People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
M'hamed Bougara University of Boumerdes



Faculty of Sciences
Department of Computer Science
Program: Computer Science.
Specialization: Computer Systems.

Final Year Dissertation for the Obtention of the Academic Bachelor's Degree

Theme

Fingerprint Presentation Attack Detection Using Vision Transformers

Presented by:
Allal Cherif Ines
Aissi Ounissa
Kezadri Chahrazed
Metiaz Elhadj Mohamed

Defended on 27/05/2026 Before the Jury Composed of:
HERBADJI Abderrahmane: Supervisor
ALOUANE Besma: Examiner

Table of Contents

List of Abbrevrations .....
General Introduction 1
Chapter 1: Biometrics & Security Threats 2
Introduction 2
1.1 Biometrics Fundamentals: 2
1.2 Fingerprint Recognition systems: 3
1.3 Biometric System Pipeline: 3
1.3.1 Acquisition: 3
1.3.2 Feature Extraction: 3
1.3.3 Matching: 3
1.4 Security vulnerabilities: 3
Conclusion 4
Chapter 2 :Fingerprint Presentation Attacks 6
Introduction 6
2.1 Presentation Attack Overview 6
2.1.1 Spoof vs. Genuine Fingerprints 6
2.1.2 Attack Goals 7
2.2 Attack Materials and Methods 7
2.2.1 Silicone 7
2.2.2 Gelatin 7
2.2.3 Latex 7
2.2.4 Printed Fingerprints 8
2.2.5 Lifted Latent Prints 8
2.2.6 Emerging Materials 8
2.3 Attack Generation Pipeline 8
2.4 Sensors and Hardware Defenses 9
2.4.1 Multi-spectral Sensors 9
2.4.2. Thermal Fingerprint Sensing: 10
2.4.3 Pressure and Capacitive Sensors 10
2.4.4 Optical Distortion Detection 11
2.5 Software-based PAD 12
2.5.1 Texture Analysis 12
2.5.2 LBP and BSIF 12
2.5.3 Deep Learning (CNN) 14
Conclusion 15
Chapter 3: Proposed PAD Detection 16
3.1 System Overview 16
Introduction 16
3.1.2 General Architecture 16
3.1.3 Input Fingerprint Image 16
3.1.4 Pre-processing Stage 16
3.1.5 ViT-Based Feature Extraction 16
3.1.6 Feature Vector Generation 17

3.1.7 ANN Classification 17
3.1.8 Decision Stage (LIVE / FAKE) 17
3.2 Proposed Method 17
3.2.1 Feature Extractor: Pretrained ViT (Frozen) 18
3.2.2 Classifier: ANN for Binary Classification (LIVE vs. FAKE) 18
3.2.3 Decision Logic: SoftMax Classification 18
3.2.4 System Architecture Summary 19
3.2.5 Evaluation Metrics 19
3.3. Dataset (LivDet 2015 and LivDet 2023) 20
3.3.1 Dataset Presentation 20
3.3.2 Sensors 21
3.3.3 Dataset Structure (LIVE / FAKE) 21
3.3.4 Split Strategy 21
3.4. Experimental Setup 22
3.4.1 Hardware Equipment 22
3.4.2 Framework 22
3.4.3 Evaluation Protocol 22
3.4.4 Metrics 22
3.5 Results and Discussion 23
3.5.1-Result Tables and Plots 23
3.5.2 Comparison with Baseline 26
Discussion 26
Conclusion 27
General Conclusion **28**
Bibliography **29**

List of Abbrevrations

CNN	Average Classification Error
ACER	Average Classification Error Rate
ANN	Artificial Neural Network
APCER	Attack Presentation Classification Error Rate
BPCER	Bona Fide Presentation Classification
BSIF	Binarized Statistical Image Features
CLS	Classification Token
CNN	Convolutional Neural Network
DNA	Deoxyribonucleic Acid
DoS	Denial of Service
EER	Equal Error Rate
FAR	False Acceptance Rate
FFT	Fast Fourier Transform
FRR	False Rejection Rate
ID	Identity Document
LBP	Local Binary Patterns
LivDet	Livness Detection Competition
LPQ	Local Phase Quantization
MHSA	Multi-Head Self-Attention
MLP	Multi-Layer Perceptron
NIST	National Institute
PAD	Presentation Attack Detection
PAI	Presentation Attack Instrument
SIFT	Scale-Invariant Feature Transform
SVM	Support Vector Machine
ViT	Vision Transformer

General Introduction

Biometric authentication systems have become essential components of modern security infrastructures due to their ability to identify individuals based on unique physiological or behavioral characteristics. Among the various biometric modalities, fingerprint recognition remains one of the most widely used because of its reliability, distinctiveness, and ease of integration into applications such as smartphones, banking systems, border control, and access management. However, despite their effectiveness, fingerprint recognition systems remain vulnerable to presentation attacks, where artificial fingerprints fabricated from materials such as silicone, gelatin, or latex deceive biometric sensors and gain unauthorized access.

To address these threats, Fingerprint Presentation Attack Detection (PAD), also known as fingerprint liveness detection, has emerged as a critical security mechanism in biometric systems. PAD techniques distinguish genuine fingerprints from spoofed ones by identifying physical, physiological, or textural characteristics of fake samples. Existing approaches fall into two categories: hardware-based methods, which use additional sensors to capture biological properties such as temperature or perspiration, and software-based methods, which analyze fingerprint images to detect anomalies related to spoof materials.

The main objective of this work is to develop and evaluate a robust fingerprint PAD system using deep learning to accurately distinguish LIVE fingerprints from FAKE ones. The proposed approach combines a pretrained Vision Transformer (ViT) for feature extraction with an Artificial Neural Network (ANN) classifier for binary classification. First, fingerprint images are preprocessed to improve quality and ensure consistency. The images are then processed by the Vision Transformer, which extracts discriminative features by capturing both local texture information and global structural relationships through self-attention mechanisms. Finally, the extracted features are fed into the ANN classifier to determine the nature of the samples.

The proposed approach was evaluated using benchmark datasets from the International Fingerprint Liveness Detection Competition, namely LivDet2015 and LivDet2023. These datasets contain fingerprint samples acquired from multiple sensors and generated using various spoofing materials, enabling realistic performance evaluation under diverse conditions. The implementation and experiments were carried out using PyTorch along with supporting libraries such as NumPy, Pandas, Scikit-learn, and Matplotlib.

This manuscript is organized as follows. Chapter 1 presents the theoretical background on biometrics and fingerprint recognition systems, as well as their main security vulnerabilities. Chapter 2 discusses fingerprint presentation attacks, spoof fabrication materials, and existing PAD countermeasures. Chapter 3 describes the proposed PAD system, including the methodology, datasets, experimental setup, evaluation metrics, and results. Finally, the general conclusion summarizes the main findings and suggests possible directions for future research.

Chapter 1: Biometrics & Security Threats

Introduction

"Fingerprints cannot lie, but liars can make fingerprints." Mark Twain's statement highlights the growing challenge of spoofing attacks in modern biometric systems. As fingerprint recognition is increasingly used in airports, smartphones, and secure environments, protecting these systems against fake fingerprints has become essential [3].

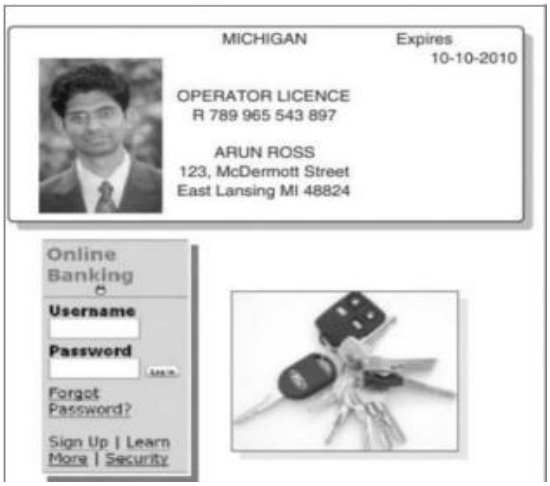
To address this issue, researchers have developed Fingerprint Presentation Attack Detection (PAD) methods capable of distinguishing genuine fingerprints from spoofed ones [4]. PAD techniques are generally divided into two categories: hardware-based approaches, which rely on additional sensors to detect signs of life such as temperature or perspiration, and software-based approaches, which analyze fingerprint images using extracted features.

Advances in artificial intelligence and deep learning have significantly improved software-based PAD systems, making fingerprint spoof detection more accurate, reliable, and efficient.

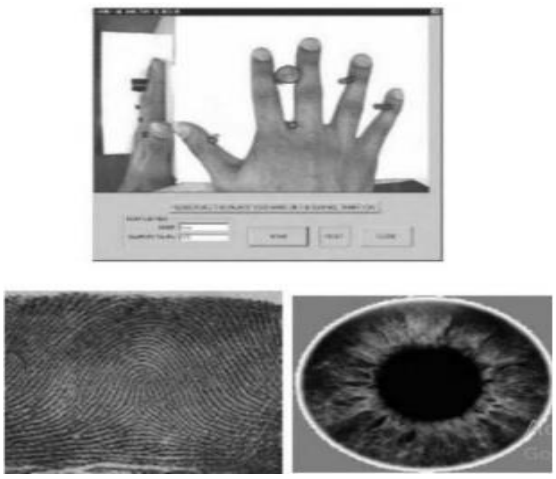
1.1 Biometrics Fundamentals:

Biometrics is the science of establishing an individual's identity based on the physical, chemical, or behavioral attributes. Its relevance in modern society has been reinforced by the need for large-scale identity management systems whose functionality relies on accurate identity determination in applications such as networked computer resources, nuclear facilities, remote financial transactions, and commercial aviation. The proliferation of web-based services and decentralized customer service centers has further underscored the need for reliable identity management systems [6].

Traditional methods of establishing a person's identity include knowledge-based (e.g., passwords) and token-based (e.g., ID cards) mechanisms, but these surrogate representations of identity can easily be lost, shared, manipulated, or stolen. Biometrics offers a natural and reliable solution by utilizing fully automated or semi-automated schemes to recognize individuals based on their biological characteristics. With biometrics, it is possible to establish an identity based on who you are, rather than what you possess or what you remember (Figure 1.1) [3][4].



(a) Traditional authentication (passwords, tokens, ID cards).



(b) Biometric authentication based on "who you are"

Figure 1.1 – Comparison of authentication schemes.

Biometrics offers key advantages such as negative recognition and non-repudiation that cannot be provided by tokens or passwords. Negative recognition determines that an individual is enrolled in the system even if they deny it critical in welfare disbursement to prevent double-dipping. Non-

repudiation guarantees that a person who accessed a facility cannot later deny doing so. Biometric systems use a variety of physical or behavioral characteristics including fingerprint, face, hand/finger geometry, iris, retina, signature, gait, palmprint, voice pattern, and DNA, to establish identity [6].

1.2 Fingerprint Recognition systems:

Fingerprint recognition is the process of comparing known and unknown fingerprints to determine whether they originate from the same person [6][3]. The literature demonstrates that fingerprint recognition algorithms generally rely on four distinct methodological families:

- **Biometric Frameworks and Multimodal Systems:** Distinguish between unimodal systems (single trait) and multimodal systems (fingerprint + face, iris, voice, or EEG). Multimodal systems significantly improve reliability but introduce high computational complexity [3][4].
- **Minutiae-Based Extraction and Local Matching:** The most prevalent family, relying on sequential pipelines involving histogram equalization, binarization, thinning, minutiae extraction, and false minutiae removal. These systems remain sensitive to dry or wet skin conditions and nonlinear deformations [4].
- **Hybrid, Correlation, and Ridge-Feature Methods:** Combine minutiae extraction with correlation-based gray-level profiles or ridge-density maps. Powerful for low-quality images, requiring less pre-processing while keeping storage costs low.
- **Transform and Core-Based Optimizations**: Enable high-speed operations via 2D Fast Fourier Transforms (FFT), Gabor Filtering, Wavelet transforms, and Spectral Minutiae representations. Enable large-scale database matching but depend critically on core point detection accuracy.

1.3 Biometric System Pipeline:

A biometric system is essentially a pattern recognition system that operates by acquiring biometric data from an individual, extracting a feature set from the acquired data, and comparing this feature set against the template set in the database [3][6]. The three core modules are:

1.3.1 Acquisition:

The sensor module captures the biometric data (e.g., a fingerprint sensor images the ridge and valley structure of a user's finger) [4][6].

1.3.2 Feature Extraction:

The acquired data is processed to extract salient or discriminatory features (e.g., the position and orientation of minutiae points in a fingerprint image) [3].

1.3.3 Matching:

Extracted features are compared against stored templates to generate matching scores. The matcher module also encapsulates a decision-making component that confirms a user's claimed identity (verification) or establishes a user's identity (identification).

1.4 Security vulnerabilities:

To systematically classify the threats facing fingerprint recognition systems, a generalized taxonomy of vulnerabilities adapted from Mirajkar's security assessment criteria is detailed below [4][6][3]:

- **Administration:** Intentional or unintentional procedural errors executed by system administrators during configuration.
- **User Privilege Escalation:** Actions taken by a legitimate user attempting to illicitly upgrade their operational privileges to an administrative level.
- **Enrollment Exploits:** Subversion or bypassing of initial biometric registration and verification procedures.

- **Spoofing:** The presentation of a synthesized, fake biometric trait (e.g., molded silicone or gelatin replicas) to impersonate a legitimate user.
- **Mimicry:** The tactical replication of a legitimate user's physical biometric features by an adversary.
- **Undetected Vulnerabilities:** Latent system flaws that fail to trigger alarms, thereby facilitating subsequent iterative attacks.
- **Fail Secure Anomalies:** System failures or vulnerabilities arising from abnormal operational conditions within the host IT environment.
- **Power Exploitation:** Intentional disruption or cutting of electrical power lines to force a system reset or failure mode.
- **Bypass Attacks:** Completely avoiding the biometric subsystem via physical barrier destruction, coercing an authorized individual, or insider cooperation.
- **Corrupt Attack:** The deliberate weakening of system integrity through the modification or substitution of core configuration parameters.
- **Degrade Vectors:** Secondary software applications running within the shared IT infrastructure that inadvertently reduce the biometric system's security boundaries.
- **Hardware Tampering:** Physical counterfeiting, modification, or probing of the biometric reader's hardware components.
- **Residual Trait Exploitation:** Capture and replication of latent fingerprints left behind on the physical sensor to manufacture artificial templates.
- **Cryptological Attacks:** Breaking cryptographic primitives or encryption keys during data transmission, leading to advanced interception vectors such as Replay Attacks.
- **Brute Force Attacks:** The iterative, high-frequency presentation of synthetic biometric variants to the sensor, a vector highly dependent on a high False Acceptance Rate (FAR) setting.
- **Evil Twin Vulnerabilities:** Exploiting high cross-class similarities where an imposter's natural biometric features closely mirror those of an enrolled user.
- **Signal Replay:** Bypassing the physical sensor entirely by directly injecting a previously intercepted and recorded legitimate biometric signal into the communication channel.
- **Fake Template Insertion:** Direct injection of fraudulent biometric templates into the centralized database storage or onto decentralized smart cards (tokens).
- **Noise Injection:** The deliberate application of electronic or signal noise to force the system into an unstable processing state conducive to unauthorized access.
- **Poor Image Enrollment:** Accepting sub-optimal, low-quality, or noisy images during registration, which broadens the matching tolerance and allows attackers to easily deceive the system.
- **Weak ID Templates**: Faking identification metrics through the utilization of highly generalized or loosely defined mathematical templates.
- **FAR/FRR Manipulation:** Adversarial targeting and optimization of attacks based on the specific operational boundaries of the system's matching thresholds.
- **Denial-of-Service (DoS):** Flood-based or algorithmic attacks aimed at completely disrupting system availability for legitimate users.

Conclusion

Biometric systems offer a powerful and convenient means of identity verification and identification, leveraging the unique biological and behavioral characteristics of individuals. From the fundamental principles of universality, uniqueness, permanence, and measurability to the intricate workings of fingerprint recognition systems and the multi-stage biometric pipeline, technology has evolved

significantly. However, the increasing reliance on biometrics necessitates a thorough understanding and robust mitigation of inherent security threats. Spoofing attacks, replay attacks, sensor attacks, and template attacks represent significant challenges that can undermine the integrity and trustworthiness of biometric systems. Continuous research and development in liveness detection, secure template protection schemes, and advanced cryptographic protocols are essential to ensure the continued efficacy and security of biometric technologies in an ever-evolving threat landscape. As biometrics become more pervasive in daily life, a balanced approach that prioritizes both convenience and robust security measures will be crucial for their successful and ethical deployment [3][4].

Chapter 2 :Fingerprint Presentation Attacks

Introduction

In this chapter, we present fingerprint presentation attacks and their impact on biometric systems. We introduce the main concepts related to these attacks, explain how they are carried out, and highlight the techniques used to detect them. Fingerprint recognition is currently one of the most widely deployed biometric modalities in real-world applications, ranging from border control to mobile authentication. This widespread use has made it a prime target for adversarial attempts, and the threat of presentation attacks has grown accordingly [2]. The formal study of these attacks was significantly advanced by the LivDet (Liveness Detection) competition series, which provided standardized benchmarks and datasets that the research community has used to evaluate and compare liveness detection algorithms. We first define the distinction between genuine and spoofed fingerprints and the goals of presentation attacks, then describe the main materials and fabrication pipelines used to create fake fingers. We then examine hardware-level countermeasures deployed at the sensor level, including multi-spectral, thermal, and capacitive sensing technologies, before turning to software-based detection approaches such as texture analysis methods and deep learning architectures [1][3].

2.1 Presentation Attack Overview

Presentation attacks consist of presenting fake fingerprints to a sensor to bypass authentication systems. These attacks aim to imitate genuine fingerprints using different techniques and materials. Understanding their principles is essential before studying how they are created and detected [1].

2.1.1 Spoof vs. Genuine Fingerprints

A genuine fingerprint presents natural ridge patterns, skin texture variations, and micro-details that are inherent to living tissue. A spoofed fingerprint, on the other hand, is an artificial replica that attempts to mimic these characteristics using synthetic materials. A Presentation Attack Instrument (PAI) is defined as any object specifically crafted to interfere with the normal operation of a biometric system at the sensor level [2].

The key differences between genuine and fake fingerprints are:

- **Genuine fingerprints**: show natural sweat pore patterns and ridge continuity.
- **Spoofed fingerprints**: often lack fine details, reflecting the properties of the fabrication material.

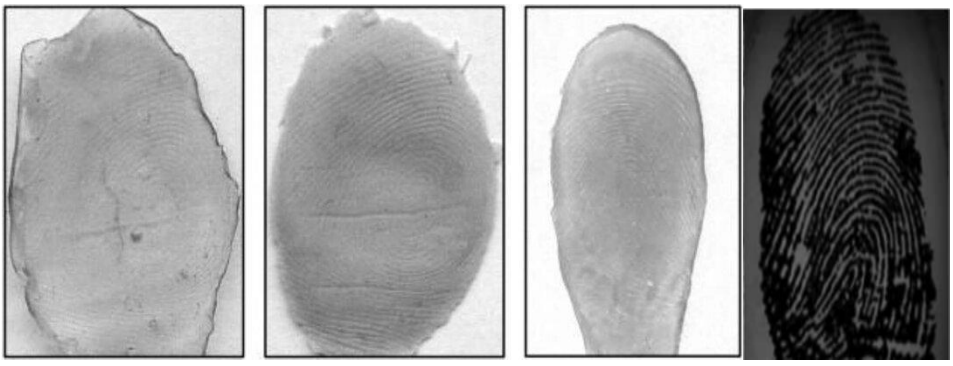


Figure 1: Comparison of 3 genuine (left) and spoofed (right) fingerprint samples.

Fake samples may exhibit surface artifacts such as air bubbles or material granularity.

Beyond visible surface differences, genuine fingers also exhibit physiological liveness indicators that artificial materials inherently lack [2]. These include perspiration patterns, elastic skin deformation under pressure, and micro-texture variations caused by the structure of living tissue. Skin distortion analysis under varying pressure levels has been shown to effectively distinguish genuine from fake fingers, since synthetic materials deform in fundamentally different ways than

human skin. Attack instruments are generally grouped by their material composition and fabrication method into what the literature refers to as PAI species, a classification that structures the materials described in the following section [6].

2.1.2 Attack Goals

The primary objective of a presentation attack is to fool the sensor-level acquisition module of the fingerprint recognition system [1]. The sensor module is considered one of the most critical vulnerability points in a biometric pipeline, as it is precisely at this stage that genuine biometric traits can be substituted with fabricated ones [1][2].

Attack goals can be summarized as follows:

- **Impersonation:** presenting a fake finger that matches a registered identity.
- **Obfuscation:** hiding one's identity from the system.
- **Unauthorized access:** bypassing security in access control applications.

It is also useful to distinguish between different attacker models based on the level of knowledge and effort involved. A zero-effort attack refers to presenting any arbitrary finger without any targeted fabrication, which rarely succeeds against a well-configured system. A targeted attack, by contrast, involves crafting a fake specifically designed to match a known enrolled identity and represents the most realistic threat in high-security scenarios [1]. A more sophisticated variant is the wolf attack, where a single fabricated artifact is designed to match multiple enrolled users simultaneously, exploiting the tolerance margins of the matcher. This type of attack is particularly relevant in large-scale deployments where matching thresholds may be set loosely to accommodate a high volume of users [6].

2.2 Attack Materials and Methods

To perform presentation attacks, different materials and techniques are used to create fake fingerprints. These methods vary in complexity, cost, and effectiveness. The most commonly encountered materials in the research literature and benchmark datasets are described below [2][6].

2.2.1 Silicone

Silicone is one of the most realistic materials for fake finger fabrication due to its flexibility and skin-like texture. Silicone-based artificial fingers are among the most effective attack materials because of their ability to replicate both the mechanical and optical properties of human skin. In practice, two-component platinum-cure silicones such as EcoFlex or Dragon Skin are commonly used, as their shore hardness can be tuned to closely approximate the compliance of real fingertip skin. This property makes silicone-based PAIs particularly dangerous against optical sensors, which rely on light reflectance from the finger surface, and capacitive sensors, which measure the dielectric properties of the material placed on the platen.

2.2.2 Gelatin

Gelatin is a widely used material that is easy to prepare and produces clear ridge patterns. It is widely adopted in liveness detection benchmarks precisely because of its accessibility and the quality of ridge detail it reproduces. Its transparency allows the resulting fake finger to capture fine level-2 and level-3 features with high fidelity. It has been shown to be one of the most challenging attack materials in benchmark evaluations, consistently achieving high spoof acceptance rates against a wide range of detection systems [1].

2.2.3 Latex

Latex offers good elasticity and is commonly used in cross-material evaluation scenarios. Its presence in standard benchmark datasets such as LivDet confirms its relevance as a reference attack material in the research community. However, latex has notable material properties that can make it easier to detect than silicone: it tends to degrade over time, its electrical properties differ significantly from biological skin, and it can leave a characteristic surface sheen that some sensors are sensitive

to. Despite these limitations, latex remains a relevant threat in practice due to its low cost and ease of fabrication, and it is regularly included as a PAI species in benchmark evaluations [1][2].

2.2.4 Printed Fingerprints

Printed fingerprints involve printing a fingerprint image on paper or transparent film, sometimes coated with conductive ink to fool capacitive sensors. Studies have shown that fingerprints reconstructed from ISO templates and printed on special media can successfully deceive optical sensors in a significant proportion of attempts. This attack method is particularly relevant in the context of digital template theft, where an attacker who has obtained a stored biometric template can reconstruct and print a usable spoof without ever physically accessing the victim. The conductive ink variant is designed specifically to circumvent capacitive sensing, which normally relies on the electrical conductivity of living skin to register a fingerprint image [1][2].

2.2.5 Lifted Latent Prints

A particularly stealthy attack method involves collecting latent fingerprints left on surfaces such as glasses or registry papers. This method is especially dangerous as it requires no direct physical contact with the victim, making it difficult to detect or prevent. The attacker can lift the latent print using adhesive film or dusting powder, scan it at high resolution, invert it digitally, and then use it to cast a mold. This attack pathway is notable because it is covert end-to-end: the victim may never be aware that their fingerprint has been captured and weaponized [1][4].

2.2.6 Emerging Materials

Beyond the classical materials described above, more recent LivDet competitions have introduced additional PAI species that reflect an evolving threat landscape. Wood glue (also referred to as body double or Elmer's glue) has appeared in several editions as a low-cost, easily accessible material that nonetheless produces surprisingly clear ridge impressions. More recently, 3D-printed molds produced from digital fingerprint models have been tested, eliminating the need for a physical mold casting step and enabling mass fabrication of attack instruments. These newer materials have been shown to pose significant challenges to detection algorithms trained exclusively on the classical set of attack materials, underscoring the importance of building liveness detection systems that generalize across unseen PAI species [1][2].

Material	Cost	Ridge Fidelity	Optical Risk	Capacitive Risk
Silicone	Medium	Very High	High	High
Gelatin	Low	High	High	Medium
Latex	Low	Medium	Medium	Low
Printed	Very Low	Medium	High	Low
Latent lift	Very Low	High	High	Medium

2.3 Attack Generation Pipeline

The fabrication of a presentation attack instrument typically follows a four-stage pipeline, consistent with the fabrication procedures documented in research literature and LivDet competition reports [1] [2][4].

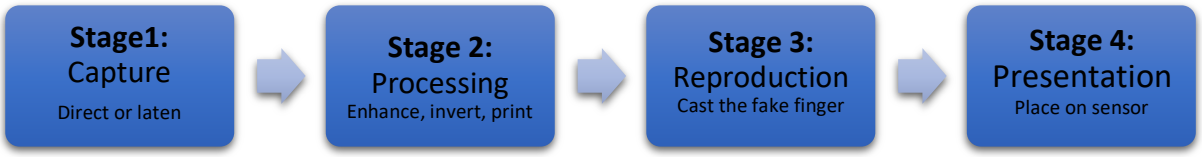


Figure 2: Attack generation pipeline.

- **Stage 1 -Capture:**

The target fingerprint is obtained either directly (with the victim's cooperation or under coercion) or indirectly through latent prints left on surfaces.

- **Stage 2 -Processing:**

The captured fingerprint is digitally enhanced and cleaned to improve ridge clarity before being used to create a mold. This typically involves threshold-based binarization to isolate ridge structures, followed by image inversion since the mold is the photographic negative of the original fingerprint and printing onto a transparency film or laser-etching the pattern onto a PCB substrate to serve as the mold master. The quality of this digital processing step directly determines the ridge fidelity of the final artifact [4].

- **Stage 3 -Reproduction:**

A physical mold is created, and the chosen material (gelatin, silicone, latex, etc.) is poured into it and allowed to set, producing the fake finger artifact. The quality of the mold at this stage directly determines the ridge fidelity of the resulting fake finger. It is important to distinguish between the negative mold, which captures the inverse of the fingerprint relief, and the positive cast, which is the final PAI that will be presented to the sensor. In some attack scenarios, a dental impression kit applied directly to the target's finger is used to obtain the mold without requiring a digital capture step, which makes the attack faster and entirely analog [1][4].

- **Stage 4 - Presentation:**

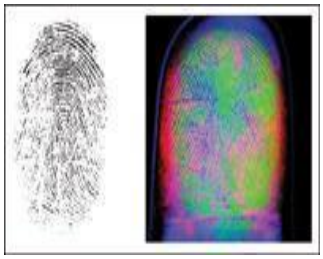
The finished artifact is presented to the fingerprint sensor. The PAI is placed on the sensor in the same way a genuine finger would be, with the goal of generating a matching fingerprint image. The success of this final stage is heavily dependent on the sensor technology being targeted. Optical sensors are generally more susceptible to presentation attacks than capacitive sensors, as they only capture a surface image without any electrical measurement. Multispectral sensors, which capture subsurface tissue information, are considered significantly more robust against PAIs precisely because synthetic materials do not have the same subsurface structure as living skin. A comprehensive analysis of how sensor type affects attack vulnerability across the different PAI species has been documented in the literature [2][4].

2.4 Sensors and Hardware Defenses

Fingerprint sensing is a widely adopted biometric modality used for identity verification and authentication systems. However, these systems are vulnerable to presentation attacks, where artificial fingerprints are used to deceive the sensor. To address this issue, Presentation Attack Detection (PAD) techniques are integrated at both hardware and software levels to improve system robustness [1].

Fingerprint acquisition systems can be differentiated according to user interaction paradigms: namely, touch-based, contactless, partial, and slap acquisition methods. These sensing systems operate by detecting the distinctive ridge valley patterns present on the fingertip surface, while also capturing additional physical properties that can assist in liveness detection [8].

2.4.1 Multi-spectral Sensors



Multi-spectral fingerprint sensors capture biometric information using multiple wavelengths of light, typically including visible and infrared spectra. Unlike conventional optical sensors that only capture surface ridge patterns, multi-spectral imaging provides both surface and subsurface skin information.

Figure 2.4.1: The print on the left of a dry fingertip. On the right is the same finger but imaged with a multispectral sensor.

The finger is illuminated using different wavelengths of light at various angles. Each wavelength penetrates the skin to different depths, producing complementary reflections. These signals are combined to generate a robust fingerprint representation that includes both external ridge structure and internal skin characteristics [7][8].

Multi-spectral sensing improves liveness detection by analyzing subsurface biological properties that are difficult to reproduce in artificial fingerprints, such as silicone, gelatin, or printed replicas [7].

Advantages

- Captures surface and subsurface fingerprint information
- Strong resistance to spoofing attacks
- High performance in challenging conditions (dry, wet, or dirty fingers)

Limitations

- High cost due to advanced optical components
- Complex processing and system design

2.4.2. Thermal Fingerprint Sensing:



Thermal fingerprint sensors operate by detecting heat patterns emitted by a finger. Human skin naturally maintains a stable temperature distribution, which allows the sensor to distinguish between live and fake fingerprints. When a finger is placed on the sensor surface [8], temperature variations are captured based on heat transfer between the skin and the sensing element, the contact interface, and the surrounding environment. These variations are captured using thermal or pyroelectric sensing elements to reconstruct the fingerprint image [7].

Figure2.4.2: Thermal sensor.

Unlike purely optical or capacitive methods, thermal sensing relies on physiological properties of living tissue rather than only surface texture, making it useful for liveness detection and presentation attack prevention [7].

However, thermal sensing performance can be affected by environmental conditions such as ambient temperature and airflow, which may affect heat distribution and measurement stability. In addition, capture speed can be slightly slower compared to conventional fingerprint technologies [7].

Advantages

- Effective in detecting artificial fingerprints that do not emit natural body heat
- Helps prevent spoofing attacks using printed or molded fingerprints
- Can capture latent fingerprints and support feature-based biometric matching

Limitation

- Sensitive to ambient temperature and airflow
- Slower capture speed
- Higher cost due to specialized materials

2.4.3 Pressure and Capacitive Sensors

Pressure and capacitive fingerprint sensing are contact-based technologies used in biometric systems to capture fingerprint patterns through physical interaction with the sensor surface. Although they are often grouped together in practice, they rely on different physical principles.

Capacitive sensing operates by measuring variations in electrical capacitance between microscopic electrodes embedded under the sensor surface. When a finger touches the sensor, the ridges come into proximity with the electrodes, while valleys remain farther away, creating measurable differences in capacitance [8]. These variations are converted into electrical signals and processed to reconstruct a fingerprint image [7].

Pressure sensing, on the other hand, measures the mechanical force distribution exerted by the finger on the sensor surface. The deformation pattern produced by the contact is captured and translated into a spatial representation of the fingerprint structure [8].

Both sensing approaches contribute to hardware-based PAD by analyzing physical interaction properties of the finger. Real human skin exhibits consistent electrical and mechanical behavior, whereas spoof materials often fail to accurately replicate both capacitance and pressure distribution patterns.

Advantages

- High accuracy in ridge–valley reconstruction (capacitive)
- Additional mechanical verification through force distribution (pressure)
- Fast response and low power consumption
- Suitable for compact and embedded devices

Limitations

- Performance affected by very dry or wet fingers
- Small sensing area may limit full fingerprint capture
- Pressure sensing alone is less commonly used and often supplementary
- Still vulnerable to advanced spoof attacks without additional PAD methods

2.4.4 Optical Distortion Detection

Optical distortion detection is a hardware-based PAD technique used in optical fingerprint systems to analyze image-level anomalies during acquisition [7].

The system examines the captured fingerprint image for irregular optical artifacts such as:

- unnatural reflections
- blurred ridge structures
- inconsistent light scattering
- printing or molding artifacts
- air gaps between fake material and sensor surface

This technique enhances spoof detection by identifying inconsistencies in how light interacts with real skin compared to artificial materials [7].

Advantages

- Easy to integrate into optical sensor pipelines and does not require additional hardware
- Effective against simple printed or molded attacks

Limitations

- Less effective against high-quality spoof materials
- Depends heavily on image quality and lighting conditions

2.5 Software-based PAD

Software-based Presentation Attack Detection (PAD) refers to algorithmic techniques designed to determine whether a presented fingerprint is bona fide or a spoofing attempt, regardless of the fingerprint capture device. Unlike hardware-based methods, these approaches rely entirely on image analysis and pattern recognition techniques. Software-based fingerprint PAD can generally be divided into two main categories: handcrafted feature-based methods and deep learning-based approaches [8][9].

2.5.1 Texture Analysis

Texture analysis examines the visual and micro-textural properties of fingerprint images to distinguish genuine fingerprints from spoofed samples by analyzing:

- Ridge smoothness
- Pore visibility
- Noise patterns
- Texture consistency
- Irregular ridge structures in real skin vs. uniform surfaces in spoof materials

Handcrafted feature-based methods are commonly employed in this context. These methods rely on manually designed descriptors to extract discriminative information from fingerprint images, including gradients, texture distributions, and fine-grained micro-textural patterns. They are primarily applied to contact-based fingerprint images and have demonstrated effective performance in spoof detection by capturing subtle differences between genuine skin and artificial materials [8][9].

The resulting texture features are subsequently utilized as input for classification algorithms to differentiate between authentic and fraudulent fingerprints.

Advantages

- Low cost
- No additional hardware required
- Easy to integrate into existing fingerprint systems
- Works with existing contact-based fingerprint images
- Effective at capturing subtle visual differences between real and fake skin

Limitations

- Sensitive to image quality and noise
- Performance can drop under poor lighting or low-resolution sensors
- May struggle with high-quality spoof materials that closely mimic real skin texture
- Less effective when texture information is limited or distorted
- Can be affected by sensor-specific variations and preprocessing methods

2.5.2 LBP and BSIF

Among the most widely used handcrafted descriptors in fingerprint PAD systems are Local Binary Patterns (LBP) and Binarized Statistical Image Features (BSIF). Both excel at capturing fine micro-textural characteristics and detecting inconsistencies in artificial fingerprints. Other common descriptors include Scale-Invariant Feature Transform (SIFT) and Local Phase Quantization (LPQ) [1].

2.5.2.1 Local Binary Patterns (LBP)

Local Binary Patterns (LBP) is a texture descriptor that encodes local image structure by comparing each pixel with its surrounding neighboring pixels. For each pixel, the intensities are thresholded against the center pixel value: values greater than or equal to the center are assigned 1, while others

are assigned 0. This process generates a binary number that represents the local texture pattern. The final representation of the image is typically obtained by building a histogram of these binary patterns [8][9].

LBP is particularly useful in fingerprint PAD because real fingerprints exhibit natural micro-texture variations such as pores, ridge edge irregularities, and sweat patterns, whereas spoof materials tend to present smoother and more uniform textures [8].

Advantages:

- Simple and computationally efficient
- Low memory and processing requirements
- Effective for capturing local texture variations
- Works well with small datasets

Limitations:

- Sensitive to noise and image quality variations
- Limited ability to capture global structural information
- Weak generalization to unseen spoof materials
- Performance degrades across different sensors

2.5.2.2 Binarized Statistical Image Features (BSIF)

Binarized Statistical Image Features (BSIF) is a learned texture descriptor based on statistical properties of natural images. Unlike LBP, BSIF uses filters learned through Independent Component Analysis (ICA). These filters are applied to local image patches, and the filter responses are binarized depending on whether they are positive or negative. The resulting binary codes are then aggregated into histograms to represent the image texture [9].

Because BSIF uses learned filters instead of predefined rules, it can capture more complex and informative texture patterns compared to traditional handcrafted descriptors [9].

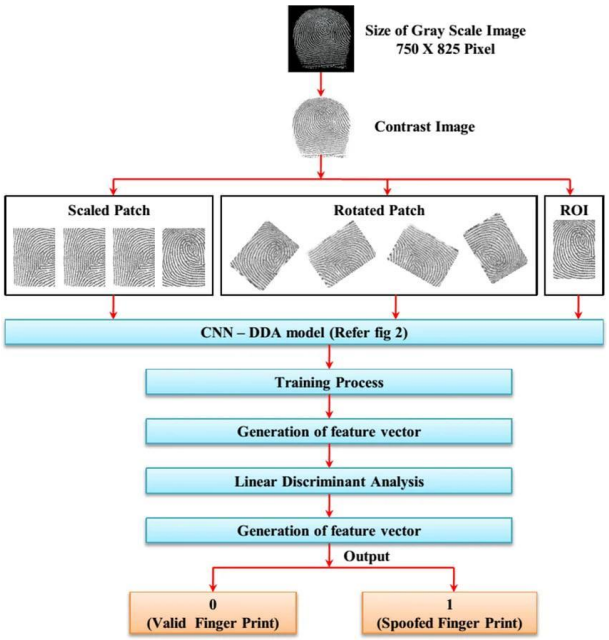
Advantages:

- Data-driven feature extraction (more expressive than LBP)
- Captures richer and more complex texture information
- Generally, more robust than basic handcrafted methods
- Performs well in texture classification tasks

Limitations:

- Depends on quality and representativeness of training data for filters
- Still lacks high-level semantic understanding
- More computationally expensive than LBP
- Limited generalization to unseen acquisition conditions

2.5.3 Deep Learning (CNN)



Convolutional Neural Networks (CNNs) have become the dominant approach in fingerprint Presentation Attack Detection (PAD) due to their ability to automatically learn hierarchical feature representations directly from raw images. Unlike handcrafted methods, CNNs eliminate manual feature design and learn discriminative patterns in an end-to-end manner [8].

CNNs are particularly effective for fingerprint PAD because they are sensitive to local micro-texture and ridge–valley structures. These patterns are often distorted in spoof fingerprints due to artifacts such as air bubbles, material granularity, surface irregularities, and abnormal reflectance. Convolutional filters can therefore learn to capture these inconsistencies for spoof detection [8].

Figure 2.5.1: CNN-DDA-based spoofed fingerprint detection

A typical CNN architecture consists of convolutional and pooling layers for feature extraction, followed by fully connected layers for classification. The final output is usually a binary decision (live/spoof) obtained using a SoftMax function. Some early approaches combined CNN feature extraction with separate classifiers (e.g., SVM), while more recent methods adopt fully end-to-end learning frameworks [8].

Patch-based strategies are widely used, where fingerprint images are divided into local regions or minutiae-centered patches. These local regions provide strong discriminative cues, and patch-level predictions are often aggregated using voting or score-level fusion. Advanced architectures such as Inception networks and lightweight models like MobileNet and SqueezeNet variants have also been applied to improve accuracy and reduce computational cost [8].

Advantages:

- Automatic feature learning without manual design
- High accuracy and strong discriminative power
- Captures both local and global fingerprint patterns
- Can be optimized in an end-to-end framework
- Effective with large and diverse datasets

Limitations:

- Requires large, labeled datasets for effective training
- High computational cost and training time
- Limited robustness to unseen spoof materials (domain shift)
- Risk of overfitting to dataset-specific or sensor-specific patterns
- Less suitable for low-resource or embedded systems without model compression

Conclusion

This chapter shows that fingerprint presentation attacks exploit fundamental differences between genuine skin and artificial materials. Attackers use materials such as silicone, gelatin, and lifted latent prints within a four-stage pipeline (capture, processing, reproduction, presentation) to deceive sensors.

Hardware defenses including multi-spectral and thermal sensing offer strong liveness detection by analyzing subsurface and physiological properties, while software methods like LBP, BSIF, and CNNs detect texture anomalies. However, no single countermeasure is sufficient; the most robust systems combine multiple hardware and software layers. The key remaining challenge is generalization: detection algorithms often fail against unseen attack materials or different sensors, demanding continuous benchmarking and adaptive solutions.

Chapter 3: Proposed PAD Detection

3.1 System Overview

Introduction

This section outlines the complete framework of the proposed fingerprint Presentation Attack Detection (PAD) system. The primary aim is to differentiate between genuine and spoof fingerprint samples through a deep learning strategy that combines Vision Transformer (ViT) feature extraction with Artificial Neural Network (ANN) classification. The proposed workflow comprises several consecutive phases: image preprocessing, feature extraction, classification, and final decision-making.

3.1.2 General Architecture

The proposed PAD system adopts a sequential processing pipeline to examine fingerprint images and identify presentation attacks. The procedure begins with the input fingerprint image, which is preprocessed to enhance its quality and align with the input specifications of the feature extraction model [8].

The enhanced image is subsequently passed into a pretrained Vision Transformer (ViT), tasked with extracting high-level, discriminative features from the fingerprint texture. These features are condensed into a compact feature vector that reflects both localized texture details and global structural patterns within the image.

This feature vector is then forwarded to an Artificial Neural Network (ANN) classifier, which performs binary classification to separate genuine fingerprints from spoof ones. The system then outputs a decision labeling the input fingerprint as LIVE or FAKE.

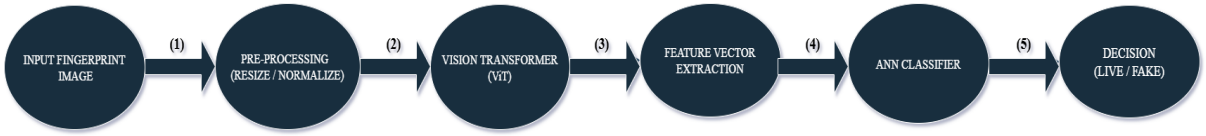


Figure 3.1.1: Proposed PAD System Architecture diagram.

3.1.3 Input Fingerprint Image

Fingerprint images serve as biometric data within authentication systems. In this study, the input samples are drawn from the LivDet 2015 and LivDet 2023 benchmark datasets, which are commonly employed for Presentation Attack Detection (PAD). These datasets contain both real fingerprints collected from genuine users and fake fingerprints fabricated from materials such as silicone, latex, and gelatin [2][5].

The images are captured using various sensors, leading to differences in quality, resolution, and acquisition conditions. Such diversity makes the datasets well-suited for assessing the robustness of biometric systems under realistic conditions.

The type of input data (whether genuine or spoof) plays a key role in system performance, as it directly affects the model's capacity to accurately differentiate between live and fake fingerprints.

3.1.4 Pre-processing Stage

Fingerprint images undergo a preprocessing stage that refines and optimizes captured data to enhance accuracy. It involves evaluating image characteristics, segmenting the area of interest, enhancing quality, and extracting discriminative features to ensure a consistent input format for the model.

The main goal of preprocessing is to enhance the quality of fingerprint images and prepare them for efficient analysis by the Vision Transformer (ViT) [8].

3.1.5 ViT-Based Feature Extraction

The Vision Transformer (ViT) is used in the proposed system as a feature extraction module. After the preprocessing stage, the fingerprint image is fed into the ViT model to extract deep and discriminative representations.

The ViT works by dividing the input image into a set of small non-overlapping patches. Each patch is then converted into a vector representation and processed as a sequence of tokens. Using a self-attention mechanism, the model learns relationships between different regions of the fingerprint image, allowing it to capture both local and global structural information [9].

Unlike traditional feature extraction methods, ViT can learn complex texture patterns present in fingerprint images, which are important for distinguishing between genuine and spoof samples. The output of this stage is a feature vector that represents the most relevant information extracted from the input image [8].

3.1.6 Feature Vector Generation

Following the processing of a fingerprint image by the Vision Transformer (ViT), the model generates a feature vector that encapsulates the most salient information extracted from the input sample. This vector is derived from the learned patch-based representations and the self-attention mechanism inherent to the transformer architecture, enabling the capture of both fine-grained textural details and overarching structural patterns within the fingerprint [8].

The resulting feature vector offers a compact yet highly discriminative numerical encoding of the fingerprint image. It encodes the essential attributes required to differentiate genuine from spoof samples and subsequently serves as the input to a downstream artificial neural network (ANN) classification stage, where the final decision is made [8].

3.1.7 ANN Classification

The feature vectors generated by the Vision Transformer are passed to an Artificial Neural Network (ANN) classifier, which is tasked with discerning discriminative patterns between genuine and spoof fingerprint samples. The ANN performs binary classification based on the representations extracted by the ViT module [8].

The network consists of an input layer, one or more hidden layers, and an output layer. The input layer receives the feature vector, while the hidden layers apply non-linear transformations via activation functions to capture complex relationships within the data. During training, the model updates its internal weights using backpropagation to minimize classification error and enhance generalization to unseen samples [8][9].

3.1.8 Decision Stage (LIVE / FAKE)

In the final stage, the ANN produces raw class scores. A SoftMax activation function is then applied to convert these scores into probability values representing the likelihood of each class (LIVE or FAKE). The system selects the class with the highest probability as the final decision.

The output of this stage constitutes the system's final classification result, determining whether the fingerprint is genuine or constitutes a presentation attack [8].

3.2 Proposed Method

The proposed system performs binary fingerprint liveness detection through a two-stage pipeline: a frozen pretrained Vision Transformer (ViT) that extracts patch-based feature representations from input fingerprint images, and a compact Artificial Neural Network (ANN) classifier that maps these representations to a binary liveness decision (LIVE vs. FAKE) via a SoftMax output layer. The architecture is designed to maximize classification performance while minimizing trainable parameters, training time, and risk of overfitting challenges that are particularly acute in fingerprint PAD due to the limited size of publicly available spoofing datasets.

3.2.1 Feature Extractor: Pretrained ViT (Frozen)

A standard Vision Transformer (ViT) backbone is adopted as the feature extractor. Following the architecture described by Dosovitskiy et al. and detailed in Raja et al., the input fingerprint image of dimensions (3 × 224 × 224) is divided into a grid of non-overlapping patches of fixed size (16 × 16) pixels. Each patch is linearly projected into a high-dimensional embedding vector. A learnable positional embedding is added to each patch token to preserve spatial location information within the original image. A special classification token ([CLS]) is prepended to the sequence of patch tokens, serving as an aggregate representation of the entire image [4][9].

The complete token sequence is passed through N transformer encoder blocks, each consisting of Multi-Head Self-Attention (MHSA) and a Multi-Layer Perceptron (MLP) sublayer, with layer normalization and residual connections. The MHSA mechanism enables the model to capture long-range dependencies between distant patches potentially encoding global fabrication artifacts not visible to convolutional receptive fields. The number of attention heads is set to 6, and the layer depth to 12 for the base configuration [4].

The ViT backbone is initialized with weights pretrained on the ImageNet dataset (or a large-scale pretraining corpus) and is kept entirely frozen during training that is, no gradient updates are applied to any ViT parameters. This design choice is motivated by two considerations. First, as demonstrated by Raja et al., pretrained ViT representations are highly transferable to fingerprint image tasks, and freezing the backbone reduces the risk of overfitting on small spoofing datasets. Second, freezing the feature extractor dramatically reduces the number of trainable parameters, lowering memory requirements and accelerating training. The output of the final transformer block specifically the [CLS] token embedding of dimension d = 768 is extracted and passed to the downstream ANN classifier. Alternatively, the mean-pooled patch token embeddings can be used as a richer patch-level feature representation [4].

3.2.2 Classifier: ANN for Binary Classification (LIVE vs. FAKE)

The ViT-extracted feature vector (of dimension d, typically 768 for the base model) is fed into a lightweight Artificial Neural Network (ANN) designed for binary classification. The network architecture consists of:

- Input layer: receives the ViT [CLS] embedding of dimension d = 768.
- Hidden Layer 1: fully connected layer with 256 neurons, ReLU activation function, and Dropout (p = 0.3) for regularization.
- Hidden Layer 2 (optional): fully connected layer with 64 neurons, ReLU activation, and Dropout (p = 0.2).
- Output layer: 2 neurons corresponding to the two classes (LIVE and FAKE).

The ANN is trained end-to-end using Cross-Entropy Loss, optimized with the AdamW optimizer. A polynomial learning rate decay schedule is applied, starting from an initial learning rate of 1e-4 and decaying with a power of 3 down to a minimum of 1e-5. Weight decay of 2e-5 is applied for regularization. The model is trained for up to 30 epochs with a batch size of 64 [3].

3.2.3 Decision Logic: SoftMax Classification

The raw output of the ANN's final linear layer (a 2-dimensional logit vector) is passed through a SoftMax activation function to produce a normalized probability distribution over the two classes. For an input fingerprint image x, let z = [z_live, z_fake] denote the logit vector. The SoftMax function computes:

$$\boldsymbol{P\ (LIVE \mid x) = \frac{exp(z_live)}{(exp(z_live) + exp(z_fake))}}$$

The predicted class is determined by:

$$\hat{c} = \boldsymbol{argmax}\{\boldsymbol{P}(\boldsymbol{LIVE} \mid \boldsymbol{x}), \boldsymbol{P}(\boldsymbol{FAKE} \mid \boldsymbol{x})\}$$

A sample is classified as LIVE if P (LIVE | x) > τ, where τ is a decision threshold defaulting to 0.5. In operational deployments, τ can be adjusted to trade-off between APCER and BPCER in accordance with the security requirements of the specific application, as recommended by ISO/IEC 30107-3. For example, a high-security access control scenario would lower τ to minimize APCER at the cost of increased BPCER [3][4][9].

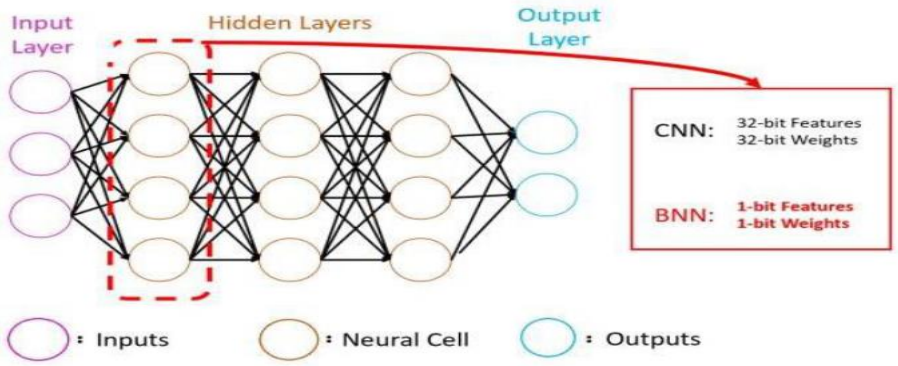


Figure 3.2.1: ANN (artificial neural network).

3.2.4 System Architecture Summary

The complete proposed pipeline operates as follows [9]:

- The input fingerprint image (3 × 224× 224) is partitioned into 14 × 14 = 196 non-overlapping patches of size (16 × 16) pixels.
- Each patch is projected into a d-dimensional embedding by the frozen ViT's patch embedding layer.
- Positional embeddings and the [CLS] token are added, and the resulting 197-token sequence is processed by N = 12 frozen transformer encoder blocks.
- The [CLS] token output (dimension d = 768) is extracted as the feature representation.
- The feature vector is passed through the ANN classifier (768 → 256 → 64 → 2). SoftMax is applied to the output logits to generate class probabilities.
- The final liveness decision (LIVE or FAKE) is made based on the predicted class probability and threshold τ.

3.2.5 Evaluation Metrics

Performance evaluation follows the ISO/IEC 30107-3 standard, as described by Raja et al. and Popli et al. The following metrics are reported [9]:

1.Attack Presentation Classification Error Rate (APCER): The proportion of presentation attack samples incorrectly classified as bona fide. A low APCER indicates strong attack detection capability.

2.Bona Fide Presentation Classification Error Rate (BPCER): The proportion of bona fide images incorrectly classified as attacks. A low BPCER indicates good retention of genuine users.

3.Average Classification Error Rate (ACER): The mean of APCER and BPCER, providing a balanced summary metric.

4.Equal Error Rate (EER): The threshold at which APCER equals BPCER, used as a summary of the overall detection-verification trade-off.

5.BPCER at fixed APCER (BPCER@APCER=5%, BPCER@APCER=10%): The bona fide error rate at a specified operational attack error rate, enabling comparison at standardized operating points.

3.3. Dataset (LivDet 2015 and LivDet 2023)

3.3.1 Dataset Presentation

To evaluate our system, we used two datasets from the International Fingerprint Liveness Detection Competition (LivDet): LivDet 2015 and LivDet 2023. LivDet is a biennial competition open to both academic and industry participants, aimed at advancing Fingerprint Presentation Attack Detection (PAD) research. Both datasets were provided as part of the competition framework [1], offering standardized fingerprint data collected from multiple sensors using different spoof materials, which enables direct comparison between our approach and other PAD methods reported in the literature [1].

LivDet 2015 is an earlier and widely cited edition of the competition, offering a solid comparative baseline for our method. LivDet 2023, the most recent edition at the time of this work, places a stronger emphasis on cross-sensor and cross-material generalization, a challenge that reflects real deployment conditions, where a PAD system trained in one environment may fail when moved to another. Together, the two datasets allow us to assess both our classifier's absolute performance and its consistency across different generations of benchmarks [2][5].

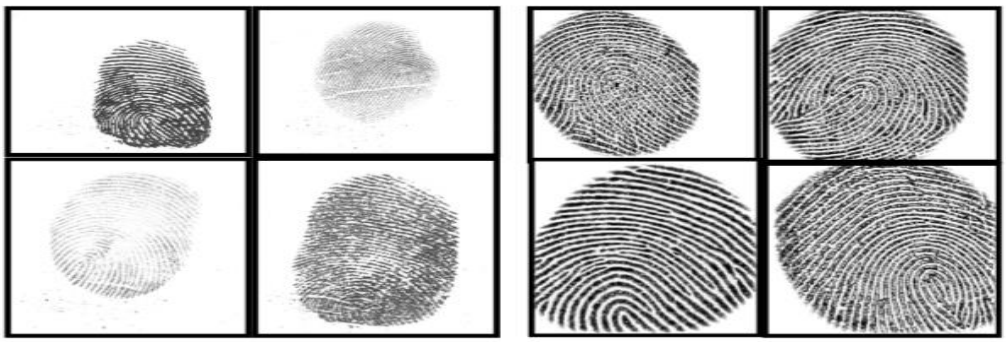
Examples of Live and Fake Fingerprint Samples



FAKE fingerprint samples from the LivDet 2023 dataset

LIVE fingerprint samples from the LivDet 2023 dataset

Figure 3.3.1: samples from the LivDet 2023 dataset.



FAKE fingerprint samples from the LivDet 2015 dataset

LIVE fingerprint samples from the LivDet 2015 dataset

Figure 3.3.2: samples from the LivDet 2015 dataset.

3.3.2 Sensors

The version of LivDet 2023 we worked with includes fingerprint images captured using two optical sensors: GreenBit DactyScan84C and Dermalog ZF1. Both produce high-resolution grayscale images at 500 dpi, which is the standard resolution in biometric identification systems. In practice, the two sensors produce noticeably different outputs. GreenBit delivers high-contrast images with sharp ridge definition, while Dermalog tends to produce slightly softer images with different background characteristics [5].

The LivDet 2015 dataset was captured using three sensors: Hi Scan, Digital Persona, and GreenBit. Each produces images with distinct characteristics in terms of contrast, resolution, and background noise, which further diversifies the range of acquisition conditions our model was exposed to [2]. Having sensors from two different competition years in our evaluation meant that our classifier had to handle real variation in image quality across both hardware generations and sensor types, a critical requirement for any PAD system intended for real-world deployment.

For both datasets, the data was provided to us as pre-extracted ViT feature vectors of 768 dimensions, and our work focused on training and evaluating the ANN classifier on top of these representations.

3.3.3 Dataset Structure (LIVE / FAKE)

Both datasets are divided into two classes: LIVE samples, collected from genuine human fingers, and FAKE samples, generated from artificial spoof materials such as silicone, gelatin, and latex. These materials simulate real-world fingerprint presentation attacks and are widely used in PAD research. The two datasets present different class distributions: in LivDet 2023, FAKE samples slightly outnumber LIVE samples, while in LivDet 2015, LIVE samples significantly outnumber FAKE samples (4000 vs 1000 per sensor), producing a more pronounced imbalance in the opposite direction. We did not apply any re-balancing to either dataset, since our goal was to evaluate our classifier under conditions that are as close as possible to the original competition protocols and to real deployment scenarios [2][5].

The distribution of samples for the LivDet 2023 dataset is presented in Table 3.1, and for the LivDet 2015 dataset in Table 3.2.

.

Dataset Split	**LIVE**	**FAKE**	**Total**
Training	875	1050	1925
Test	375	450	825

Table 3.1 Distribution of LIVE and FAKE samples in the LivDet 2023 dataset.

Dataset Split	**LIVE**	**FAKE**	**Total**
Training	4000	1000	5000
Test	4000	1000	5000

Table 3.2 Distribution of LIVE and FAKE samples in the LivDet 2015 dataset.

For both datasets, we followed the official LivDet train/test partitioning protocols. The training set was used to learn the ANN classifier parameters, while the test set was kept entirely separate and used only for final performance evaluation, ensuring that our reported metrics reflect genuine generalization rather than overfitting to seen data.

Overall, the diversity of sensors and spoof materials across LivDet 2015 and LivDet 2023 made them demanding and well-suited benchmarks for our work. Using both allowed our results to be directly compared against other methods evaluated on the same standards.

3.3.4 Split Strategy

The training set was used to learn the ANN classifier parameters, while the test set was kept entirely separate and used only for final performance evaluation, ensuring that our reported metrics reflect genuine generalization rather than overfitting to seen data [9].

Overall, the diversity of sensors and spoof materials across LivDet 2015 and LivDet 2023 made them demanding and well-suited benchmarks for our work and using both allowed our results to be directly compared against other methods evaluated on the same standards [2][5][9].

3.4. Experimental Setup

The presentation attack detection system was built and examined on a set of fixed equipment.

3.4.1 Hardware Equipment

The experiments were performed on a CPU-only system to simulate resource-constrained deployment environments. The hardware configuration consisted of an Intel® Core™ i5-1240P processor (base frequency 1.70 GHz, turbo up to 4.40 GHz, 12 cores, 16 threads), 16 GB of LPDDR5 RAM, a 239 GB NVMe SSD, and the Windows 11 Pro operating system.

3.4.2 Framework

The system was implemented in Python 3.10 using PyTorch 2.11, along with Scikit-learn, NumPy, Pandas, and Matplotlib, within a dedicated virtual environment (venv).

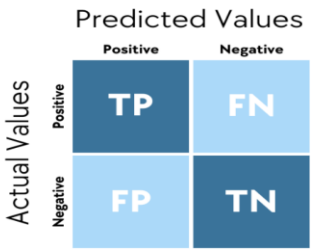
3.4.3 Evaluation Protocol

The experiments were executed in three steps (training, validation, testing) on different datasets using different sensors. Features were extracted using a ViT (vision transformer) and a CNN (convolutional neural network). The models were trained using three fully connected layers, with a dropout layer between each pair to prevent overfitting. We used the BCEWithLogitsLoss loss function, and the Adam optimizer, with a maximum of 40 epochs and a learning rate 0.0001. The train set was divided into batches during training (batch size = 25) to improve efficiency.

The decision threshold was set to 0.5 on the sigmoid output of the ANN.

3.4.4 Metrics

To evaluate model performance, we used a confusion matrix to compute four characteristics (APCER, BPCER, ACEURACY, ACE) [8] [4].



1-APCER (Attack Presentation Classification Error Rate): measures the proportion of attack presentations (spoof/fake fingerprints) that are incorrectly classified as bona fide (live) presentations.

2-BPCER (Bona Fide Presentation Classification Error Rate): measures the proportion of bona fide (genuine/live) presentations that are incorrectly classified as attacks (fake).

3-ACCURACY: is the overall proportion of correct classifications (both live and fake samples) out of all presented samples

4-ACE: is widely used in LivDet competitions and research papers because it gives equal importance to security (APCER) and usability (BPCER).

3.5 Results and Discussion

3.5.1-Result Tables and Plots

Dermalog Sensor (LivDet 2023):

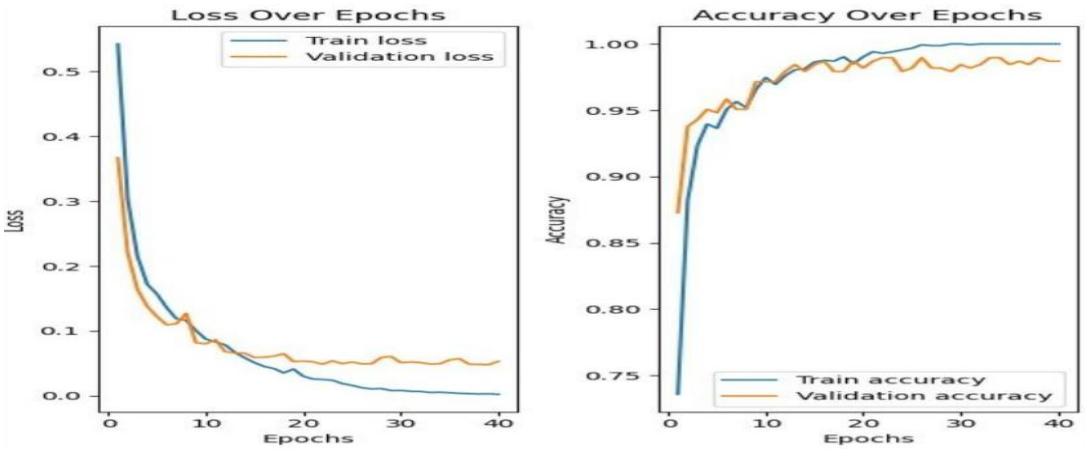


Figure 3.5. 1: Loss&Accuracy in Train&Validation Dermalog Sensor (LiveDet 2023)

Dataset	Sensor	Performance	Train %	Validation %	Test %
LivDet 2023	Dermalog	Accuracy	100	98.96	98.06
		APCER	0	0	1.56
		BPCER	0	2.29	1.6
		ACE	0	1.14	1.98

Table 3.1: Performance metrics for the Dermalog Sensor (LivDet 2023).

GreenBit Sensor (LivDet 2023):

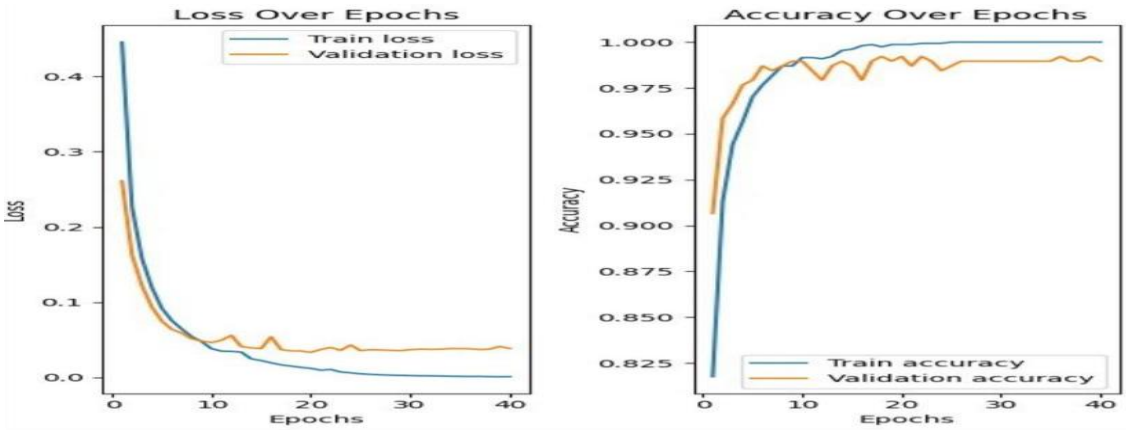


Figure3.5. 2: Loss&Accuracy in Train&Validation GreenBit Sensor (LivDet 2023)

Dataset	Sensor	Performance	Train %	Validation %	Test %
LivDet 2023	GreenBit	Accuracy	100	98.96	98.55
		APCER	0	0.95	1.33
		BPCER	0	1.14	1.6
		ACE	0	1.05	1.47

Table 3.2: Performance metrics of the GreenBit Sensor (LivDet 2023)

Digital Persona Sensor (LivDet 2015):

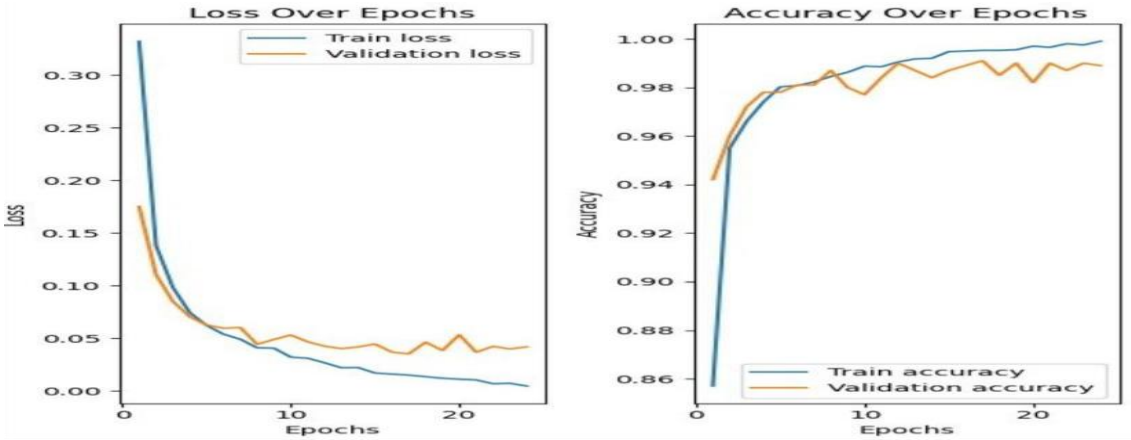


Figure 3: Loss&Accuracy in Train&Validation Digital Persona Sensor (LivDet 2015)

Dataset	Sensor	Performance	Train %	Validation %	Test %
LivDet 2015	Digital Persona	Accuracy	100	98.9	94.74
		APCER	0.03	1.00	2.75
		BPCER	0.37	1.5	15.3
		ACE	0.2	1.25	9.02

Table 3.3: Performance metrics of the Digital Persona Sensor (LiveDet 2015)

Hi Scan Sensor (LivDet 2015):

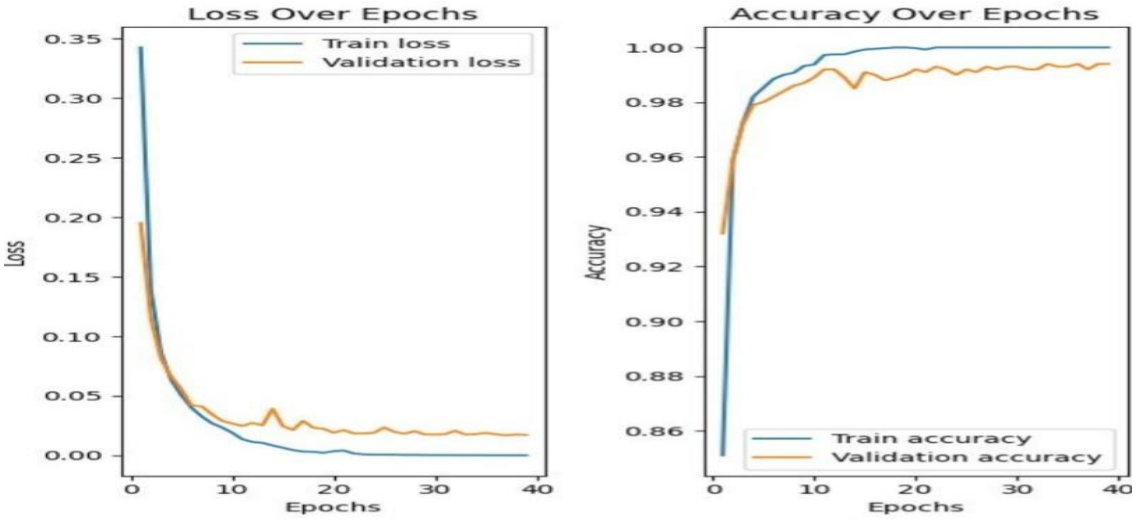


Figure 4: Loss&Accuracy in Train&Validation Hi Scan Sensor (LivDet 2015)

Dataset	Sensor	Performance	Train %	Validation %	Test %
LivDet 2015	Hi Scan	Accuracy	100	99.4	95.74
		APCER	0	0.12	1.87
		BPCER	0	2.5	13.8
		ACE	0	1.31	7.84

Table 3.4: Performance metrics of the Hi Scan Sensor (LivDet 2015)

Greenbit Sensor (LivDet 2015):

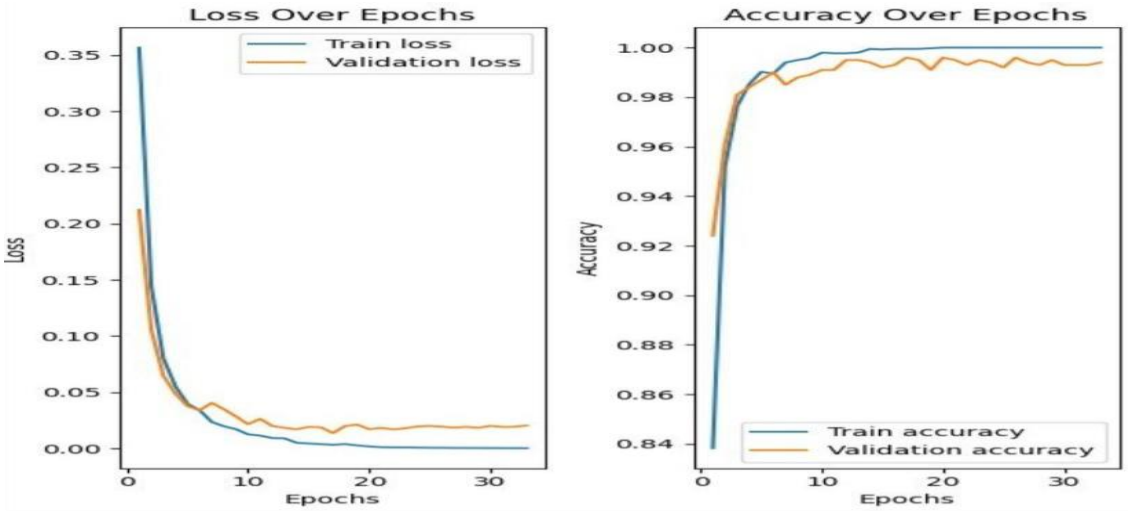


Figure 3.5. 5: Loss&Accuracy in Train&Validation GreenBit Sensor (LiveDet 2015)

Dataset	Sensor	Performance	Train %	Validation %	Test %
LivDet 2015	Greenbit	Accuracy	100	99.3	96.2
		APCER	0	0.37	0.83
		BPCER	0	2.00	15.7
		ACE	0	1.19	8.26

Table 5: Performance metrics of the GreenBit Sensor (LivDet 2015)

3.5.2 Comparison with Baseline

The features of LivDet 2023 are extracted using the Vision Transformer (ViT) architecture, whereas the features of LivDet 2015 are extracted using a Convolutional Neural Network (CNN) architecture. As shown in the figures above, the ViT-based approach achieves superior performance and efficiency, primarily due to its capacity to model long-range dependencies, which makes it particularly suitable for fingerprint presentation attack detection.

Discussion

The obtained results demonstrate that the proposed fingerprint Presentation Attack Detection (PAD) system achieved strong performance on both LivDet 2023 and LivDet 2015 datasets. For LivDet 2023, the Dermalog and GreenBit sensors produced excellent results, with stable training and validation curves indicating effective learning and good generalization without severe overfitting. The proposed model achieved test accuracies of 98.06% and 98.55% respectively, with low APCER, BPCER, and ACE values, confirming strong spoof detection capability while maintaining good usability for genuine users.

In contrast, the LivDet 2015 dataset produced slightly lower testing performance, particularly in BPCER values, although APCER values remained relatively low across all sensors. This indicates that the system still detects spoof fingerprints effectively but faces more difficulty with genuine fingerprint classification, likely due to older sensor technology and higher dataset variability.

A clear difference was also observed between the feature extraction approaches. The ViT-based approach used in LivDet 2023 consistently achieved better accuracy and lower ACE values than the CNN-based approach used in LivDet 2015. This improvement is mainly due to the ability of Vision Transformers to capture both local and global fingerprint patterns more effectively. Overall, the experimental results confirm that the proposed ViT-ANN framework provides a robust and reliable solution for fingerprint presentation attack detection.

Conclusion

The proposed fingerprint Presentation Attack Detection (PAD) system leverages a pretrained Vision Transformer (ViT) for robust feature extraction combined with a lightweight Artificial Neural Network (ANN) for binary classification. The architecture processes fingerprint images through preprocessing, patch-based ViT feature representation, and a fully connected classifier with dropout regularization to distinguish between live and fake samples. Extensive experiments conducted on the LivDet 2023 and LivDet 2015 datasets across multiple sensors (Dermalog, GreenBit, Digital Persona, and Hi Scan) demonstrated excellent performance. On the LivDet 2023 dataset using ViT features, the system achieved test accuracies exceeding 98%, with notably low APCER and ACE values reaching as low as 1.47%. These results significantly outperform the CNN-based approach applied to the LivDet 2015 dataset, highlighting the advantage of ViT's long-range modeling capabilities in capturing critical global patterns in fingerprint images. The consistently low APCER across all sensors reflects strong security against presentation attacks, while the overall balanced metrics confirm the system's practical viability. Training dynamics remained stable with minimal overfitting, further supporting the reliability of the proposed methodology.

General Conclusion

This research addressed a critical vulnerability in modern biometric authentication infrastructures: the threat of presentation attacks using common physical spoofing materials such as silicone, gelatin, and latex. To build a robust defense, we developed and evaluated a two-stage pipeline combining a frozen **Vision Transformer (ViT-B/16)** architecture with a lightweight **Artificial Neural Network (ANN)** classifier.

The experimental results validate this design strategy. By maintaining a frozen, ImageNet-pretrained ViT backbone, the system avoids the need for massive, specialized biometric training data. This architectural choice reduced trainable parameters by approximately 99%, mitigating overfitting on narrow spoof subsets and resolving a major real-world implementation barrier: allowing the complete pipeline to run on basic CPU hardware without sacrificing model depth. More importantly, the self-attention mechanism proved highly effective at identifying macro-level manufacturing flaws and long-range spatial anomalies across fingerprint patches that traditional convolutional frameworks regularly miss.

When tested against official LivDet benchmarks across five diverse sensor environments, the model demonstrated solid defensive capabilities, achieving a peak classification accuracy of **98.55%** and a restricted **APCER of 1.33%** on modern GreenBit hardware. Minimizing the APCER is central to biometric security, as it directly reduces the likelihood that an unauthorized spoof attempt will bypass structural defenses.

However, the experimental data also highlights clear operational boundaries. Performance drops on older sensor models visible in the sharp rise of BPCER to 13–16% within the LivDet 2015 dataset reveal a sensitivity to acquisition shifts caused by variation in legacy hardware configurations. Furthermore, ensuring cross-material generalization when confronting entirely novel, unseen fabrication materials remains an ongoing challenge.

Bibliography

1. D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*, 2nd ed. London: Springer, 2009.
2. L. Ghiani, D. Yambay, V. Mura, G. L. Marcialis, F. Roli, and S. Schuckers, "LivDet 2015 Fingerprint Liveness Detection Competition 2015," *2015 International Conference on Biometrics (ICB)*, pp. 1–6, 2015.
3. J. Galbally, J. Fierrez, R. Cappelli, and G. L. Marcialis, "Introduction to Presentation Attack Detection in Fingerprint Biometrics," in *Handbook of Biometric Anti-Spoofing*, 3rd ed., S. Marcel, J. Fierrez, and N. Evans, Eds. Singapore: Springer Nature, 2023, pp. 3–20.
4. ISO/IEC 30107-3:2017, "Information Technology Biometric Presentation Attack Detection Part 3: Testing and Reporting," International Organization for Standardization, 2017.
5. R. Casula, M. Micheletto, G. Orrù, G. L. Marcialis, and S. C. Schuckers, "LivDet2023 Fingerprint Liveness Detection Competition: Advancing Generalization," *arXiv preprint arXiv:2309.15578*, 2023.
6. A. K. Jain, A. A. Ross, and K. Nandakumar, Handbook of Biometrics. New York, NY: Springer, 2008.
7. Meng, X., Jiang, Y., Wang, X., & Lu, H. (2023). A Review of Fingerprint Sensors: Mechanism, Characteristics, and Applications. *Micromachines*, 14(6), 1253.
8. Li H, Ramachandra R. A Survey on Deep Learning Techniques for Fingerprint Presentation Attack Detection. *Sensors*. 2026;26(4):1283. doi:10.3390/s26041283.
9. Sumalatha, U., Prakasha, K. K., Prabhu, S., & Nayak, V. C. (2024). A Comprehensive Review of Unimodal and Multimodal Fingerprint Biometric Authentication Systems: Fusion, Attacks, and Template Protection. *IEEE Access*, 12, 64300–64334. doi:10.1109/ACCESS.2024.3395417.