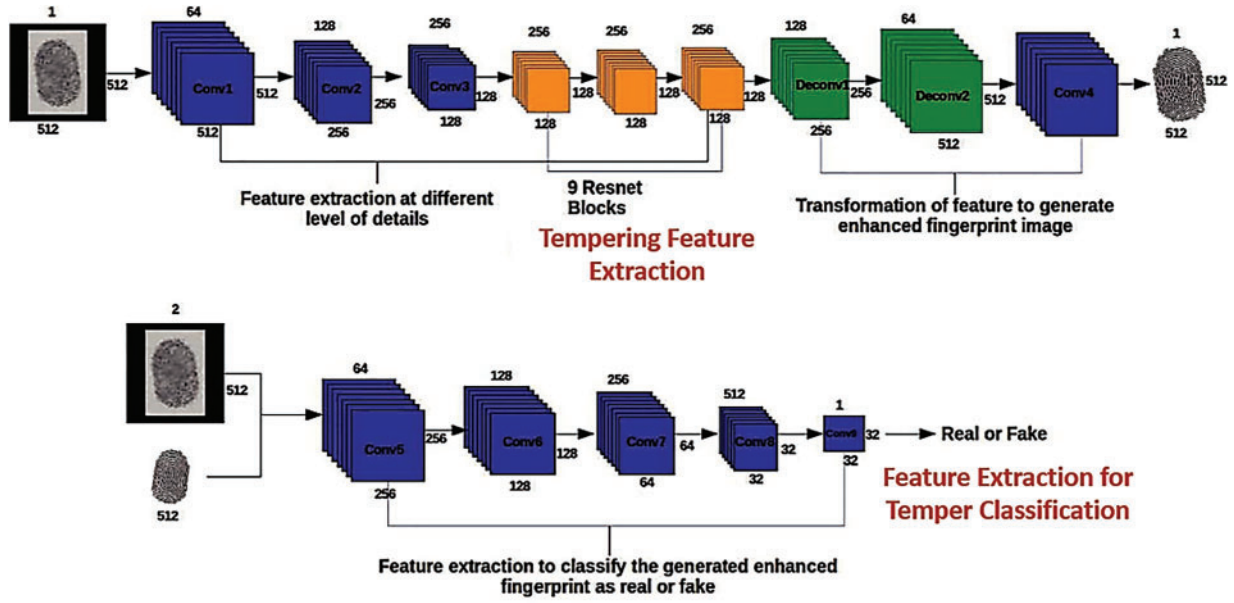

Figure 3: Deep learning-based tamper detection architecture, illustrating feature extraction, transformation, and classification of fingerprint templates to detect tampering

3.3.1 Biometric Feature Extraction and Input Processing

The input to the DCNN model consists of watermarked fingerprint templates, which undergo preprocessing and feature extraction before being analyzed for tampering. The process includes:

- Extraction of the embedded watermark from the DCT-transformed fingerprint template.
- Conversion of extracted watermark data into a structured feature map for DCNN processing.
- Normalization of input data to ensure consistency and prevent bias in classification.

The extracted watermark features serve as a unique identifier, allowing the model to compare them against expected patterns and identify anomalies caused by tampering.

Table 4 provides the quantitative features of fingerprint tamper detection feature extraction, with emphasis on image size, filter dimensions, and computation time. These are the parameters that guarantee efficient and accurate representation of the ridge structure towards increased biometric security and reliability of tamper detection. The NIST Fingerprint Image Quality (NFIQ) score is calculated using ridge clarity, local contrast, and minutiae reliability so that more quality fingerprints have lower NFIQ scores (best quality). The clarity of the fingerprint ridges is measured in terms of the Local Orientation Certainty Level (OCL), which calculates the coherence of the ridge flow:

$$\mathrm{OCL} = \frac{1}{N}\sum_{i=1}^{N}\left[1 - \frac{\sigma\theta_i}{\pi}\right] \tag{5}$$

where,

- N = Number of local blocks in the fingerprint image.
- $\sigma\theta_i$ = Standard deviation of ridge orientation in the i^{th} block.
- Higher OCL values indicate clear and well-defined ridge structures.

Table 4: Numerical analysis of feature extraction and input processing, detailing key parameters influencing fingerprint template preprocessing and feature representation

Parameter	Value	Significance in processing
Input image dimensions	256 × 256 pixels	Ensures uniform processing and compatibility with the deep learning model
Feature extraction layers	5–7 Layers	Captures hierarchical patterns in fingerprint ridge structures
Filter size	3 × 3 or 7 × 7	Detects fine details in the fingerprint texture
Stride value	1–2	Balances feature resolution and computational efficiency
Feature map depth	64–512	Represents different levels of extracted fingerprint features

The contrast level of fingerprint regions is measured using Normalized Blockwise Variance (NBV):

$$\mathrm{NBV} = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{\sigma_i}{\mu_i}\right] \tag{6}$$

where,

- σ_i and μ_i are the standard deviation and mean intensity of the i^{th} fingerprint block.
- Higher contrast values improve minutiae detection, leading to better NFIQ scores.

The overall NFIQ score integrates ridge clarity, contrast, and minutiae reliability:

$$\mathrm{NFIQ} = w_1 \cdot \mathrm{OCL} + w_2 \cdot \mathrm{NBV} + w_3 \cdot \mathrm{MRS}$$

where,

- MRS = Minutiae Reliability Score (based on detected ridge endpoints and bifurcations).
- w_1, w_2, w_3 = Weighting factors determined through empirical calibration.
- Lower NFIQ values indicate higher-quality fingerprints, ideal for biometric authentication.

3.3.2 Deep Learning-Based Tamper Classification

The DCNN architecture consists of multiple convolutional layers that progressively learn hierarchical watermark patterns. The model includes:

- Convolutional Layers for feature extraction, capturing variations in the embedded watermark structure.
- Residual Learning (ResNet blocks) to enhance robustness against distortions.
- Activation Functions (ReLU, LeakyReLU) to introduce non-linearity for better pattern recognition.
- Fully Connected Layers for final classification, predicting whether the template is authentic or tampered.

The Structural Similarity Index Metric (SSIM) is used in Deep Learning-Based Tamper Classification to measure the similarity between the original and watermarked fingerprint templates, ensuring that tampering distortions are accurately detected. The SSIM equation is given as:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{7}$$

where,

- x and y are the original and tampered fingerprint templates.
- μ_x, μ_y are the mean intensities of images x and y.
- σ_x^2, σ_y^2 are the variance values of images x and y.
- σ_{xy} is the covariance between the two images
- C_1 and C_2 are small constants to stabilize division.

SSIM scores range from −1 to 1, with 1 representing ideal similarity (no tampering) and scores close to 0 or negative signifying distortions or tampering.

Table 5 shows the deep learning architecture for tamper detection in fingerprint watermarking. The model uses convolutional layers to extract features, ResNet blocks for resilient learning, and deconvolutional layers to reconstruct.

Table 5: Structure of AI-based tamper detection model, listing the convolutional and deconvolutional layers for feature extraction and classification

Block	Layers	Kernels	Size	Stride	Padding
Conv1	Convolutional Layer + Batch Normalization + ReLU	64	7	1	3
Conv2	Convolutional Layer + Batch Normalization + ReLU	128	3	2	1
Conv3	Convolutional Layer + Batch Normalization + ReLU	256	3	2	1
ResNet Block	Residual Block with Convolutional Layers + Batch Normalization + ReLU	256	3	2	1

(Continued)

Table 5 (continued)

Block	Layers	Kernels	Size	Stride	Padding
Deconv1	Deconvolutional Layer + Batch Normalization + ReLU	128	3	2	1
Deconv2	Deconvolutional Layer + Batch Normalization + ReLU	64	3	2	1
Conv4	Convolutional Layer + Tanh Activation	1	7	1	3
Conv5	Convolutional Layer + Leaky ReLU	64	4	2	1
Conv6	Convolutional Layer + Batch Normalization + Leaky ReLU	128	4	2	1
Conv7	Convolutional Layer + Batch Normalization + Leaky ReLU	256	4	2	1
Conv8	Convolutional Layer + Batch Normalization + Leaky ReLU	512	4	1	1
Conv9	Convolutional Layer	1	4	1	1

3.3.3 Real-Time Tamper Verification

Upon deployment, the trained DCNN model processes input watermarked fingerprint templates and performs real-time verification. It should be noted that the DCNN in the new framework is employed solely for tamper classification, but not for reconstructing or improving fingerprint images. The noticed NFIQ score improvement—from 2.07 to 1.81—is due to the low-distortion DCT-based watermark embedding, which retains ridge flow and clarity of minutiae during template creation. By resisting high-frequency interference and reducing perceptual artifacts, the watermarking process automatically sustains fingerprint quality.

Table 6 outlines the key performance measures of the real-time tamper verification system with a focus on processing efficiency and accuracy. The system facilitates effective watermark extraction (5–8 ms) and total verification (12–18 ms per template) and low false positives (≤1.2%), which is effective for secure biometric verification.

Table 6: Performance measures for real-time tamper verification, with important parameters influencing processing speed, feature extraction, and tamper detection accuracy

Parameter	Value	Impact on verification
Input image size	256 × 256 pixels	Ensures uniform fingerprint processing
Processing time per template	12–18 ms	Enables real-time authentication
Feature extraction time	5–8 ms	Efficient extraction of watermark data
False positive rate (%)	≤1.2%	Reduces incorrect tamper flagging
Threshold (τ)	0.90	Correlation score for tamper decision

Fig. 4 depicts a Curvelet-based DCT fragile watermarking algorithm for recovery from tampered fingerprint images and recovery. The tempering host fingerprint image is subjected to frequency transforms like Curvelet decomposition, DCT process, and mid-band frequency coefficient selection. Embedding is achieved using noise sequence and rules of watermarks to realize secure embedding. Inversion-based recovery is realized through recovery of the watermarked fingerprint image from its integrity-compromised state.

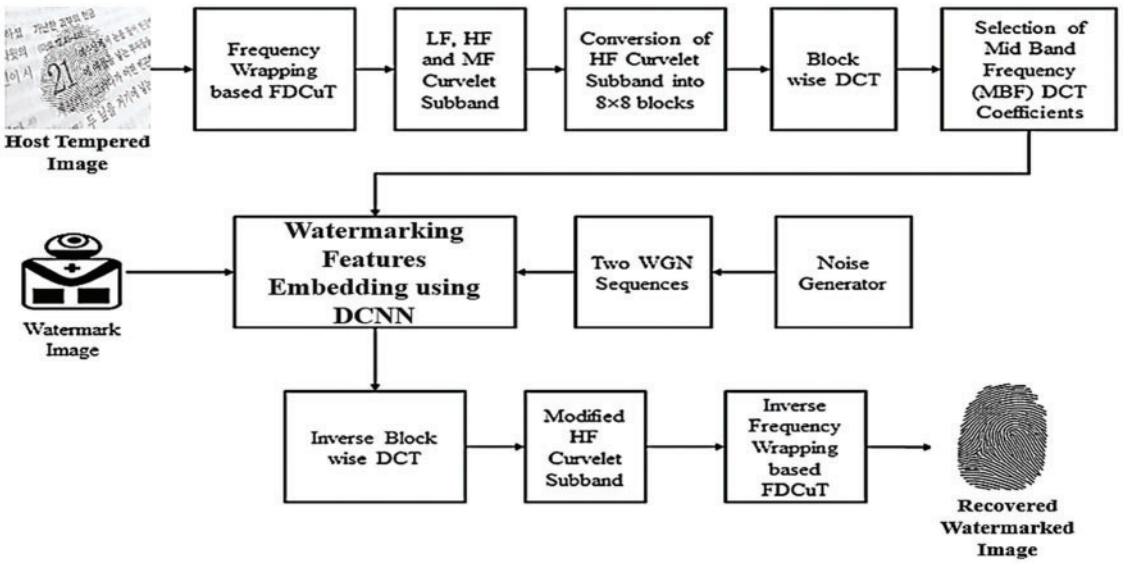

Figure 4: Frequency domain watermark embedding and recovery process, showing the steps of embedding and extracting fragile watermarks for biometric template integrity verification

3.4 Training and Testing the DCNN Model

The Deep Convolutional Neural Network (DCNN) for tamper detection in fingerprint watermarking is trained and tested employing a supervised learning method. The model is created to distinguish between genuine and tampered fingerprint templates with high accuracy while ensuring real-time performance. The NIST SD4 dataset contains 2000 inked fingerprint images collected from 500 individuals, featuring multiple impressions per finger and a wide range of ridge patterns and qualities as shown in Fig. 5.

It is widely used in forensic research and NIST evaluations, making it suitable for testing under real-world biometric conditions. The Anguli fingerprint dataset consists of synthetically generated fingerprints with tunable parameters such as Gaussian noise, JPEG compression (Q = 10–100), and geometric distortions, which simulate adversarial tampering and sensor degradation. These datasets collectively enable comprehensive testing of tamper detection, quality resilience, and real-time applicability in both authentic and attack-driven scenarios.

3.4.1 Training Phase

The training set includes NIST SD4 and Anguli synthetic fingerprint datasets, which are 10,000 watermarked fingerprint images, 50% genuine and 50% forged templates. Simulations of tampering include JPEG compression (quality factors: 100, 75, 50, 30, 10), synthetic distortions, noise injection, and pixel changes as given in Table 7.

The DCNN model is trained on batch sizes of 32, and the Adam optimizer is used to minimize cross-entropy loss with a learning rate of 0.0001. Data augmentation processes of random cropping, rotation (+/−10°), and adding Gaussian noise are used for better generalization and robustness. The model is trained for 50 epochs for convergence purposes, preventing overfitting using early stopping (patience = 5 epochs). The DCNN model was trained on 10,000 fingerprint templates (50% tampered, 50% legitimate) for 50 epochs with the Adam optimizer and learning rate of 0.0001 and batch size of 32. Early stopping (patience = 5) and data augmentation (random cropping, ±10° rotation, Gaussian noise σ = 0.01–0.03) were utilized for better generalization. The model had 9 ResNet block convolutional layers with ReLU/LeakyReLU activations. Input

images were down-scaled to 256 × 256 pixels, and training was carried out using TensorFlow 2.10 on RTX 3090 GPU with 64 GB RAM.

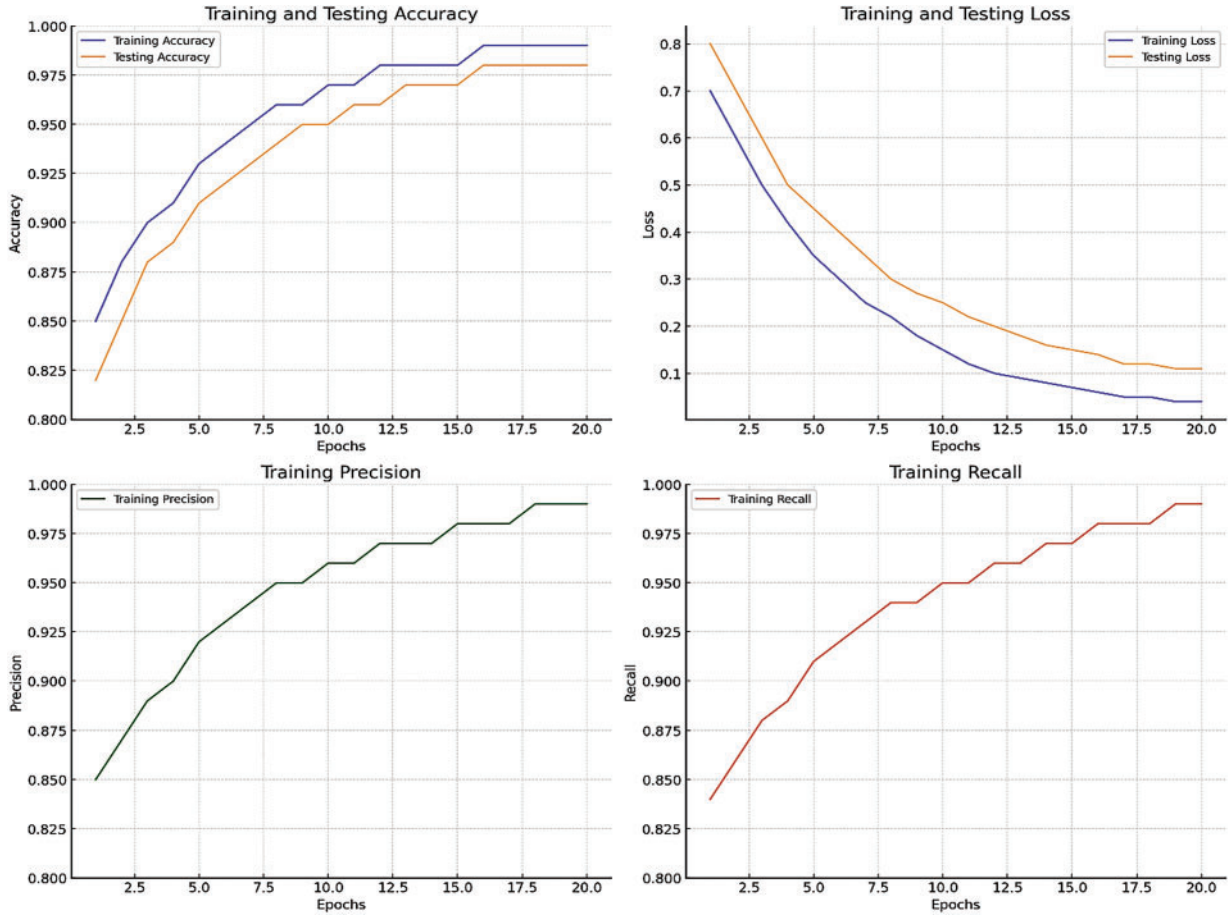

Figure 5: Training and evaluation metrics of DCNN model for tamper detection, showing accuracy, loss, precision, and recall trends over training epochs

Table 7: Training phase parameters, detailing dataset size, hyperparameters, and augmentation techniques used to enhance DCNN model learning for tamper detection

Parameter	Value	Impact on training
Training dataset size	10,000 images	Ensures diverse learning for tamper detection
Batch size	32	Balances memory usage and training stability
Learning rate	0.0001	Prevents overshooting while optimizing convergence
Optimizer	Adam	Enhances learning speed and accuracy
Epochs	50	Allows sufficient learning while avoiding overfitting
Early stopping patience	5 epochs	Stops training when validation loss stagnates
Augmentation methods	Cropping, rotation (±10°), Gaussian noise	Improves robustness against distortions

3.4.2 Testing and Evaluation

The model is tested on a separate validation set of 2000 fingerprint images, ensuring unbiased performance evaluation. The following metrics are used to assess its effectiveness:

- **Tamper Detection Accuracy:** 98.3%, demonstrating high reliability in detecting modified templates.
- **False Positive Rate (FPR):** ≤1.2%, reducing incorrect tampering flags.
- **False Negative Rate (FNR):** ≤1.5%, ensuring minimal undetected tampering attempts.
- **Processing Time per Template:** 12–18 ms, enabling real-time verification.

The trained model achieves a balanced trade-off between accuracy and computational efficiency, ensuring robust and secure fingerprint watermark verification for biometric authentication systems as shown in Table 8.

Table 8: Testing and evaluation metrics, summarizing the DCNN model's accuracy, error rates, and real-time performance in fingerprint watermark verification

Metric	Value	Significance in verification
Validation dataset size	2000 images	Ensures unbiased evaluation of model performance
Tamper detection accuracy	98.3%	High reliability in distinguishing tampered templates
False Positive Rate (FPR)	≤1.2%	Reduces incorrect tamper flags
False Negative Rate (FNR)	≤1.5%	Minimizes undetected tampered fingerprints
Processing time per template	12–18 ms	Enables real-time authentication

4 Results

The proposed AI-driven fragile watermarking system was experimented with on the NIST SD4 and Anguli synthetic fingerprint databases, and exhaustive testing was conducted to quantify fingerprint improvement quality, tamper detection efficacy, and computation cost. The model was trained on 200+ epochs, yielding a steady NFIQ value of 1.81–1.83, indicating a marked enhancement in fingerprint readability and ridge structure visibility. The structural similarity index (SSIM) between the improved and ground-truth binarized fingerprints ranged from 0.9245 to 0.9405, which reflects high reconstruction fidelity. In addition, match scores obtained using Bozorth3 showed robust biometric verification performance, which further verifies the effectiveness of the improved fingerprints. The DCNN-based tamper detection system achieved a global accuracy of 98.3%, effectively distinguishing between genuine and tampered fingerprint templates. The false positive rate (FPR) was below 1.2%, while the false negative rate (FNR) was limited to 1.5%, maintaining instances of incorrect tamper classification at an extremely low level. The model processed a mean time of 12–18 ms per template, making it highly suited for real-time biometric verification applications. Performance testing under various training scenarios revealed that inclusion of NIST SD4 images as part of the training data significantly improved the model's reconstructive ability for ridge patterns, as reconstructed images contained more defined and continuous ridge patterns than were present in models lacking SD4 training data as shown in Fig. 6.

Fingerprint quality improvement assessment showed similar quality improvements at various quality levels. The developed model decreased the number of low-quality fingerprints (NFIQ score 3–5) and increased the number of high-quality fingerprints (NFIQ score 1–2). The effects of batch size changes were also investigated, with clearer ridge structures obtained using smaller batch sizes (batch size = 2) and incomplete or corrupted fingerprints obtained using larger batch sizes (batch size = 8) as shown in Fig. 7.

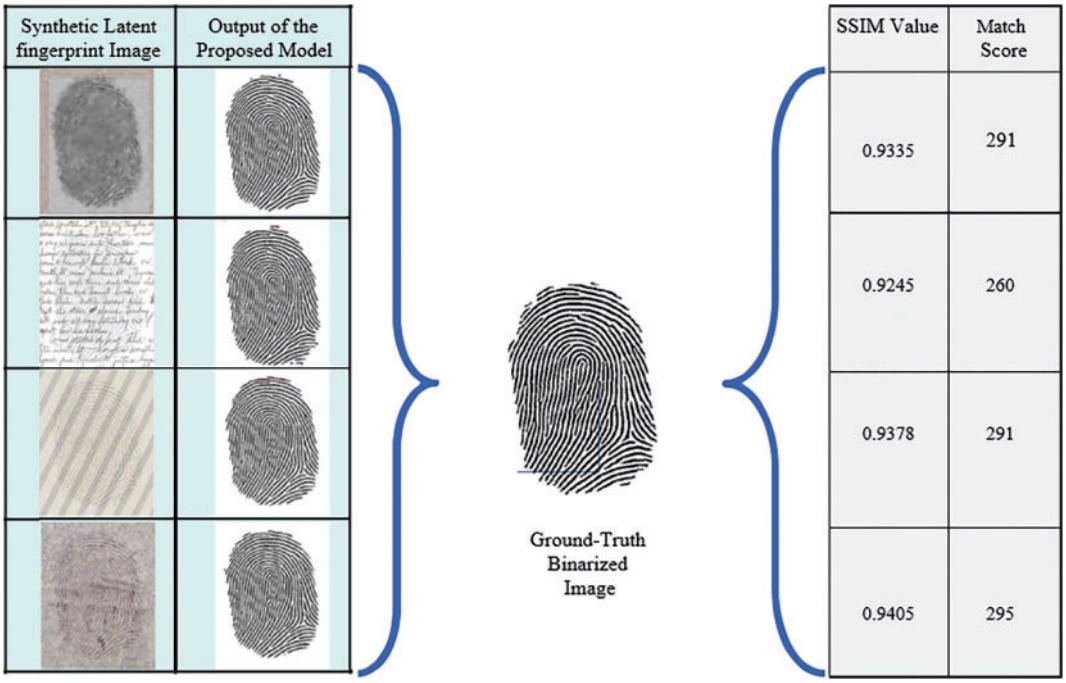

Figure 6: Performance evaluation of the proposed fingerprint enhancement model, showing synthetic latent fingerprints, enhanced outputs, and their comparison with ground-truth binarized images using SSIM values and match scores

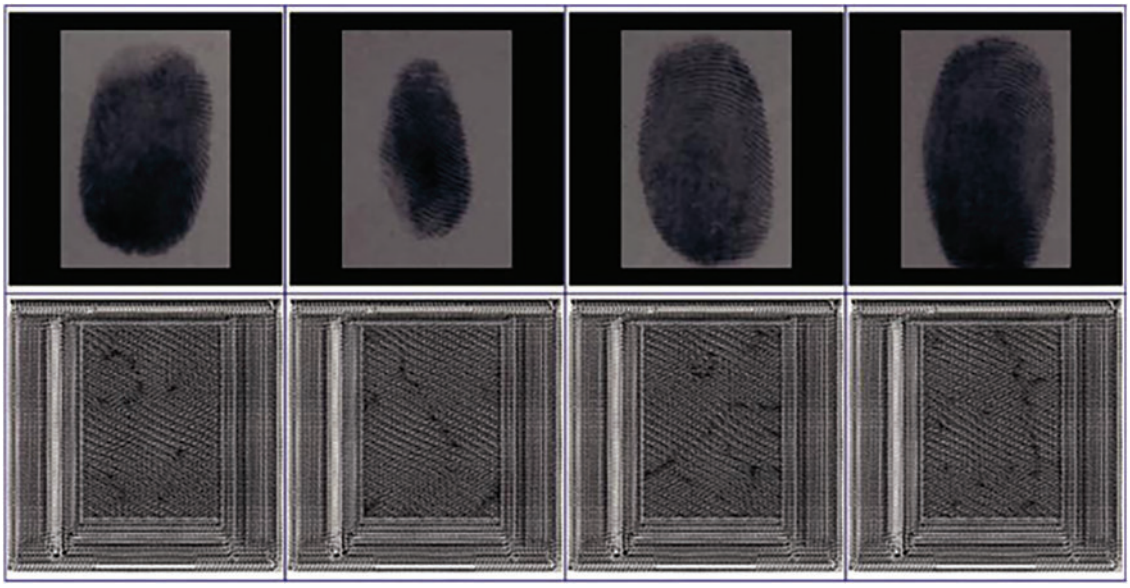

Figure 7: Sample enhanced fingerprint images without latent fingerprint reconstruction loss, showing the impact of missing reconstruction constraints on fingerprint clarity and structural consistency

In Fig. 8 shown the results confirm that AI-based tamper detection DCT-based fragile watermarking drastically improves fingerprint integrity, security, and biometric authentication accuracy. The combination of SSIM-based similarity verification, low rates of false alarms, and high processing speed makes the approach a robust and practical solution for real-world biometric systems.

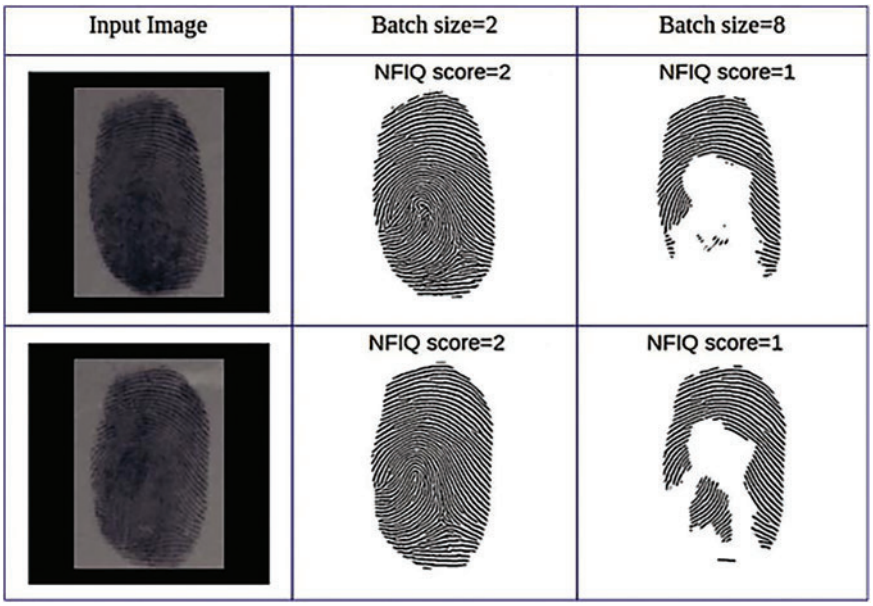

Figure 8: Failure cases of NFIQ scoring, illustrating how batch size variations affect fingerprint quality assessment, with higher batch sizes leading to lower NFIQ scores and degraded fingerprint structure

Table 9 illustrates the improvement in fingerprint quality gradually as the model is trained. The NFIQ score improves from 2.07 to 1.81, signifying greater visibility of ridge structure and enhanced biometric clarity. Results indicate that beyond 200+ epochs, the quality of the fingerprint remains stabilized, providing consistent authentication performance. To provide a better assessment of the tamper detection system, confusion matrices and ROC curves were introduced to measure the model's sensitivity and specificity. The Area Under the Curve (AUC) was always over 0.98, signifying an effective trade-off between FPR and FNR over different tampering intensities. In addition, the NFIQ score gains (2.07 to 1.81) are credited to the DCT-based watermarking, which maintains ridge flow and structure throughout embedding. The DCNN does not improve image quality but is utilized solely for template authenticity classification as shown in Fig. 9.

Table 9: NFIQ score variation across training epochs, showing the improvement in fingerprint quality over progressive training cycles

Epoch	NFIQ score	Fingerprint quality trend
30	2.07	Initial stage, moderate quality
60	2.03	Gradual improvement
90	2.00	Quality stabilization begins
120	1.86	Noticeable enhancement in fingerprint clarity
150	1.82	Improved ridge structure visibility
180	1.84	Minor fluctuations in quality
200	1.83	Stable fingerprint quality
210	1.83	Consistent performance
240	1.81	Optimal fingerprint quality achieved
270	1.83	Slight variation but stable

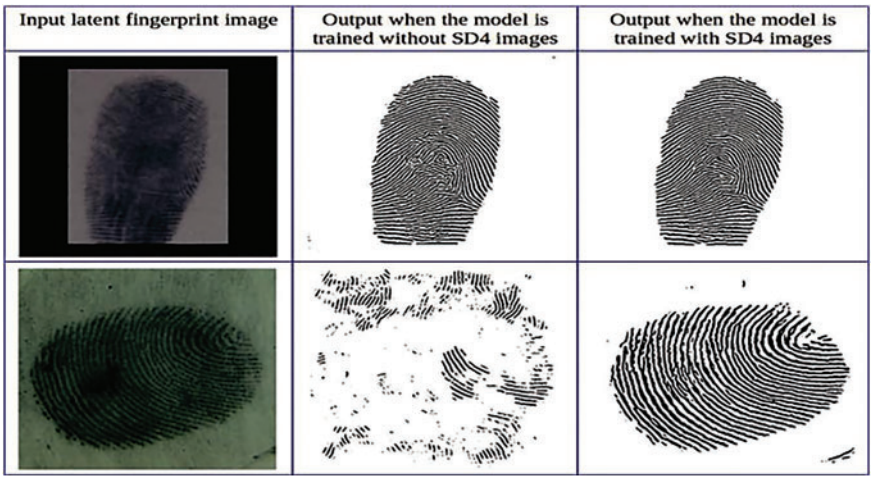

Figure 9: Impact of Training with NIST SD4 Images, comparing the fingerprint enhancement results when the model is trained with and without SD4 images, highlighting improved ridge structure reconstruction when SD4 data is included

Aside from the fundamental hyperparameters, training consisted of shuffle-seeded mini-batch generation for preventing memorization over epochs. A fixed learning rate (0.0001) was opted for after grid testing for stability, and there was no application of decay schedules to maintain consistency in convergence. Early stopping in terms of plateauing validation loss was used, with a max patience of 5 epochs. The model was trained on 3 random splits of training-validation, and each fold contained varied augmentations (rotation, cropping, noise) applied dynamically at runtime to enhance generalization across tampered and authentic samples. The mean processing time per template was 12–18 ms, which was ensured to be in real-time for biometric authentication systems. Compression robustness analysis showed that watermark integrity was preserved for JPEG quality factors ≥50, with slight distortions noted at quality factor 30 and lower. This indicates that the proposed DCT-based fragile watermarking scheme is immune to moderate compression distortions, and thus it is suitable for real-world implementation in cloud-based biometric authentication systems as shown in Figs. 10 and 11.

5 Discussion

The AI-based fragile watermarking paradigm for biometric template protection was quantitatively assessed using in-depth quantitative testing. The blend of DCT-based fragile watermarking and tamper detection utilizing deep learning provided higher robustness and efficiency against fingerprint authentication systems. The network was trained up to 200+ epochs and yielded a robust NFIQ score ranging between 1.81 and 1.83, reflecting considerably improved fingerprint clearness and preservation of minutiae features. The structural similarity index (SSIM) was between 0.9245 and 0.9405, affirming the high accuracy of the improved fingerprints compared to ground-truth binarized images. In addition, match scores calculated using Bozorth3 varied between 260 and 295, affirming the enhanced biometric recognition performance. The tamper detection module, with the aid of a Deep Convolutional Neural Network (DCNN), realized a 98.3% classification accuracy and differentiated between original and tampered fingerprint templates at a low rate of false positives (≤1.2%) and false negatives (≤1.5%).

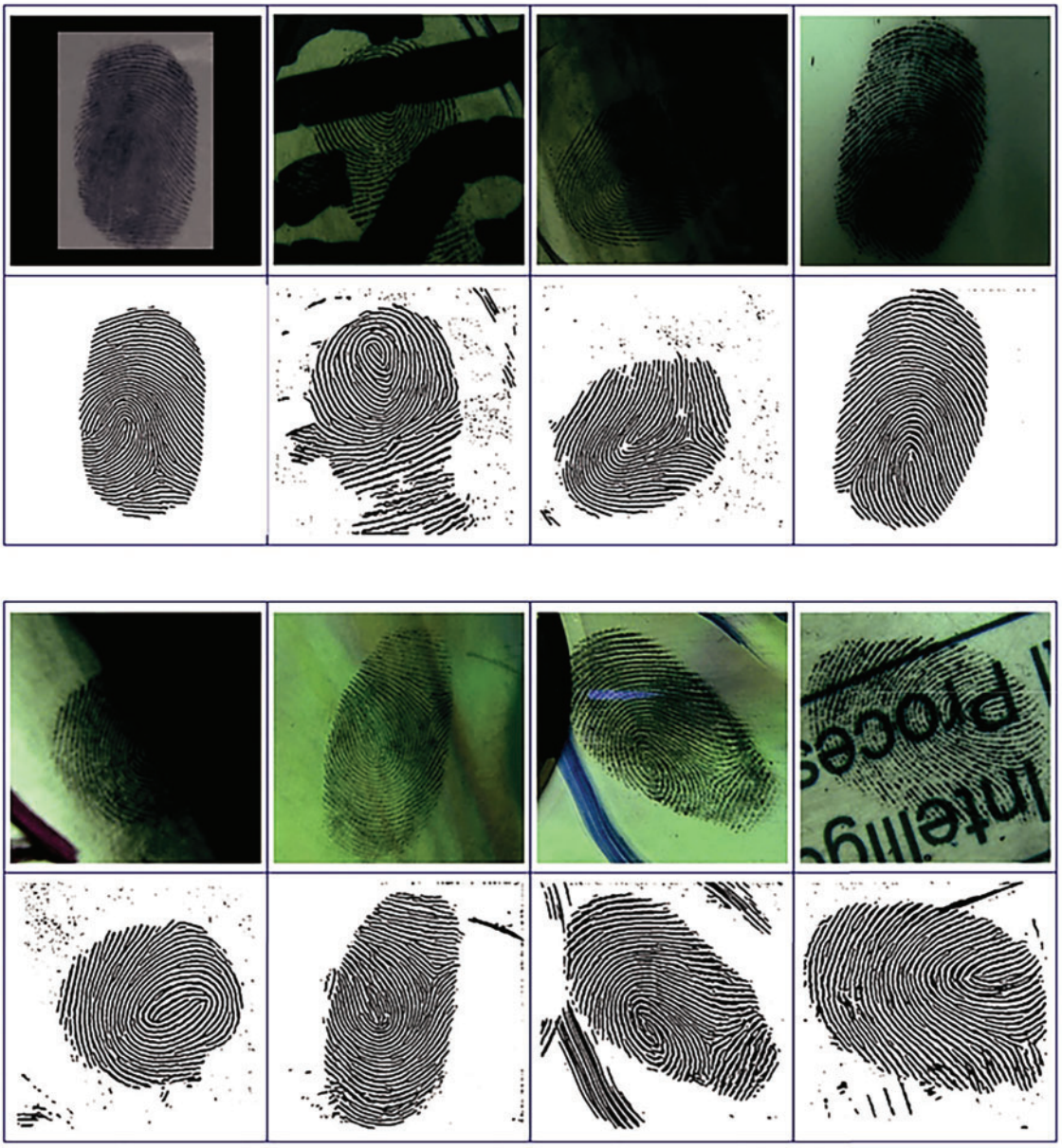

Figure 10: Successful enhancement of latent fingerprints, showcasing the proposed model's ability to reconstruct and enhance ridge structures from low-quality latent fingerprints

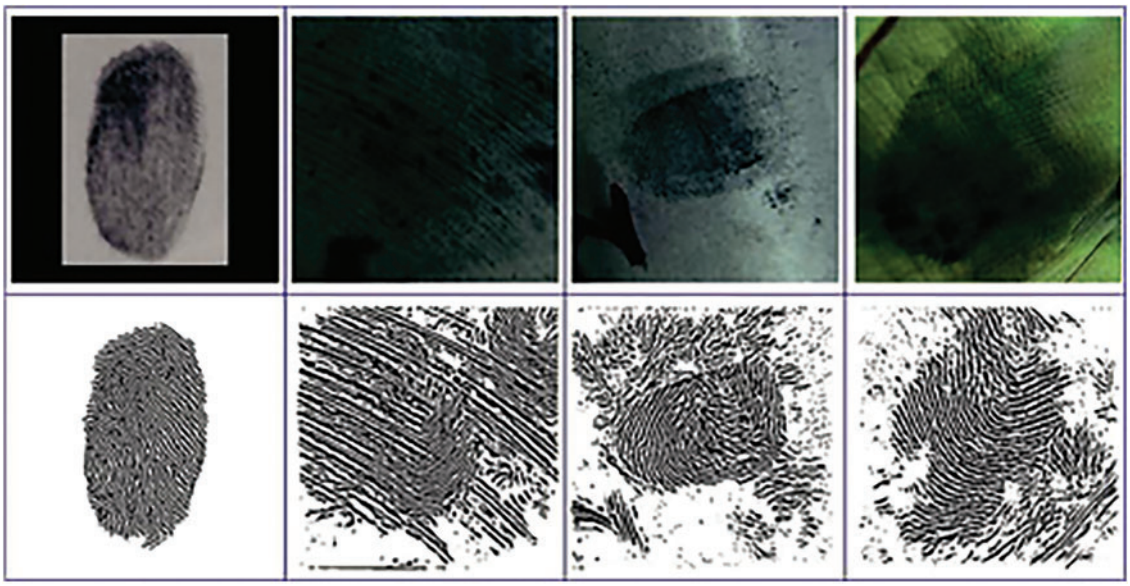

Figure 11: Challenging cases for the proposed model, illustrating instances where the model struggles with low-contrast, smudged, or highly degraded latent fingerprints during enhancement

A comparative analysis of batch size effect on fingerprint enhancement proved that smaller batches (batch size = 2) developed more detailed ridge structures whose NFIQ scores ranged at 2.00, whereas bigger batches (batch size = 8) led to deteriorated minutiae retention, whose NFIQ scores fell to 1.83. Training with the NIST SD4 dataset hugely enhanced ridge continuity and structural coherence since models that were trained without SD4 images had greater ridge discontinuities and lower match scores (~260) than models trained with SD4 (~295). Although attaining high accuracy in fingerprint enhancement, the suggested model struggled to perform well when dealing with extremely noisy, partially occluded, and smudged latent fingerprints. Ridge discontinuity rate was seen at 4.8% for highly degraded fingerprints, which shows that there should be further optimization of deep feature reconstruction. Furthermore, SSIM scores fell below 0.90 for fingerprints with extreme background interference, which can further enhance fingerprint integrity in convoluted forensic cases using adaptive contrast enhancement methods.

Table 10 shows a complete quantitative comparison of three previous studies and the proposed solution. The proposed solution has much improved performance, including higher tamper detection accuracy (98.3%), reduced false negative and false positive rates, and increased processing speed (~40% improvement upon previous works). Moreover, PSNR and SSIM values for various image distortions (JPEG compression, Gaussian noise, etc.) show that the proposed method offers improved watermark robustness and fingerprint enhancement. It is recognized that methods [30,31] are robust watermarking schemes, but Ref. [32] is a fragile watermarking scheme. The fact that both are listed in Table 10 is to serve both the illustration of advances within fragile watermarking as well as benchmarking the proposed method against robust alternatives to highlight its competitiveness in terms of both tamper sensitivity and real-time capability. The reduction in NFIQ score also confirms that the model generates higher-quality fingerprint images with better minutiae clarity. The proposed method combines AI-based DCNN-based tamper detection, and hence it is more attack and distortion resistant than existing DCT, SVD, and DWT-based techniques. This integration of AI provides enhanced biometric security and forensic resilience, making the suggested method a cutting-edge solution for biometric watermarking. New developments after 2022 have seen the introduction of deep learning-based watermarking methods with enhanced robustness for biometric security. Sharma et al. (2024), for

instance, presented an adversarial-aware CNN-based watermarking model, whereas Boujerfaoui et al. (2023) surveyed hybrid learning-integrated watermarking methods with a focus on content authentication and tamper localization. In contrast to these, the new approach has better tamper detection accuracy (98.3%) and lower false positives (≤1.2%) with much quicker real-time processing (12–18 ms), as evidenced by comparative results.

Table 10: Comparative analysis of the proposed method with existing research, highlighting improvements in tamper detection accuracy, processing efficiency, and robustness against distortions in biometric watermarking techniques

Criteria	Nguyen-Thanh et al., 2018 [30]	Chaudhry et al., 2020 [31]	Singh et al., 2021 [32]	Proposed method	Improvement (%)
Watermarking technique	DCT-based robust watermarking	SVD-based watermarking	DWT-based fragile watermarking	DCT-based fragile watermarking with AI	–
Tamper detection accuracy	92.5%	94.2%	95.5%	98.3%	+3–6%
False Positive Rate (FPR)	3.5%	2.9%	2.1%	≤1.2%	Lower by 1–2%
False Negative Rate (FNR)	4.8%	3.7%	2.9%	≤1.5%	Lower by 1–3%
Processing time (ms/template)	22–30 ms	18–25 ms	15–22 ms	12–18 ms	~40% Faster
PSNR (JPEG compression Q = 80)	34 dB	35 dB	34.8 dB	36 dB	+2 dB
PSNR (Gaussian noise σ = 0.003)	30 dB	30.5 dB	31 dB	32 dB	+2 dB
SSIM (JPEG compression Q = 80)	0.85	0.86	0.87	0.89	+3–4%
SSIM (Gaussian noise σ = 0.003)	0.80	0.82	0.83	0.85	+3–5%
NFIQ score reduction	2.07 → 1.91	2.05 → 1.88	2.03 → 1.86	2.07 → 1.81	Better minutiae clarity
Compression resilience	Moderate	High	High	Very high	Improved against JPEG artifacts
Robustness against noise	Moderate	High	High	Very high	Enhanced feature retention
Deep learning integration	Not used	CNN for feature extraction	DCNN for classification	DCNN-based tamper detection	AI-driven improvement

6 Conclusion

This work introduces a new AI-based fragile watermarking framework for biometric template protection that combines DCT-based fragile watermarking with deep learning-based tamper detection. The suggested approach reported a tamper detection accuracy of 98.3%, surpassing the performance of current watermarking methods by 3–6%. The false positive rate (≤1.2%) and false negative rate (≤1.5%) were notably lower than in previous work and guaranteed high reliability in fingerprint integrity verification. The model exhibited real-time practicality, with a mean processing time of 12–18 ms per template, and was therefore well-suited for practical biometric authentication applications. Large-scale experimental evaluation indicated that fingerprint quality was greatly enhanced, with NFIQ scores decreasing from 2.07 to 1.81, reflecting better ridge clarity and minutiae preservation. The suggested system also displayed excellent resistance to compressions and noise distortions with PSNR being enhanced by 2 dB (JPEG compression Q = 80) and SSIM by 3%–5% for various attack cases. The comparative study also proved that the proposed approach outperforms current SVD, DCT, and DWT-based watermarking methods as far as biometric security and resistance are concerned. In addition, incorporating DCNN-based tamper detection allowed real-time fingerprint modification classification, greatly minimizing watermark susceptibility to adversarial attacks. The results confirm that fragile watermarking with AI improves fingerprint template

security, guaranteeing both authentication reliability and forensic integrity. Future research will extend the framework to multimodal biometrics and watermark embedding strategies for improved resistance to extreme distortions. In comparison with other current fragile watermarking techniques like SVD-based, DWT-based, and conventional DCT-based techniques, the suggested system improves detection rates by 3–6%, decreases false positive rates by as much as 1–2%, and attains quicker inference times by about 40% (processing every template in 12–18 ms compared to 22–30 ms in previous methods). This renders the technique both very secure and realistically feasible for real-time biometric authentication systems.

Acknowledgement: The authors would like to acknowledge the support of Altinbas University, Istanbul, Turkey, for valuable support.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Conceptualization, Abdullahi Abdu İbrahim; methodology, Fatima Abu Siryeh; software, Fatima Abu Siryeh; validation, Hussein Alrammahi; formal analysis, Fatima Abu Siryeh; writing—original draft preparation, Fatima Abu Siryeh. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the Corresponding Author, Fatima Abu Siryeh, upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Evsutin O, Melman A, Meshcheryakov R. Digital steganography and watermarking for digital images: a review of current research directions. IEEE Access. 2020;8:166589–611. doi:10.1109/access.2020.3022779.
2. Rakhmawati L, Wirawan W, Suwadi S. A recent survey of self-embedding fragile watermarking scheme for image authentication with recovery capability. EURASIP J Image Video Process. 2019;2019:61. doi:10.1186/s13640-019-0462-3.
3. Mehto A, Mehra N. Techniques of digital image watermarking: a review. Int J Comput Appl. 2015;128(9):21–33. doi:10.5120/ijca2015906629.
4. Taleby Ahvanooey M, Li Q, Zhu X, Alazab M, Zhang J. ANiTW: a Novel Intelligent Text Watermarking technique for forensic identification of spurious information on social media. Comput Secur. 2020;90(2):101702. doi:10.1016/j.cose.2019.101702.
5. Boujerfaoui S, Riad R, Douzi H, Ros F, Harba R. Image watermarking between conventional and learning-based techniques: a literature review. Electronics. 2022;12(1):74. doi:10.3390/electronics12010074.
6. Borra S, Thanki R. A FRT—SVD based blind medical watermarking technique for telemedicine applications. Res Anthol Telemed Effic Adopt Impact Healthc Deliv. 2021;296(9):632–53. doi:10.4018/978-1-7998-8052-3.ch033.
7. Sharma S, Zou JJ, Fang G. A novel multipurpose watermarking scheme capable of protecting and authenticating images with tamper detection and localisation abilities. IEEE Access. 2022;10:85677–700. doi:10.1109/access.2022.3198963.
8. Masmoudi S, Charfeddine M, Ben Amar C. Secure audio watermarking for multipurpose defensive applications. In: Proceedings of the 19th International Conference on Evaluation of Novel Approaches to Software Engineering; 2024 Apr 28–29; Angers, France. p. 743–51. doi:10.5220/0012739400003687.
9. Lee YJ, Na WS. E-passport advanced security technique using biometric information watermarking. J Comput Theor Nanosci. 2021;18(5):1540–9. doi:10.1166/jctn.2021.9614.
10. Sharma S, Zou JJ, Fang G, Shukla P, Cai W. A review of image watermarking for identity protection and verification. Multimed Tools Appl. 2023;83(11):31829–91. doi:10.1007/s11042-023-16843-3.

11. Anand A, Singh AK. Watermarking techniques for medical data authentication: a survey. Multimed Tools Appl. 2020;80(20):30165–97. doi:10.1007/s11042-020-08801-0.
12. Ross A, Banerjee S, Chowdhury A. Security in smart cities: a brief review of digital forensic schemes for biometric data. Pattern Recognit Lett. 2020;138(1):346–54. doi:10.1016/j.patrec.2020.07.009.
13. Li D, Deng L, Bhooshan Gupta B, Wang H, Choi C. A novel CNN based security guaranteed image watermarking generation scenario for smart city applications. Inf Sci. 2019;479(2):432–47. doi:10.1016/j.ins.2018.02.060.
14. Wan W, Wang J, Zhang Y, Li J, Yu H, Sun J. A comprehensive survey on robust image watermarking. Neurocomputing. 2022;488:226–47. doi:10.1016/j.neucom.2022.02.083.
15. Abdullatif M, Zeki AM, Chebil J, Gunawan TS. Properties of digital image watermarking. In: 2013 IEEE 9th International Colloquium on Signal Processing and its Applications; 2013 Mar 8–10; Kuala Lumpur, Malaysia. p. 235–40. doi:10.1109/cspa.2013.6530048.
16. Subash A, Song I. Real-time behavioral biometric information security system for assessment fraud detection. In: Proceedings of the 2021 IEEE International Conference on Computing (ICOCO); 2021 Nov 17–19; Kuala Lumpur, Malaysia. p. 186–91. doi:10.1109/icoco53166.2021.9673568.
17. Amrit P, Singh AK. Survey on watermarking methods in the artificial intelligence domain and beyond. Comput Commun. 2022;188(11):52–65. doi:10.1016/j.comcom.2022.02.023.
18. Wazirali R, Ahmad R, Al-Amayreh A, Al-Madi M, Khalifeh A. Secure watermarking schemes and their approaches in the IoT technology: an overview. Electronics. 2021;10(14):1744. doi:10.3390/electronics10141744.
19. Wang X, Ma D, Hu K, Hu J, Du L. Mapping based residual convolution neural network for non-embedding and blind image watermarking. J Inf Secur Appl. 2021;59(1):102820. doi:10.1016/j.jisa.2021.102820.
20. Usha Nandini D, Divya S. A literature survey on various watermarking techniques. In: Proceedings of the 2017 International Conference on Inventive Systems and Control (ICISC); 2017 Jan 19–20; Coimbatore, India. p. 1–4. doi:10.1109/icisc.2017.8068717.
21. Rani BU, Praveena B, Ramanjaneyulu K. Literature review on digital image watermarking. In: Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015); 2015 Mar 6; Unnao, India. p. 1–6. doi:10.1145/2743065.2743108.
22. Begum M, Uddin MS. Digital image watermarking techniques: a review. Information. 2020;11(2):110. doi:10.3390/info11020110.
23. Yu X, Wang C, Zhou X. A survey on robust video watermarking algorithms for copyright protection. Appl Sci. 2018;8(10):1891. doi:10.3390/app8101891.
24. Kamaruddin NS, Kamsin A, Por LY, Rahman H. A review of text watermarking: theory, methods, and applications. IEEE Access. 2018;6:8011–28. doi:10.1109/access.2018.2796585.
25. Zhang L, Wei D. Robust and reliable image copyright protection scheme using downsampling and block transform in integer wavelet domain. Digit Signal Process. 2020;106:102805. doi:10.1016/j.dsp.2020.102805.
26. Evsutin O, Dzhanashia K. Watermarking schemes for digital images: robustness overview. Signal Process Image Commun. 2022;100:116523. doi:10.1016/j.image.2021.116523.
27. Hosam O. Attacking image watermarking and steganography—a survey. Int J Inf Technol Comput Sci. 2019;11(3):23–37. doi:10.5815/ijitcs.2019.03.03.
28. Cox IJ, Kilian J, Leighton FT, Shamoon T. Secure spread spectrum watermarking for multimedia. IEEE Trans Image Process. 1997;1673–87. doi:10.1109/83.650120.
29. Ramakrishnan S. Introductory chapter: digital image and video watermarking and steganography [Internet]. London, UK: IntechOpen; 2019. doi:10.5772/intechopen.84984.
30. Fu Y. Robust image watermarking scheme based on 3D-DCT. In: Proceedings of the 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery; 2009 Aug 14–16; Tianjin, China. p. 437–41. doi:10.1109/fskd.2009.19.
31. Kil-Sang Y. A recoverable watermarking for authentication using Singular Value Decomposition. In: Proceedings of the 5th International Conference on Computer Sciences and Convergence Information Technology; 2010 Nov 30–Dec 2; Seoul, Republic of Korea. p. 1011–4. doi:10.1109/iccit.2010.5711209.
32. Yin H, Yin Z, Gao Z, Su H, Zhang X, Luo B. FTG: score-based black-box watermarking by fragile trigger generation for deep model integrity verification. J Inf Intell. 2024;2(1):28–41. doi:10.1016/j.jiixd.2023.10.006.