Order N$^o$ : ../../..

**PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA**
**MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH**
**UNIVERSITY KASDI MERBAH OF OUARGLA**

**Faculty of New Information Technologies and Communication**
**Department of Electronics and Telecommunications**

# ACADEMIC MASTER'S DEGREE

Field : **Sciences and Technology**

Branch : **Electronics**

Option : **Electronics of Embedded systems**

presented By :

**Benglia Aymen**

**Khelifa Abd EL Karim**

## *Theme*

# Person Identification Biometric Systems from Local Finger-Knuckle Prints Based on Deep Learning

Publicly defended on : 25/06/2024 before the jury composed of :

| | | | |
|---|---|---|---|
| *LATI Abdelhai* | *MCA* | *University Kasdi Merbah - Ouargla* | *President* |
| *Djamel SAMAI* | *Pr.* | *University Kasdi Merbah - Ouargla* | *Supervisor* |
| *Nasri Nadjib* | *MAA* | *University Kasdi Merbah - Ouargla* | *Examiner* |

# Dedication

**I dedicate this significant event in my life to :**

**My parents, who have been my steadfast supporters and the guiding lights of my life. My dear brother and sisters, whose unwavering belief and support have been invaluable.**
**To my colleague Benglia Aymen I wish him a bright and happy future.and my friends and all the acquaintances I met during my university studies.**
**My professor, for believing in me and providing the support and guidance needed to write this productive and enjoyable thesis.**
**I extend my heartfelt gratitude to my colleague Benglia Aymen for his invaluable assistance. I wish him a bright and successful future.**

***KHELIFA ABD EL KARIM***

# Dedication

**With sincere gratitude and admiration, I dedicate this modest work :**

**to my wonderful mother, whose unwavering care, sincere prayers, unending love, and unmatched tenderness have motivated my efforts to finish this research. To my dad, who urged me to seek after my fantasies and imparted in me the upsides of persistence and difficult work. to my brothers and sister, my life's guiding lights, whose unwavering love and support have given me strength and direction**

**To my colleague KHELIFA ABD EL KARIM who helped me.I wish him a happy future,to the director of this theme prof. SAMAI Djamel, we thank him for having guided, directed, helped and advised us.**

**Last but not least, I'd like to express my sincere gratitude to my friends and coworkers for all of their patience with us.**

***AYMEN BENGLIA.***

# Acknowledgments

First of all, we express our gratitude to God who gave us patience and courage throughout these years of study. We extend our sincere thanks to Prof. Samai Djamel, who supervised this project and provided invaluable advice and basic guidance. We are very grateful to the professors working in the labo Ingénierie électrique at the Scientific Research Center for their patience and efforts. And their great contributions to the development of this research. Their support was instrumental in achieving the positive results of our study and providing all the necessary resources. We would also like to thank the jury members, Mr. Nasari Nadjib and DR. Lati Abdelhay, for their willingness to review and evaluate our work. In addition, we take this opportunity to express our sincere thanks to all teachers for their cooperation, presence and compassion throughout our education.

***To all my colleges and friends– Thank you so much.***

# الملخص

تتضمن القياسات الحيوية تحديد هوية الأفراد تلقائيًا باستخدام خصائصهم الفسيولوجية أو السلوكية. وتعتبر الأنظمة البيومترية الحيوية متعددة الوسائط التي تدمج أساليب التعرف على طرائق متعددة الحل الأمثل لتحديد الهوية بدقة. يركز هذا البحث على استخدام بصمات مفاصل الأصابع لتحديد الهوية البيومترية وتقييم أداء نموذج سي إن إن المقترح. تستكشف الدراسة تقنيات مختلفة لاستخراج السمات وطرق التصنيف، بهدف تحديد التركيبة الأكثر فعالية لبصمات مفاصل الأصابع لتحديد الهوية.

**الكلمات المفتاحية** التعرف، التعلم العميق، SVM ، القياسات الحيوية ، CNN ، التعلم المنقول، أحادي الوضع، متعدد الوضع، FKP .

# *Abstract*

Biometrics involves automatically identifying individuals using their physiological or behavioral characteristics. Multimodal biometric systems integrating multiple modalities recognition methods are considered the optimal solution for accurate identification. This research focuses on using knuckle prints for biometric identification and evaluates the performance of a proposed CNN model. The study explores different feature extraction techniques and classification methods, aiming to determine the most effective combination of knuckle prints for identification.

**Keywords :** Recognition, deep learning, SVM, biometrics, CNN, transfer learning, unimodal, multimodal, FKP.

# *Résumé*

La biométrie consiste à identifier automatiquement des individus à l'aide de leurs caractéristiques physiologiques ou comportementales. Les systèmes biométriques multimodaux intégrant plusieurs méthodes de reconnaissance sont considérés comme la solution optimale pour une identification précise. Cette recherche se concentre sur l'utilisation des empreintes des articulations pour l'identification biométrique et évalue les performances d'un modèle CNN proposé. L'étude explore différentes techniques d'extraction de caractéristiques et méthodes de classification, dans le but de déterminer la combinaison la plus efficace d'empreintes de phalanges pour l'identification.

**Mots-clés :** Reconnaissance, apprentissage profond, SVM, biométrie, CNN, apprentissage par transfert, unimode, multimode, FKP.

# Contents

# List of Figures

# *Abbreviations*

| | |
|---|---|
| **AI :** | Artificial Intelligence |
| **CMC :** | Cumulative Match Characteristic |
| **CNN :** | Convolution Neural Network |
| **DL :** | Deep Learning |
| **EER :** | Equal Error Rate |
| **FAR :** | False Accept Rate |
| **FKP :** | Finger Knuckle Print |
| **FRR :** | False Reject Rate |
| **GAR :** | Genuine Acceptance Rate |
| **ML :** | Machine Learning |
| **ReLU :** | Rectification Linear Unit |
| **ROC :** | Receiver Operating Characteristic |
| **ROR :** | Rank One Recognition |
| **RPR :** | Rank of Perfect Recognition |
| **SUM :** | Summation |
| **SVM :** | Support Vector Machine |
| **T0 :** | Threshold |

Chapter 1

# General introduction

The security and safety of people, property, and information are paramount societal concerns. Traditional identity verification and identification methods, such as passports, access cards, usernames, and passwords, are increasingly vulnerable to compromise. Passwords can be forgotten or hacked, and access cards can be falsified or duplicated, leading to identity theft. In response to these challenges, biometric authentication has emerged as a robust solution, leveraging physiological and behavioral traits to verify identities. Biometrics, encompassing traits such as facial features, fingerprints, iris patterns, and voice, ensures higher security and convenience as it requires physical presence for identification, significantly reducing the risk of identity theft.

Single-mode or unimodal biometric systems, although beneficial, face limitations such as low public acceptance and high error rates. This has led to the development of multimodal biometric systems, which integrate multiple biometric sources to enhance accuracy and reliability. One notable system in this realm is the Finger Knuckle Print (FKP) system, which uses finger joint impressions. FKP systems are user-friendly and have proven to be effective in establishing reliable biometric systems when combining various modalities .

Concurrently, the fields of artificial intelligence (AI), machine learning (ML), and deep learning (DL) have seen substantial advancements. AI enables machines to perform tasks

akin to human capabilities, while ML focuses on enabling computers to learn from data with minimal human intervention. Deep learning, a subset of ML, utilizes artificial neural networks to process large amounts of data, mimicking human brain functions. These technologies have revolutionized various sectors, including business, education, manufacturing, banking, military, and healthcare.

Deep learning has significantly enhanced diagnostic assistance algorithms and biosignal processing methods such as electroencephalography and electrocardiography in the medical field. For medical image processing, deep learning techniques have improved the accuracy of automatic characterization, segmentation, and classification. Convolutional neural networks (CNNs), inspired by human brain function, are particularly effective for image classification tasks, often surpassing traditional methods in performance and accuracy.

This work integrates unimodal and multimodal biometric identification systems with advanced deep-learning techniques to develop more secure and efficient authentication systems. By exploring the Finger Knuckle Print system and leveraging CNNs for image classification, this research aims to contribute to the growing field of biometric security and its applications.

This chapter is organized as follows:

**THE FIRST CHAPTER** General Introduction

**THE SECOND CHAPTER** explains the fundamental ideas of biometrics, as well as its properties, criteria, and classifications. We also present the architecture and major module of a biometric system, as well as the processes involved in biometric authentication systems. The chapter concludes with the performance evaluation of biometric systems.

**THE THIRD CHAPTER** covers deep learning, delving deeper into transfer learning (TL) concepts, which is how we conducted our research, and convolutive neural networks (CNN).

**THE FOURTH CHAPTER** shows the results of the final FKP picture recognition experiments. Next, based on transfer learning, we explore a deep learning technique

for people identification employing four pre-trained CNN networks (AlexNet, VGG19, ResNet50, and DenseNet201).

**THE FIFTH CHAPTER** covers the general conclusions and future work. It summarizes the main findings of the study and provides recommendations for further research in this area.

Chapter 2

# Basic of Biometrics

## 2.1 Introduction

The fast growth of modern human civilization has led to an increasing demand for new and efficient technologies to sustain it. Alongside, security and privacy concerns have emerged, and the usage of highly reliable and accessible individual authentication and identification techniques has become crucial. Biometrics has emerged to address this need and has become a science that studies the physiological and behavioral characteristics of the human body to recognize an individual's identity [1]. This chapter will provide an introduction to biometry and its many modalities. First, we will define and discuss biometrics. We then go on to discuss the various biometric modalities. We then present the metrics used to evaluate biometric systems, along with their meanings and other associated information.

## 2.2 definition of biometrics

Biometrics is the quantitative analysis of biological or behavioral characteristics to identify or verify an individual's identity [2].

Instead of relying on traditional identification methods such as passwords or identification cards, biometric systems analyze distinctive features like fingerprints, facial characteristics, iris patterns, or voiceprints. By capturing and comparing these unique traits, biometrics provides a secure and efficient means of authentication (Figure 2.1).

**Figure 2.1** – Biometric Technology

## 2.3 Biometrics classifications

The following categories are the most widely used biometric methods; however, they are applied in many other fields (Figure 2.2).

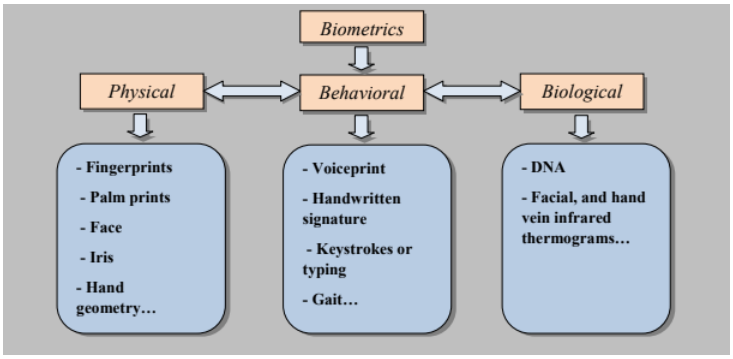



**Figure 2.2** – Biometrics categories

### 2.3.1 Physical characteristics

Anything that has to do with how the different organs look on the outside interests them, such as the following:

#### 2.3.1.1 Fingerprint

Fingerprints are the patterns of ridges on our fingers' tips. It is one of the most mature biometric technologies and is considered a legitimate proof of evidence in courts of law all over the world [3] (Figure 2.3).

**Figure 2.3** – images of fingerprint

#### 2.3.1.2 Finger Knuckle Print (FKP)

People commonly utilize FKP as a biometric due to its ease of use and low cost. The FKP surface has different patterns in terms of direction, points of detail, and pores, and these features are unique to all people. However, the identification process using the FKP image may still encounter challenges due to variations in the image's lighting quality. Because of the noisy sensor data, the extraction of features is one of the most crucial stages of FKP image recognition. This leads to an increase in research in this field in order to create a system that can be used for accurate identification [4](Figure 2.4).

**Figure 2.4** – Biometric system based on finger joints

#### 2.3.1.3 Palmprint

Palmprint recognition is based on palm texture, which has unique line features. It can also extract relatively stable features from low-resolution images and has strong anti-noise ability. Compared with other commonly used biometric recognition technologies, palmprint recognition has many unique advantages. Unlike face recognition [5], palmprint features remain relatively stable and unaffected by ornaments, expressions, and gestures. Compared with fingerprint technology [6], the effective area of a palmprint is much larger and contains more information.(Figure 2.5).

**Figure 2.5** – Images of palm print

#### 2.3.1.4 Iris

The iris is a thin, circular structure in the eye that is a protected internal organ, so it is not affected by environmental conditions [7]. Amongst all the biometric recognition systems, Iris is the most promising solution because of its uniqueness, reliability, and stability over its lifetime. Even the genetically identical twins have different iris textures [8](Figure 2.6).

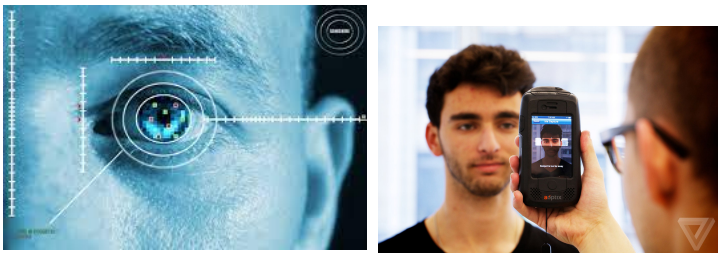


**Figure 2.6** – images of fingerprint

#### 2.3.1.5 Face

The face contains many biometric features based on the face, such as the gap between the two eyes, the geometry of the face, the facial thermography, or the width of the mouth [9]. Face recognition works by capturing and analyzing the unique facial features of an individual. These features include the distance between the eyes, the shape of the nose, and the contours of the face. This information is then transformed into a digital representation, creating what is commonly known as a "face template"(Figure 2.7).


**Figure 2.7** – Face recognition

### 2.3.2 Behavioral characteristics

This kind concentrates on analyzing an individual's behavior:

#### 2.3.2.1 Keystrokes

Every person types on a keyboard in a unique way, according to a theory. Although it is not anticipated that this behavioral biometric would be exclusive to each person, it does provide enough discriminating information to enable identity verification. Keystroke dynamics is a behavioral biometric; one could expect to see significant differences in certain people's usual typing patterns. Furthermore, a system may discreetly record someone's keystrokes as they enter data [10] (Figure 2.8).

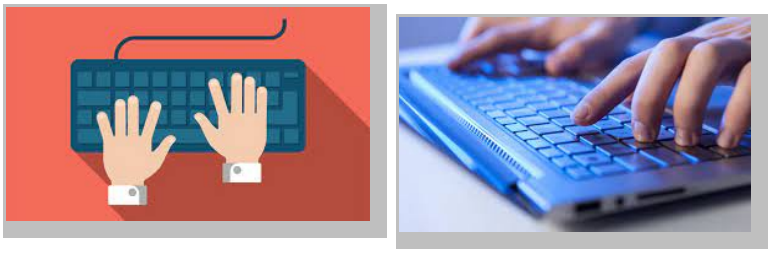


**Figure 2.8** – Typing image

#### 2.3.2.2 Voice

With a passphrase, voice or speaker recognition recognizes people based on their vocal characteristics. One of the many inexpensive and easily deployable technologies is the ability to use a telephone or microphone as a sensor. Environmental elements like background noise, however, might have an impact on speech recognition. In order to increase dependability, the U.S. government's intelligence community and the telecoms sector have dedicated a significant amount of resources to this technology [11] (Figure 2.9)

**Figure 2.9** – voice recognition image

#### 2.3.2.3 Signature

Dynamic Signature: Everyone has their own writing style. We can define a model to identify a person based on their signature. Since signatures are used in many countries as a legal or administrative element, they are used to justify a person's position or to confuse a person with previously signed documents [12](Figure 2.10).

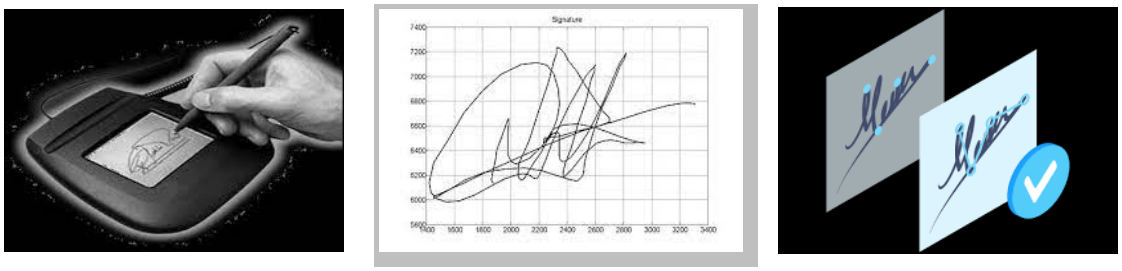


**Figure 2.10** – Signature images

#### 2.3.2.4 Gait

Gait is a sophisticated spatiotemporal biometric that describes each individual's unique walking style. Gait, while not highly precise, is sufficiently discriminating in some low-security situations to allow verification. As a behavioral biometric, gait may change over time, particularly in response to significant injuries to the joints or brain, changes in body weight, or intoxication. Since learning a person's stride is like learning a face, gait analysis might be a valid biometric. Gait-based systems are computationally costly and need a lot of data since they analyze several distinct motions of each articulating joint using video-sequence footage of a walking human [10](Figure 2.11).

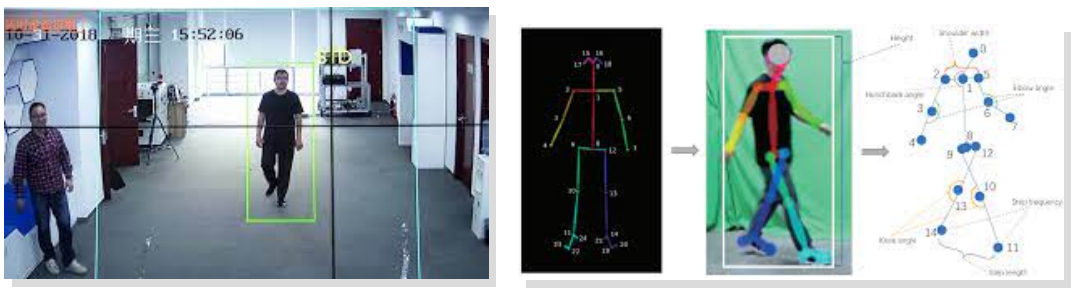


**Figure 2.11** – Images of gait system

### 2.3.3 Biological characteristics

This type of biometric modality, incorporating distinctive features like DNA, saliva, and odor, serves as a robust foundation for precise individual identification.

#### 2.3.3.1 DNA

Forensic applications primarily use DNA, a unique identity code, for person recognition. However, it faces issues like contamination, sensitivity, automatic real-time recognition, and privacy concerns. DNA theft can lead to unauthorized use, and current technology isn't suitable for online noninvasive recognition[10](Figure 2.12).

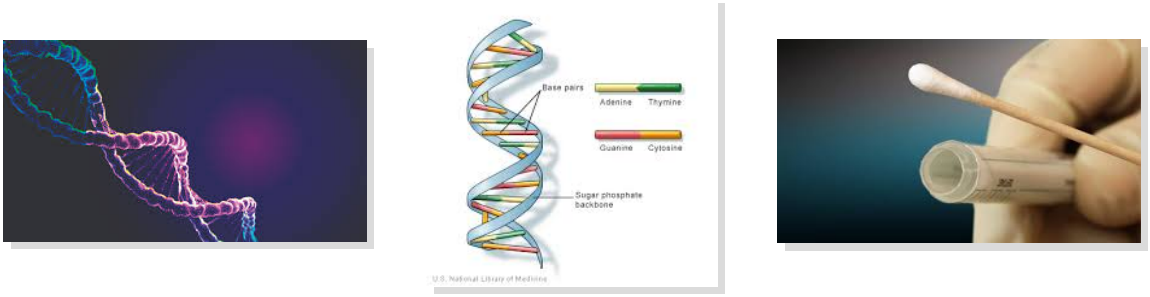



**Figure 2.12** – DNA images

## 2.4 Properties of biometric modalities

Biometric modalities have some properties that differ from one modality to another.

- **Universality:** The entire community must possess the same modality.
- **Distinctiveness:** In that aspect, any two individuals should differ enough from one another.
- **Consistency:** It implies that the characteristic should not vary significantly over time or in reaction to external factors.
- **Collectability:** It means that the trait is statistically measurable.
- **Performance:** Biometric identification must be precise, quick, and resilient to environmental and operational changes.
- **Acceptance:** This expresses how much acceptance there is for the use of a specific biometric identifier (characteristic) in day-to-day life.
- **Circumvention:** It should be difficult to manipulate the biometric modality.

| Biometric | Universality | Distinctiveness | Consistency | Collectability | Performance | Acceptance | Circumvention |
|---|---|---|---|---|---|---|---|
| Iris | High | High | High | Medium | High | Low | Low |
| Fingerprint | Medium | High | High | Medium | High | Medium | Medium |
| Face | High | Low | Medium | High | Low | High | High |
| DNA | High | High | High | Low | High | Low | Low |
| Voice | Medium | Low | Low | Medium | Low | High | High |
| Gait | Medium | Low | Low | High | Low | High | Medium |
| Keystroke | Low | Low | Low | Medium | Low | Medium | Medium |
| Signature | Low | Low | Low | High | Low | High | High |

**Figure 2.13** – Biometric Characteristics Evaluation

## 2.5 Operating modes of a biometric system

A biometric system can operate in either verification or identification mode. Both modes essentially necessitate a recruitment process that records the biometric data acquired.

Enrollment: Entering an individual's biometric information into the system's database is the first step in the enrollment process. A biometrical reader records an individual's biometric characteristics during the enrollment process. After that, a quality measurement is often performed to guarantee the acquisition's high quality [9](Figure 2.14).

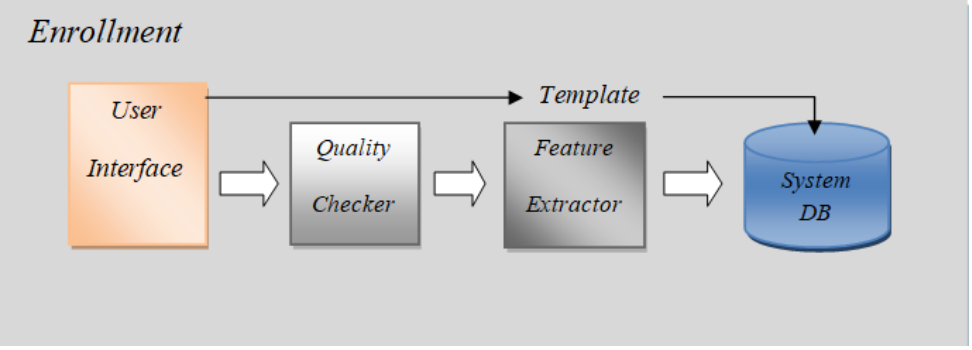



**Figure 2.14** – mode Enrollment

### 2.5.1 Verification

VThe process of verifying an individual's claimed identity involves comparing their collected biometric data to matching templates recorded in the system database. This process is known as identity verification.

A person looking for recognition in such a system presents their identity using a PIN or smart card, and the system compares the two in a one-to-one fashion to ascertain

whether the claim is accurate or not (by providing an answer to the question, "Does this biometric data really correspond to Mr. Bob?"). Identity verification is a crucial process that tries to prevent many persons from using the same identity [13] (Figure 2.15).

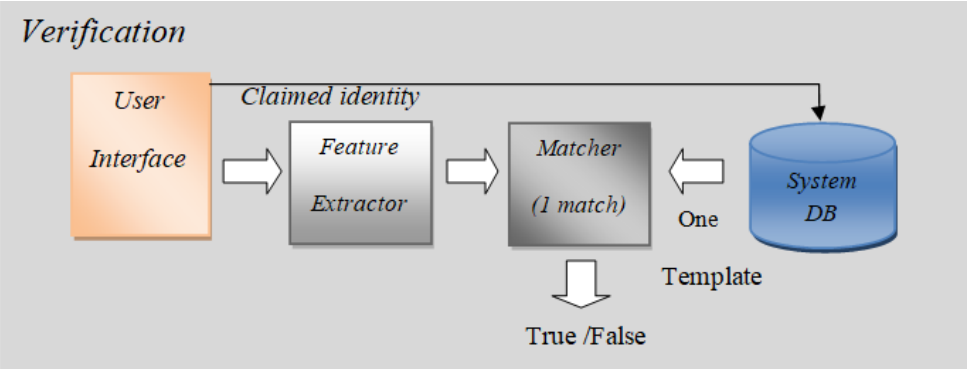



**Figure 2.15** – mode Verification

### 2.5.2 Identification

In identification mode, the system looks for matches in all user templates stored in the database to identify an individual. Thus, without requiring the subject to assert their identify, the system does a one-to-many comparison to determine an individual's identity (or fails if the subject is not enrolled in the system database) (e.g., "Whose biometric data is this?").

In negative recognition applications, identification plays an important part in determining if the subject is who she explicitly or implicitly claims not to be. Negative recognition aims to stop an individual from utilizing several identities [13] .

Positive recognition can also utilize identification for convenience, but it does not require the user to disclose their identity. On the other hand, traditional recognition methods such as PINs and cards may function correctly for positive recognition, but biometrics can only determine negative recognition (Figure 2.16).

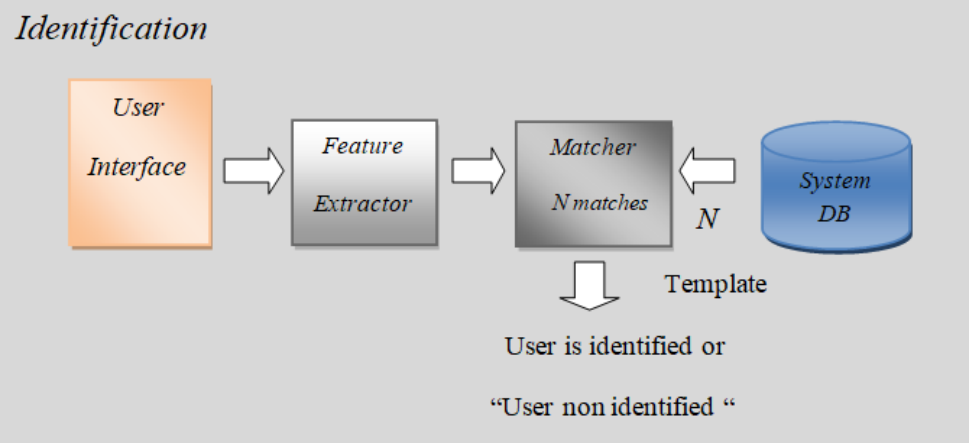



**Figure 2.16** – mode Identification

## 2.6 Modules of Biometric Systems

There are four primary elements that make up a standard biometric system:

- **Capture module:** A device, such as a camera, fingerprint reader, or security camera, is responsible for collecting a person's biometric information [14].
- **Features extraction module:** creates a new representation of the data by taking the biometric data that was collected by the capture module and extracting just the pertinent information. This new representation should ideally be specific to each individual and mostly unaffected by differences within the same class [14].
- **Matching Module:** It assesses the degree of similarity (or divergence) between the extracted characteristics and the model recorded in the system database [14].
- **Decision Module:** The system verifies the claimed identity of a user or identifies a person based on the degree of similarity between the extracted characteristics and the stored model [14].

## 2.7 Evaluation of biometric systems

We must clearly establish the following primary criteria in order to evaluate the effectiveness of a biometric system:

### 2.7.1 FRR (False Rejection Rate)

The FRR is a measure of how likely the system is to reject a legitimate request from a "Genuine" user. A Genuine user may try the transaction one or more times [15]. One

way to calculate FRR is:

$$FRR = \frac{\text{Total number of False Rejections}}{\text{Total number of 'Genuine' Attempts}} \qquad (1.1)$$

### 2.7.2 FAR (False Acceptance Rate)

The FAR measures the probability that the system may mistakenly provide access to an "Imposter." An Imposter may make one or more attempts at the transaction [15]. Performing the FAR calculation entails:

$$FAR = \frac{\text{Total number of False Acceptances}}{\text{Total number of 'Imposter' Attempts}} \qquad (1.2)$$

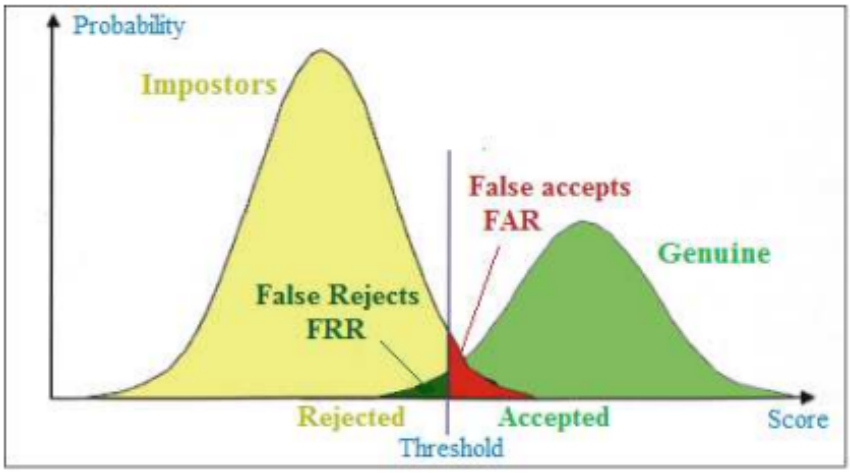



**Figure 2.17** – FAR and FRR diagrams

Figure 2.17 shows FAR and FRR diagrams according to distributions of genuine and imposter scores, While The EER is represented in Figure 2.17.

### 2.7.3 GAR (Genuine Acceptance Rate)

The GAR is an overall accuracy measurement of a biometric system. It represents the proportion of genuine attempts that are correctly identified and accepted by the system [16].It is calculated by the equation :

$$\text{GAR} = 1 - \text{FRR} \qquad (1.3)$$

### 2.7.4 ERR (Equal Error Rate)

Equal Error Rate The system obtains this quantitative measure when the False Acceptance Rate (FAR) and False Rejection Rate (FRR) distribution curves intersect at a specific point. Here, FAR corresponds to imposter cases accepted as genuine, while FRR shows the distribution of genuine attempts rejected by the system. This metric fulfills the criteria for measuring accuracy. Both protected and unprotected systems must use the same protocol and database when performing the equal error rate measurement [17].

### 2.7.5 Receiver Operating Characteristics Curve (ROC)

The ROC curve is an additional statistic used to assess accuracy. The obtained FAR is plotted against the verification rate (1-FRR). Since there is a trade-off between FAR and FRR, the ROC curve is essential in determining the biometric system's acceptable operating point where this trade-off may be handled. Every biometric system generally aims to produce a curve close to the upper left corner. An appropriate biometric system is represented by the opportunistic point that can be found on the curve in the top left corner[17](Figure 2.18).

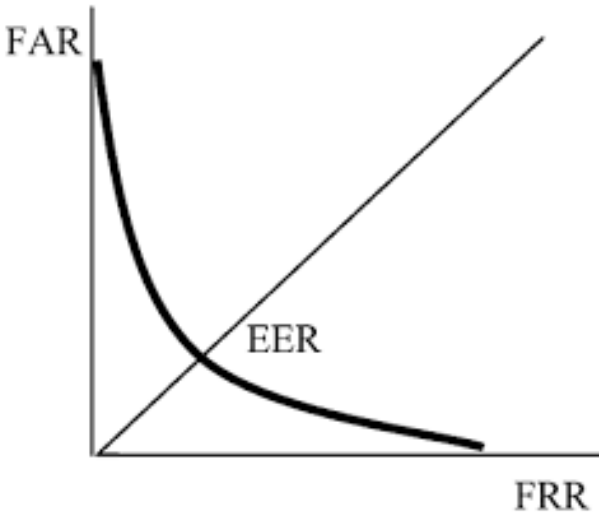



**Figure 2.18** – ROC curve

### 2.7.6 Cumulative Matching characteristic Curve (CMC)

This curve aids in the assessment of system throughput. This curve illustrates the identification rate of the biometric system when it compares a probe to each enrolled user. The system uses the matching score to rank the registered user in this case. Therefore, for inquiries whose enrollment is among the top r matches, the CMC curve gives the rank r identification rate. The CMC curve is then generated by plotting the acquired rank against the recognition rate [17] (Figure 3.1).

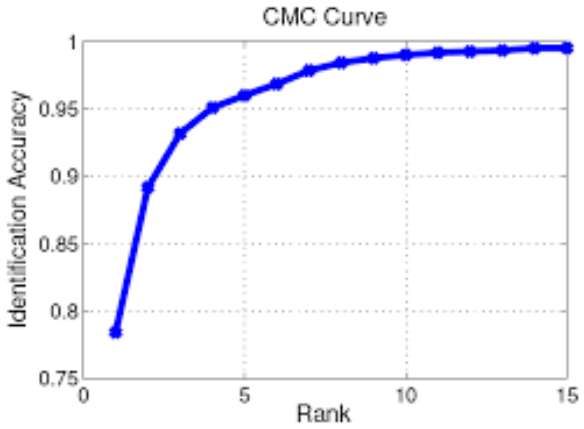



**Figure 2.19** – CMC curve

## 2.8 Conculsion

This chapter served as a foundational exploration into the realm of biometric systems. We delved into the fundamental concepts underlying these systems, examining their architecture and diverse applications across various domains. We realized that numerous factors intricately link the performance of biometric systems, leading to variations in efficacy among different systems. The following chapter will explore the use of Convolutional Neural Networks (CNNs) in biometrics, focusing on their potential applications in feature extraction and pattern recognition, and their impact on future biometric recognition systems.

Chapter 3

# Deep learning for biometrics

## 3.1 Introduction to AI

Determining the definition of artificial intelligence is a challenging task, with numerous definitions emerging, especially in recent years when the topic has garnered significant attention. The definition provided in the European Commission's communication dated April 25, 2018, which states that the term "refers to systems that display intelligent behavior by analysing their environment and taking actions—with some degree of autonomy—to achieve specific goals," is among the most recent and sufficient to mention [18].

Machine learning is a subfield of artificial intelligence (AI) that enables applications to anticipate outcomes with increasing accuracy without requiring explicit programming. In order to anticipate new output values, machine learning (ML) algorithms require the utilization of input data as features. Machine learning employs several techniques such as decision trees, support vector machine learning (SVM), and artificial neural networks (ANN), among others [19].

Deep learning is one area of machine learning. It is more widely used than ML algorithms and has grown in popularity in many areas. Convolutional neural networks (CNNs) are the most widely used algorithms in deep learning [20]. The CNN automatically extracts features from a set of raw image data based on the succession of convolutional layers, eliminating the need for human feature extraction for things like first- and second-

order statistical features, LBP, LPQ, etc. DL models come in a variety of forms, including AlexNet, GoogleNet, RestNet18, ResNet50, etc....[21].

In this chapter, we will explore more about deep learning, show the difference between machine learning and deep learning, and explain how deep learning works.

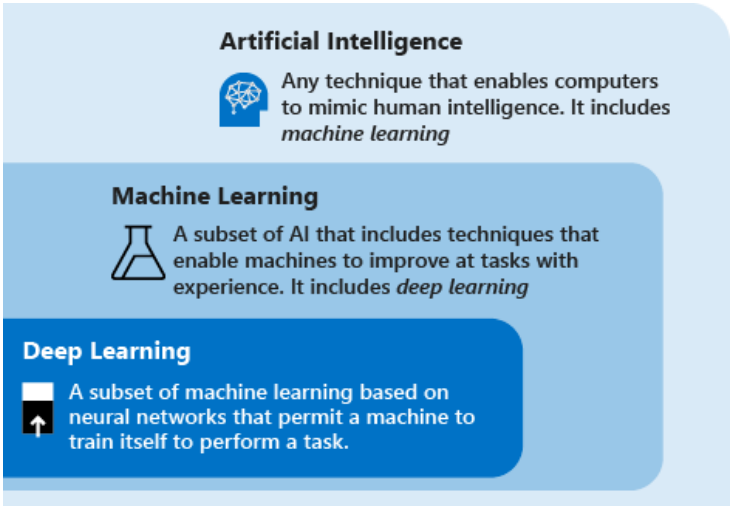



**Figure 3.1** – ai-vs-machine-learning-vs-deep-learning

## 3.2 Deep learning

Deep learning is a machine learning approach that uses automated algorithms to extract data insights to support decision-making. Deep learning techniques gradually extract higher-level information from unprocessed input using successive layers [22].

What sets deep learning apart is its progressive nature. It prioritizes the acquisition of successive layers of progressively more meaningful representations from the provided data. A deep learning model learns from the data at each layer, transferring knowledge from one layer to the next. This process is akin to how lower layers in image processing could recognize edges, while higher levels might recognize ideas that are important to humans, like faces, characters, or numbers [22].

Deep neural networks are large artificial neural networks composed of several hidden layers between the input and output layers [23](Figure 3.2).

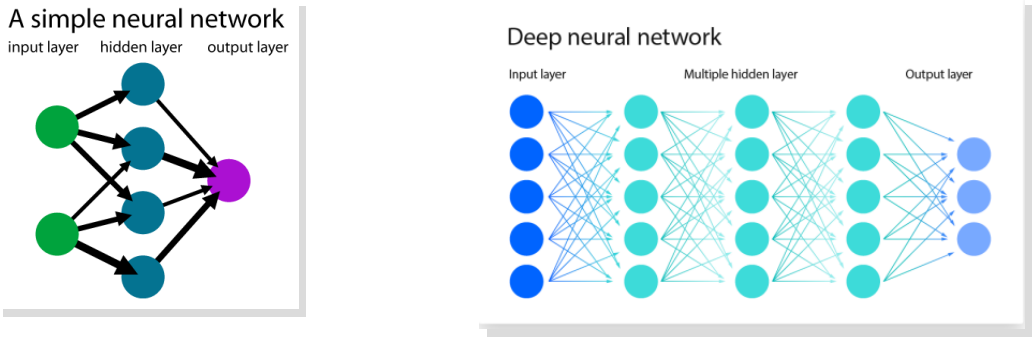



**Figure 3.2** – Dieffrence between a simple neural network and deep neural network

### 3.2.1 How deep learning work?

Deep learning methods are based on artificial neural networks, which mimic the human brain's hierarchical structure. During training, these networks automatically learn numerous hidden layers stacked on each other. The hierarchical structure of deep networks allows for nonlinear data processing, with each layer learning more complex concepts. The first layers detect low-level features as edges, while the deeper layers identify more complex features by combining features from the preceding layer. This hierarchical learning eliminates the need for hand-crafted feature extraction in the network [24].


**Figure 3.3** – Deep learning method

Figure 3.3 As one delves further into the network, a deep neural network of stacking layers learns increasingly intricate and unique traits. The lowest-level layers identify the edges of an input picture, and the next layers identify the curves and corners.

Combining the features derived from the earlier levels can lead to more complicated characteristics. The input is categorized by the output layer [24].

### 3.2.2 Deep learning vs machine learning

Even though deep learning is a branch of machine learning, it differs from traditional methods because deep learning algorithms automatically extract features and make predictions based on them. In contrast, traditional machine-learning methods require human

intervention [25].

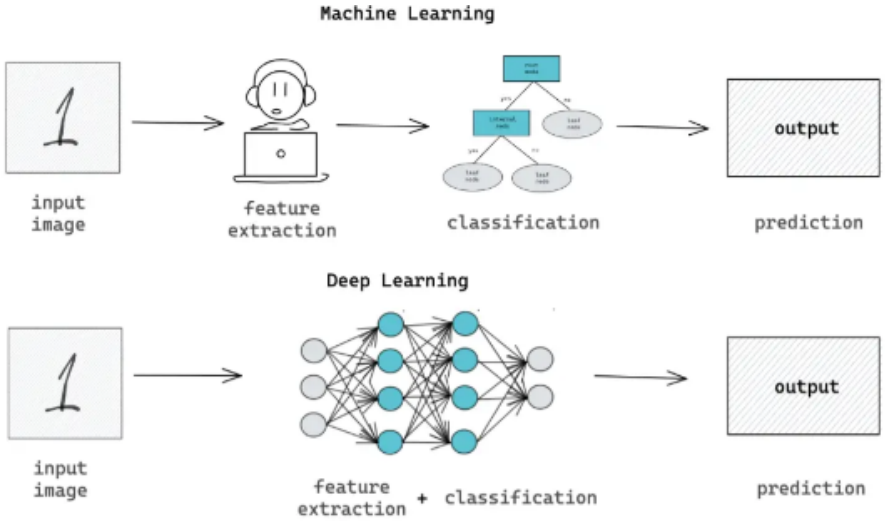



**Figure 3.4** – Deep learning methode

Another important difference is that deep learning algorithms perform better as data sizes grow, but typical machine learning approaches experience a performance plateau.

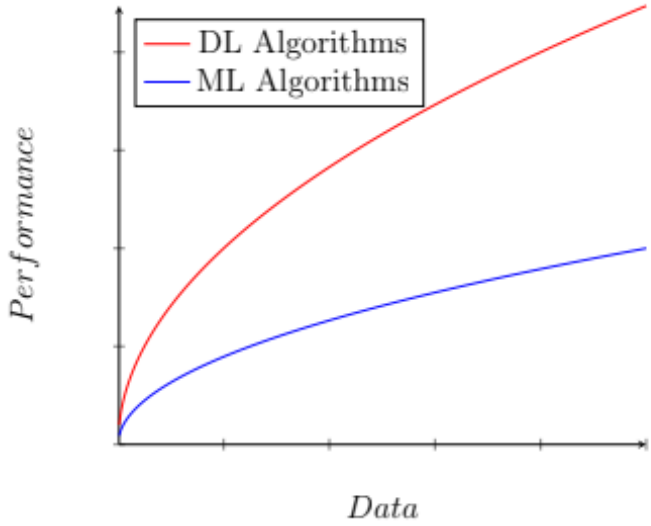



**Figure 3.5** – ML vs DL algorithms performance [24]

According to Figure 3.5, machine learning performance reaches a certain certification level and then plateaus once a certain amount of data is obtained, whereas deep learning performance rises with more data [25].

## 3.3 Neural networks basics

The development of ANN was based on the similar functioning of human neuronal networks [26]. Artificial neuron networks, which represent an advanced data modeling tool that can collect and disclose complicated input/output correlations, have provided several benefits in replicating many recent intelligent systems. Each node in an artificial neural network performs a basic calculation, and each connection transfers a signal from one node to another. Each connection is identified by a value known as the "connection strength" or weight, which indicates how much the signal is amplified or lessened by the connection [27].

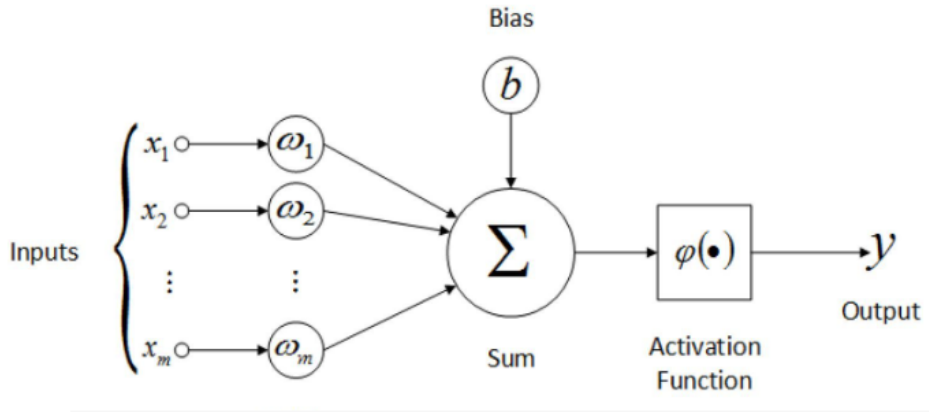



**Figure 3.6** – Structure of ANN

Figure 3.6 shows a network topology in which neuron j connects to inputs $(x_1, x_2, ...x_i)$ with weights $(w_{j1}, w_{j2}, ...w_i)$ on each connection. The neuron sums all the signals it receives, with each signal multiplied by its associated weights on the connection.

The final output,(y), is obtained by passing this output (V) through a transfer (activation) function, $f(v)$, which is often non-linear [24].

$$y = f\left(\sum_{i=1}^{n} x_i w_i + \text{bias}\right) \tag{2.1}$$

## 3.4 Deep neural network structure

There have been two major trends in deep learning approaches: supervised and unsupervised.

- **Supervised learning:** In supervised learning, input data is provided to the model

along with the output (e.g., Convolutional Neural Networks [28] and Recurrent Neural Networks [29]).

- **Unsupervised learning:** In unsupervised learning, the model receives only input data. (e.g., Deep Belief Network [30] and Deep (Sparse/Denoising) AutoEncoder [31]).

In our work, we use convolutional neural networks, which we will discuss in more detail in the section that follows.

## 3.5 Convolutional neural networks

Convolutional neural networks, sometimes called ConvNet architectures, are feed-forward neural networks typically used to process and interpret visual pictures. A CNN network detects and classifies objects in an image [32].

A Convolutional Neural Network is a deep learning algorithm model that accepts an image as input and uses it to give weights and biases—which may be learned—to distinct parts in the picture so that they can be distinguished from one another [33]. A ConvNet requires much less pre-processing than other classification techniques. One important feature of ConvNets is their ability to learn many visual features using various filter modifications, including max-pooling and convolutional layers.

CNNs are characterized by their convolution and pooling layers. In order to decrease the number of parameters and improve the sharing of common traits, these layers introduce partial correlations [22].

### 3.5.1 CNN blocs

Convolutive neuron networks are composed of two main blocks:

#### 3.5.1.1 Feature Extraction Bloc

The feature extraction block in a convolutional neural network (CNN) is crucial for extracting distinctive features from input data, especially images. It uses convolutional layers, pooling layers, and activation functions like ReLU to process the data efficiently. Convolutional layers detect patterns and features in the input data while pooling layers re-

duce computational complexity while preserving important spatial information. Common pooling operations include max pooling and average pooling, consolidating information from neighboring regions in feature maps [34].

#### 3.5.1.2 Classification Bloc

After feature extraction, convolutional neural networks (ConvNets) use fully connected layers to learn extracted features. These layers connect each neuron to the next, allowing for comprehensive learning. These layers use the detected features to make decisions about input data, predicting the strength of a particular input matching a specific class. A classifier receives the output of the last layer and generates class scores or probabilities to identify the most suitable class for the input data [35].

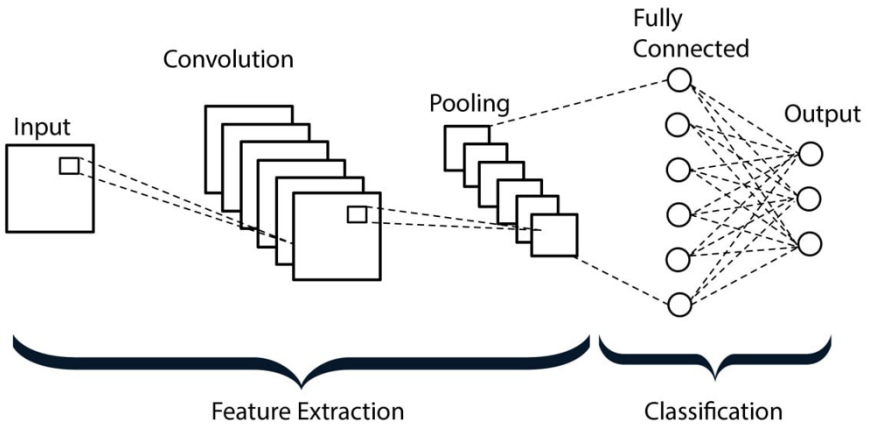



**Figure 3.7** – Structure of CNN

### 3.5.2 Layers in a CNN Network

A typical convolutional neural network consists of the following layers:

#### 3.5.2.1 Convolutional layer

The convolutional layer (CONV) is the core of a CNN used to extract features from the input image. It consists of learnable filters or kernels, each with a width and height, applied through the full depth of the volume. The convolution is performed between the input image and K kernels of size $M \times M$, each sliding across the input image and convolving with it. The result is a 2D output called a feature map, which represents

specific characteristics of the input. These feature maps are then fed to the next layer in the network, serving as feature extractors [24].

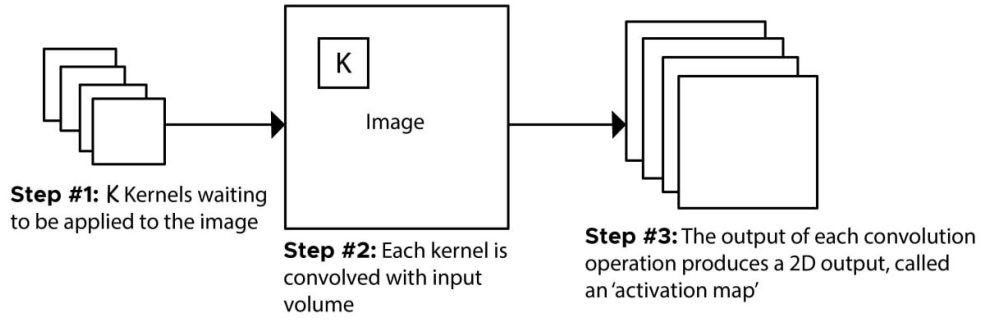



**Figure 3.8** – Convolutional Neural Network (CNN) Kernel Operation Steps

#### 3.5.2.2 Activation layer

An activation function layer receives the output from every convolutional layer. An activation function makes up the activation function layer. It takes the convolutional layer's feature map as input and outputs the activation map. The activation function produces an output signal from a neuron's activation level. It describes a neuron's output about a certain input. Usually, an activation function produces a squashing effect after receiving an input (a number), processing it mathematically, and then returning the activation level of a neuron within a specified range, such as $0 \leq x \leq 1$ or $-1 \leq x \leq 1$ [35]

#### 3.5.2.3 Pooling layer

ConvNets use a convolution and activation function layer followed by an optional pooling or down-sampling layer to reduce the input size and parameter count. The pooling layer summarizes a region of neurons in the convolution layer, with the most common technique being max-pooling. This technique outputs the maximum value in the input region, typically $2 \times 2$. The pooling layer discards less significant data but preserves detected features in a smaller representation. The reasoning behind the pooling operation is that feature detection is more important than feature location [35].

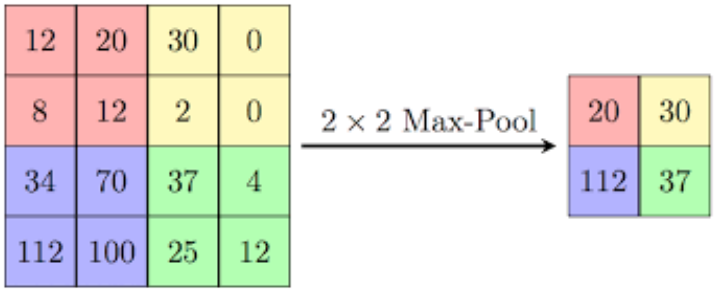



**Figure 3.9** – Example of max pooling

This strategy works well for simple problems but has its limitations and may not be suitable for some problems.

#### 3.5.2.4 Fully connected layers

The final few levels of the network, known as fully connected layers (FC), are often positioned before the output layer. They comprise neurons fully coupled to every activation in the layer above, weight, and biases. The output from the preceding layer is flattened and supplied to the FC layers. They are employed in classifying images into various classifications. A softmax classifier that calculates the probability of each class comes after [24].

### 3.5.3 Activation function

The output of a neural network is defined by mathematical models called activation functions. They determine whether or not to activate the neuron. They are non-linear transformations that are used to shorten computation times, normalize the output between [0,1] and [1,-1], and stop the network from converging. Different activation types exist. The most popular ones are the rectified linear unit, hyperbolic tangent, sigmoid function, and step function [24].

#### 3.5.3.1 Step function

This is the most basic activation function. simply outputs binary values $0 or 1$ based on threshold(Figure 3.10).often used in binary classification tasks [24] defined by :

$$f(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases} \tag{2.2}$$

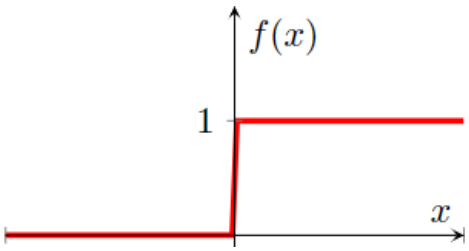



**Figure 3.10** – Step function

#### 3.5.3.2 Sigmoid function

The sigmoid function is mathematically represented as :

$$f(x) = \frac{1}{1+e^{-x}} \tag{2.3}$$

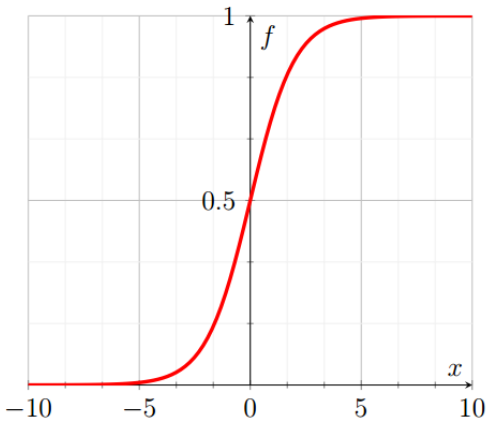



**Figure 3.11** – sigmoid activation function

It is an S-shaped curve as shown in (Figure 3.11). The sigmoid function squashes the input into the range [0, 1] [35].

#### 3.5.3.3 Hyperbolic tangent

Similar to the sigmoid function, the hyperbolic tangent function's output ranges between -1 and 1. Tanh over sigmoid has the advantage of mapping negative inputs strongly

negative and zero inputs near zero in the tanh graph [35], as shown in (Figure 3.12)

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2.4}$$

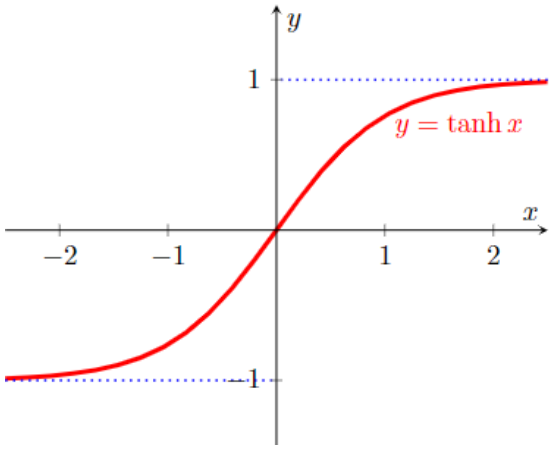



**Figure 3.12** – Hyperbolic tangent

#### 3.5.3.4 ReLU Activation Function

ReLu is a piecewise linear function that returns zero if the input value is negative and directly if the input value is positive (Figure 3.13). Because it makes training easier and frequently performs better than other activation functions, it is the default activation function for a lot of networks [22]. It is mathematically given as :

$$f(x) = \max(0, x) \tag{2.5}$$

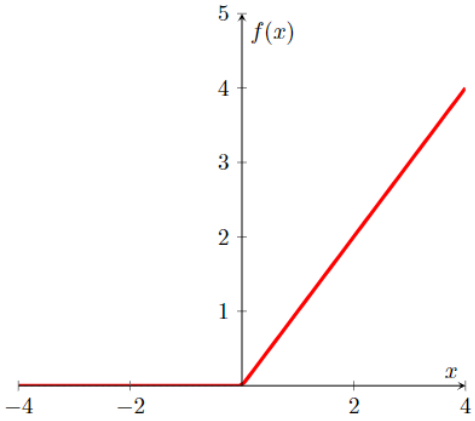



**Figure 3.13** – Rectified Linear Unit (ReLU) activation function

## 3.6 Transfer learning

Deep CNNs require a lot of compute power because of their depth and quantity of fully linked nodes, so training them on big datasets from the start might take days or weeks. Starting from the model weights of pre-trained networks on benchmark datasets, the process of picture classification can be sped up by minimizing these enormous time and computation resources. These models can be applied straight to additional classification problems, or they can be integrated into a new model. We refer to this as transfer learning. As a result, the most widely applied deep learning technique is transfer learning, which involves using a model that has been trained on one job as an initialization for a different task [24].

Transfer learning has become more and more common, particularly with convolutional neural networks. It efficiently shortens training times and improves the precision of models created for jobs with little or no training data. Some uses for transfer learning are as follows:

- **Pretrained model as fixed feature extractor**

In this case, a new linear classifier is added as the final fully connected layer (classifier layer), and it is then trained using fresh datasets. In this way, just the classifier is refined while the feature extraction layers stay constant. This approach works well in situations when the new dataset is small but comparable to the old dataset [35].

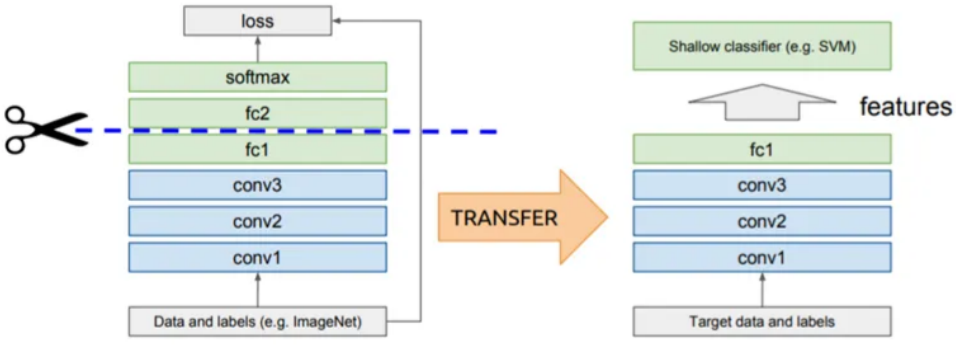



**Figure 3.14** – Fine Tuning

- **Fine-tune whole Model**

with a pre-trained model, create a new fully connected layer to replace the classifier layer, and then retrain the whole network with a fresh dataset by extending backpropa-

gation to the upper layers. This adjusts all the weights for each new task [35].

## 3.7 CNN architectures

Convolutional neural networks (CNNs) are a domain with many different architectures. The most typical ones are:

### 3.7.1 AlexNet

Alex Krizhevsky and associates developed Alexnet in 2012. Though it has eight layers and learnable parameters, its architecture is similar to that of LeNet but deeper. The network receives RBG photos as input. This model has three fully connected layers with a Softmax classifier, five convolution layers, and max-pooling layers. They use ReLU as an activation function. Two dropout layers are also present in the network [28].

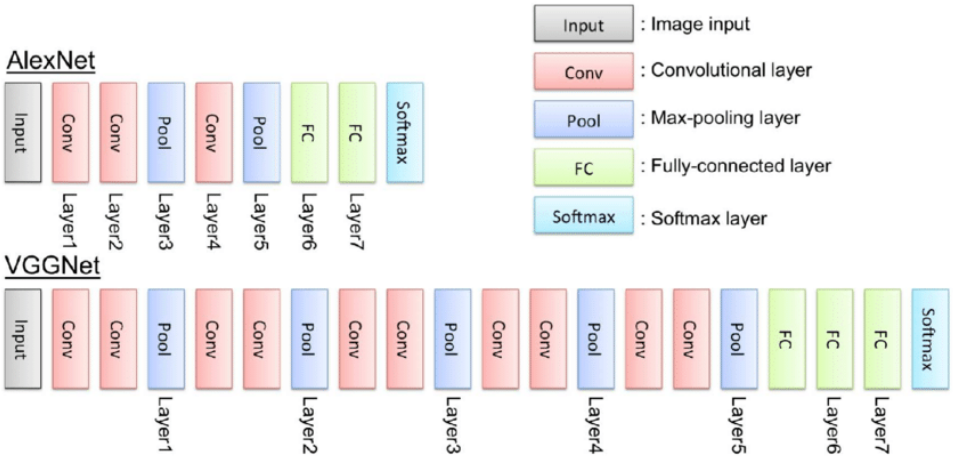



**Figure 3.15** – Alexnet architecture

### 3.7.2 VGGNet

VGGNet, based on the principles of AlexNet, aims to create deep configurations (16 to 19 layers) using structural stabilization techniques to control parameters and mitigate overfitting risks. By reducing filter sizes from $7 \times 7$ and $5 \times 5$ to $3 \times 3$, VGGNet enables the addition of more intermediate layers without exponentially increasing parameters. Comparing parameters in three stacked convolutional layers with $3 \times 3$ filters to one layer with a $7 \times 7$ filter shows that smaller filters reduce parameters [26].

For each convolutional layer with depth $C$, the parameter count in three stacked $3 \times 3$ convolutional layers is $3(3^2C^2)$, while for one layer with $7 \times 7$ filters, it's $7^2C^2$.

Overall, stacking convolutional layers with smaller receptive fields reduces parameters and enhances network non-linearity through additional activation functions (ReLU) [26].

### 3.7.3 ResNet

Xiangyu Zhang, Jian Sun, Kaiming He, and Shaoqing Ren created ResNet, or Residual Network, in 2015. This architecture was the winner of the ILSVRC competition. ResNet uses batch normalization and skip connections as its foundation. Its architecture is influenced by the VGG-19, although there are no FC layers at the network's end. An average pooling layer takes its place. There is just one max pooling layer in all ResNet models after the initial one. It There are various ResNet topologies, ranging in number of layers from 18 to 152 [36].

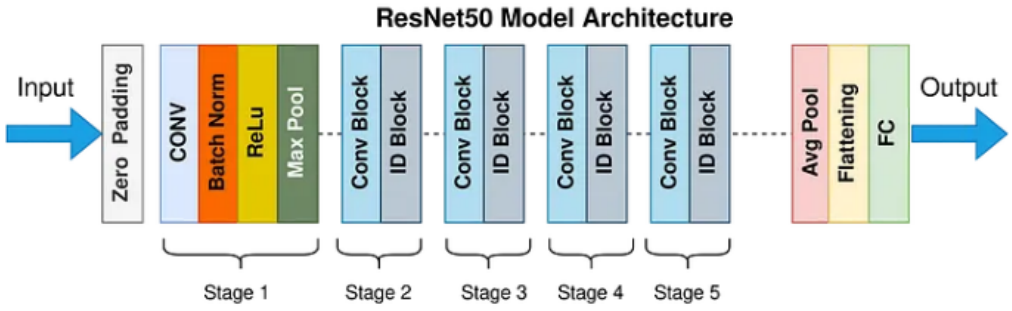



**Figure 3.16** – RESNet architecture

### 3.7.4 DenseNet

DenseNet201 is a convolutional neural network architecture developed by Huang et al. in 2017. It features a densely connected structure, with each layer receiving inputs from preceding layers and passing feature maps to subsequent layers. This design improves information flow and reduces the vanishing gradient problem.

DenseNet201 is parameter-efficient, achieving performance on par with deeper networks like ResNet-101 but with fewer parameters. Transition layers manage complexity, ensuring improved gradient flow, reduced overfitting, and enhanced model generalization[37].

## 3.8 Support Vector Machines (SVM)

Pattern recognition problems are addressed by creating Support Vector Machines (SVM). By employing support vectors—specific points from the training dataset—to cre-

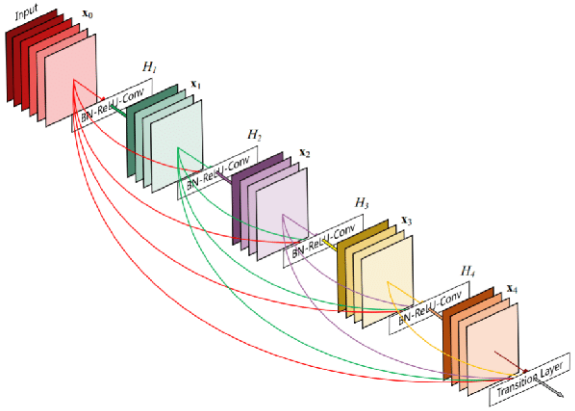



**Figure 3.17** – Architecture of DenseNet201

ate a decision boundary, this technique recognizes patterns for binary classification. The decision boundary is optimized to maximize the margin between the two classes. The ideal separation hyperplane is the decision boundary, which symbolizes this margin. As seen (Figure 3.18), the points that are closest to this hyperplane are known as support vectors [38].

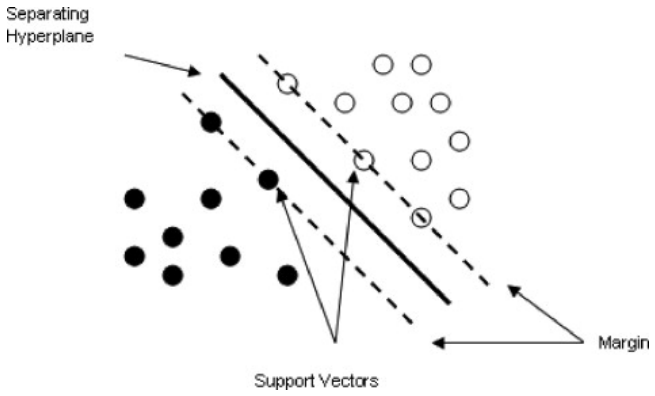



**Figure 3.18** – Separating hyperplane of the SVM

## 3.9 Conclusion

In this chapter, we have presented the various principles of deep learning. Deep learning networks are machine learning algorithms based on neural networks. We then explored some fundamental principles by detailing the structure of a Convolutional Neural Network (CNN), including the layers of a CNN, how they function, and how to progress towards the transfer learning process, the method we used in our study. Finally, we analyzed some well-known CNN architectures. In the next chapter, we will discuss the experimental results.

Chapter 4

# Results and discussions

## 4.1 Introduction

This chapter represents the experimental results of unimodal and multimodal biometric identity identification systems that employ finger-knuckle prints (FKPs) patterns as biometric information. We used four pre-trained CNNs (AlexNet, VGG19, ResNet50 and Densenet201) to extract and classify the features from the images in the data set. The trials compared each CNN's performance on the index, middle, and ring finger images, evaluated improvement through score-level fusion and examined overall systems performance. The objective is to create reliable and efficient biometric identification systems that can handle the unpredictability and complexity of biometric data.

## 4.2 System Overview and Block Diagram

Figure 4.1 shows a simplified block diagram of a biometric system designed to authenticate individuals based on the unique patterns of their finger knuckle prints (FKPs). In this system, three fingers—index, middle, and ring—are utilized to enhance the reliability and accuracy of the biometric identification process.

"The system begins with the Pre-processing Stage, where Region of Interest (ROI) extraction is performed to ensure the most relevant data is obtained. This process is carried out by the app while the photos are being taken, ensuring that the critical areas are focused on. Deep learning techniques are used in the Feature-Extraction Stage to extract meaningful features from ROIs using pre-trained networks like AlexNet, VGGNet19,

ResNet50, and Densenet201. The classification stage uses extracted characteristics to distinguish different people using classifiers like Support Vector Machines (SVM) and a fully connected layer. These classifiers produce scores indicating the likelihood of FKPs belonging to specific people. The final choice stage involves fusing the classifiers' results at the score level, using the advantages of multiple classifiers and fingers to improve system accuracy. The system completes the biometric authentication process by accepting or rejecting users based on their scores.

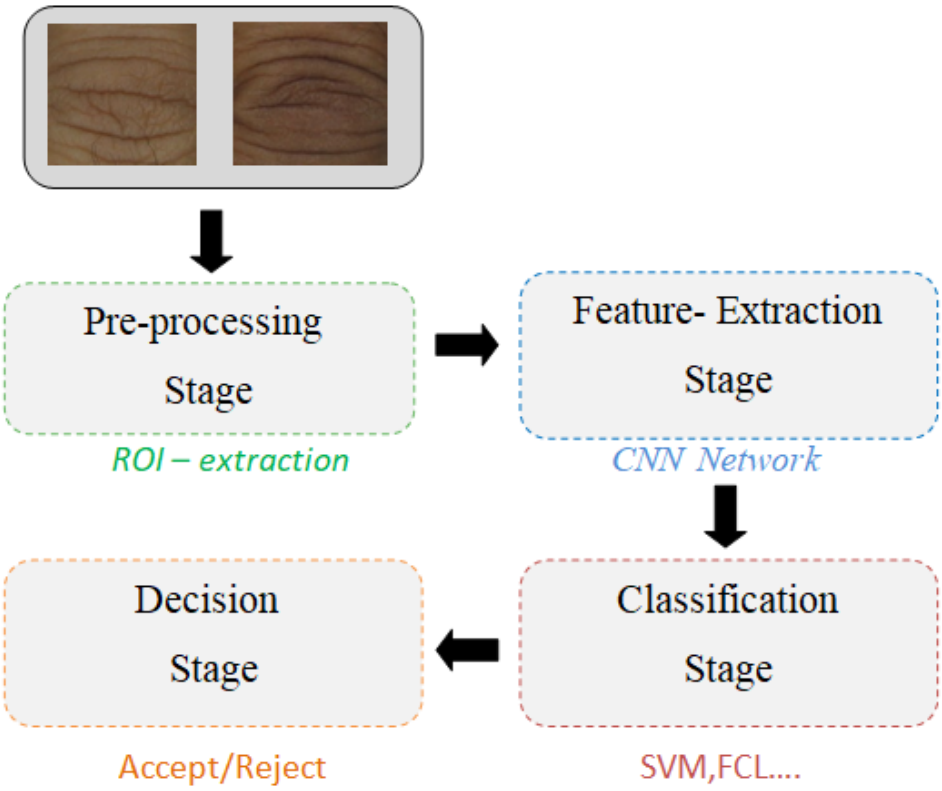



**Figure 4.1** – Biometric system structure

## 4.3 Datasets used

We meticulously evaluated the system under consideration using the FKP and palmprint datasets, which include images of finger knuckle prints (FKP) and palmprints. We painstakingly generated these datasets using a custom-developed Android application, which allowed us to capture FKP and palmprint images at specified dimensions. Forty-four individuals contributed to the database, and we acquired images in two sessions spaced 15 to 30 days apart.

Each participant's process involved FKP images: precisely six images of each of the index, middle, and ring fingers of both hands (right and left) in each session. We also collected the hand-palm dataset, using 96 images for each individual. Although we did not use the palmprint data in our study due to the time-consuming editing process, we collected it with the same meticulousness.

The FKP dataset contains images from 44 individuals. We effectively generated 88 files by looking at each finger (index, middle, and ring) on both right and left hands and grouping them. For each finger, we used six images (1st, 3rd, 5th, 7th, 9th, and 11th) for training purposes and the remaining six for testing.

During the data collection process, we encountered challenges, particularly in securing the participation of all volunteers in the second session. However, despite these hurdles, it was our first attempt at such a study, and we remained dedicated, achieving satisfactory and acceptable results. We take pride in our unwavering commitment to this research, even in adversity.

**(a)** 1 **(b)** 2

**Figure 4.2** – Example des Images FKP

## 4.4 Separation of Databases

The finger knuckle images used in our study consist of standard FKP (Finger Knuckle Print) images. These images capture the natural patterns and features present on the finger knuckles.

For data separation, we employ the following strategy:

- **Training Images**: The training set consists of images captured during odd sessions. Specifically, images from sessions 1, 3, 5, 7, 9, and 11 are utilized for training.
- **Testing Images**: The testing set comprises images captured during even sessions. This includes images from sessions 2, 4, 6, 8, 10, and 12.

Using this method, we ensure that the system's performance is systematically evaluated over many sessions, making it easier to evaluate its accuracy and resilience in identifying finger knuckle patterns.

## 4.5 Work environment

- **Hardware environment :**

- PC: DESKTOP-ITHQKBS
- Memory (RAM): 8.00 GB.
- Processor: Intel(R) Core(TM) i7-7700 CPU @ 2.80GHz 2.81 GHz.
- System type: 64-bit operating system.

- **Software environment :**

We have employed Matlab R2024a as the logic tool in our method.

## 4.6 Experimental Results

### 4.6.1 Uni-modal biometric identification system

We determined the open set's threshold (T0) and equal error rate (EER). We calculated the closed set's Ranke of Perfect recognitio (RPR) and recognition rate (ROR). The performance of the systems based on the four modalities—Index, Middel, Ring—was evaluated after adapting each pre-trained network (AlexNet, VGG19, ResNet50, and DenseNet201) with a CNN model trained from scratch (CNN Scratch).Table 4.1 shows the results for the index dataset.

**Table 4.1** – Performance of the uni-modal identification system based on index finger

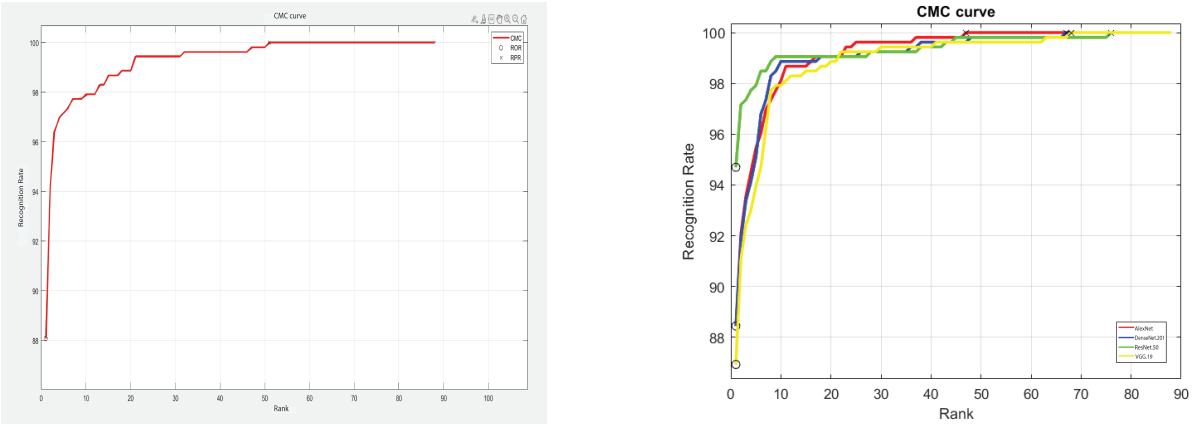
| Transfer Learning | EER | $T_0$ | ROR | RPR |
|---|---|---|---|---|
| AlexNet.50 | 3.97945 | 0.795 | 88.4470 | 86 |
| VGG.19 | 4.38001 | 0.814 | 86.9318 | 68 |
| ResNet.50 | 3.66162 | 0.837 | 86.9318 | 63 |
| DenseNet.201 | 3.40909 | 0.777 | 88.447 | 67 |
| CNN Scratch | 1.92006 | 0.051 | 88.0682 | 51 |

**ERR** CNN Scratch outperforms pre-trained models with the lowest EER at 1.92006, with DenseNet-201 having the lowest EER at 3.40909.

**$T_0$** The CNN Scratch model achieves a low threshold of 0.051, demonstrating precise decision boundaries, while pre-trained models have higher thresholds at ResNet-50, VGG19, AlexNet, and DenseNet-201.

**ROR** The CNN Scratch model performs well with ROR of 88.0682%, but DenseNet-201 and AlexNet achieving the highest ROR at 88.4470%.

**RPR** CNN Scratch outperforms all models with an RPR of 51, while ResNet-50 has the lowest RPR at 63, followed by DenseNet-201, VGG19, and AlexNet.




**(a)** 1 **(b)** 2

**Figure 4.3** – UNI-modal system performance based on index data set, (a)CMC curve for CNN scratch ,(b) CMC curves for transfer learning architectures

Based on the results and Figure 4.3, the CNN from scratch shows a superior error rate and recognition efficiency performance, with DenseNet-201 being the best performer among pre-trained models. However, it falls short compared to the custom-trained model. The analysis highlights the advantages of training a model from scratch for specific biometric tasks, provided sufficient data and computational resources are available.

**Table 4.2** – Performance of the uni-modal system based on middle modulate

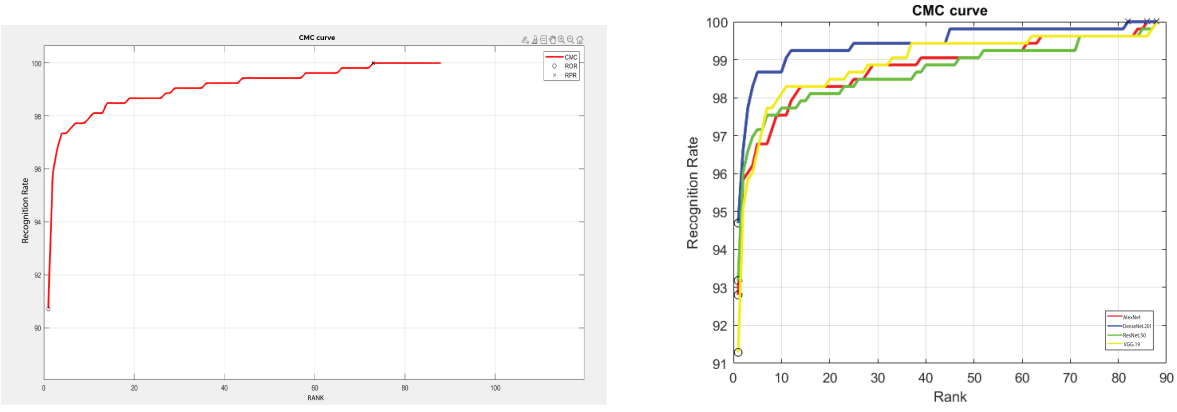
| Transfer Learning | EER | T₀ | ROR | RPR |
|---|---|---|---|---|
| AlexNet.50 | 3.60066 | 0.814 | 92.8030 | 86 |
| VGG.19 | 3.26106 | 0.775 | 91.2879 | 88 |
| ResNet.50 | 2.67328 | 0.776 | 93.1818 | 88 |
| DenseNet.201 | 1.37583 | 0.698 | 94.697 | 82 |
| CNN Scratch | 1.92006 | 0.042 | 90.7197 | 73 |

**ERR** The DenseNet-201 model has the best error rate performance at 1.37583, followed by the CNN Scratch model with an EER of 1.92006, and AlexNet-50 with the highest EER at 3.60066.

**T₀**The CNN Scratch model achieves a low threshold of 0.042, demonstrating precise decision boundaries, while DenseNet-201 has the lowest threshold at 0.698 among pre-trained models.

**ROR** DenseNet-201 outperforms other models with a high ROR of 94.697%, followed by ResNet-50, AlexNet-50, and Vgg-19, while CNN Scratch model achieves a ROR of 90.7197

**RPR** The CNN Scratch model has the highest recognition precision (RPR) at 73, followed by ResNet-50 and Vgg-19 at 88, DenseNet-201 at 82, and AlexNet-50 at 86.




**(a)** **(b)**

**Figure 4.4** – UNI-modal system performance based on middle data set, (a)CMC curve for CNN scratch ,(b) CMC curves for transfer learning architectures

Depending on the results and Figure 4.4, the CNN Scratch model shows low EER and precise decision boundaries, proving its effectiveness when trained from scratch. DenseNet-201 is the top-performing pre-trained model, excelling in EER and ROR. However, it achieved the lowest RPR compared to ResNet-50 and Vgg-19.

**Table 4.3** – Performance of the uni-modal system based on ring modulate

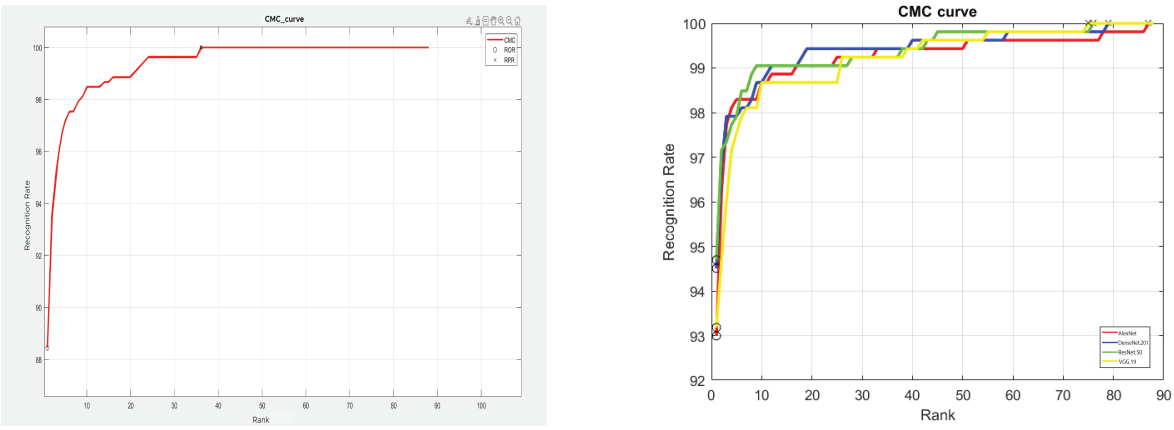
| Transfer Learning | EER | $T_0$ | ROR | RPR |
|---|---|---|---|---|
| AlexNet | 2.79084 | 0.783 | 92.9924 | 87 |
| Vgg.19 | 2.26402 | 0.76 | 93.1818 | 75 |
| ResNet.50 | 1.52821 | 0.732 | 94.697 | 76 |
| DenseNet.201 | 1.89829 | 0.692 | 94.5076 | 79 |
| CNN Scratch | 2.25531 | 0.042 | 88.4470 | 36 |

**ERR** The ResNet-50 model outperforms DenseNet-201 with an EER of 1.89829, while AlexNet-50 has the highest EER at 2.79084.

**$T_0$** The CNN Scratch model achieves a low threshold of 0.042, demonstrating precise decision boundaries, while DenseNet-201 has the lowest threshold at 0.692 among pre-trained models.

**ROR** ResNet-50 outperforms DenseNet-201, Vgg-19, and AlexNet-50 in recognizing instances with a high ROR of 94.697%, while CNN Scratch has a lower ROR of 88.4470%.

**RPR** The CNN Scratch model has the lowest recognition precision at 36, with AlexNet-50 having the highest RPR at 87.




**(a)** **(b)**

**Figure 4.5** – UNI-modal system performance based on ring data set,(a)CMC curve for CNN scratch ,(b) CMC curves for transfer learning architectures

Considering the analysis and Figure 4.5, the ResNet-50 model is the best overall, with the lowest EER and highest ROR, indicating superior error rate and recognition accuracy. The CNN Scratch model has the lowest threshold and best recognition precision, but its overall performance does not surpass ResNet-50. DenseNet-201 is a strong pre-trained option.

### 4.6.2 Multi-modal biometric identification system

This subsection aims to assess and enhance the unimodal biometric identification system's performance by utilizing data from several modalities produced from various finger kinds. Multiple multimodal systems can be constructed using multiple CNN networks for feature extraction and classification (AlexNet, ResNet50, VGG19, DenseNet201) and modalities (Index, Middle, and Ring fingers). However, we concentrated on a multi-sample system approach in our work.

In our methodology, score-level fusion was tested. Multiple combinations can be obtained using each person's modalities. In our study, we considered two cases:

**Case 01:** We input all the data (Index, Middle, and Ring fingers) simultaneously into each model (AlexNet, VGG19, ResNet50, and DenseNet201) for every evaluation. The performance results were obtained for each model using this comprehensive dataset.

**Case 02:** We separated the data by finger type, inputting the Index finger data into our custom CNN models and saving the results. The same procedure was repeated for the Middle and Ring finger data. After obtaining the results for each finger type, we performed a fusion of these three results using the SUM fusion rule to obtain our final results.

#### 4.6.2.1 Case 01 - Simultaneous Input of All Finger Data into CNN Models

**Table 4.4** – Performance of multimodal biometric identification systems (open set and closed set), identification test results for AlexNet, VGG19, Resnet50, and Densenet201 networks

| Transfer Learning | EER | $T_0$ | ROR | RPR |
|---|---|---|---|---|
| AlexNet | 0.940439 | 0.758 | 97.1591 | 82 |
| VGG.19 | 0.936085 | 0.722 | 97.5379 | 81 |
| ResNet.50 | 0.76193 | 0.703 | 97.3485 | 70 |
| DenseNet.201 | 0.570359 | 0.699 | 97.5379 | 29 |

The table presents the performance of various CNN models (AlexNet, VGG19, ResNet-50, and DenseNet-201) when simultaneously using data from Index, Middle, and Ring fingers. The metrics included are EER (Equal Error Rate), Threshold, ROR (Rank-One Recognition), and RPR ( Ranke of Perfect recognitio).

**AlexNet:** It has an EER of 0.940439, demonstrating a generally tall blunder rate. It requires the most elevated limit esteem of 0.758 among the models, recommending

that it needs a higher edge to separate accurately. Despite this, it accomplishes a ROR of 97.1591%, illustrating great rank-one acknowledgement execution. The RPR for AlexNet-50 is 82, reflecting lower exactness compared to DenseNet-201.

**VGG19:** It has an EER of 0.936085, slightly less than AlexNet—50 but still fairly high. $T_0$ value is 0.722, lower than AlexNet—50 but more advanced than ResNet—50 and DenseNet—201. VGG19 achieves a high ROR of 97.5379, meaning strong performance in identifying accurate existence on the first try. Its RPR is 81, analogous to AlexNet—50, indicating moderate perfection.

**ResNet-50:** It improves over both AlexNet-50 and VGG19 with an EER of 0.76193. $T_0$ esteem is 0.703, lower than AlexNet-50 and VGG19, proposing superior separation capability. ResNet-50 has an ROR of 97.3485%, slightly lower than VGG19 and DenseNet-201 but still very high. The RPR is 70, which is way better than VGG19 and AlexNet-50 but less exact than DenseNet-201.

**DenseNet-201:** It stands out with the minimum EER at 0.570359, indicating the best performance in minimizing errors. It has the smallest threshold value of 0.699, demonstrating the ability to identify true and false matches. DenseNet-201 matches VGG19 with a high ROR of 97.5379, establishing a strong rank-one recognition performance. It has an RPR of 29, significantly lesser than the others, indicating the highest perfection rate.

**Summary:**

DenseNet-201 is the best model for biometric identification because it has the lowest error rate, threshold, and highest accuracy. ResNet-50 does an excellent job with minor mistakes and is accurate. AlexNet and VGG19 are good at recognizing the most common objects but make more mistakes and have higher limits so that they could be better overall. Figure 4.6 The CMC graph shows that DenseNet-201 is the best, followed by ResNet-50, and then AlexNet and VGG19 are also good but not as good as the others.

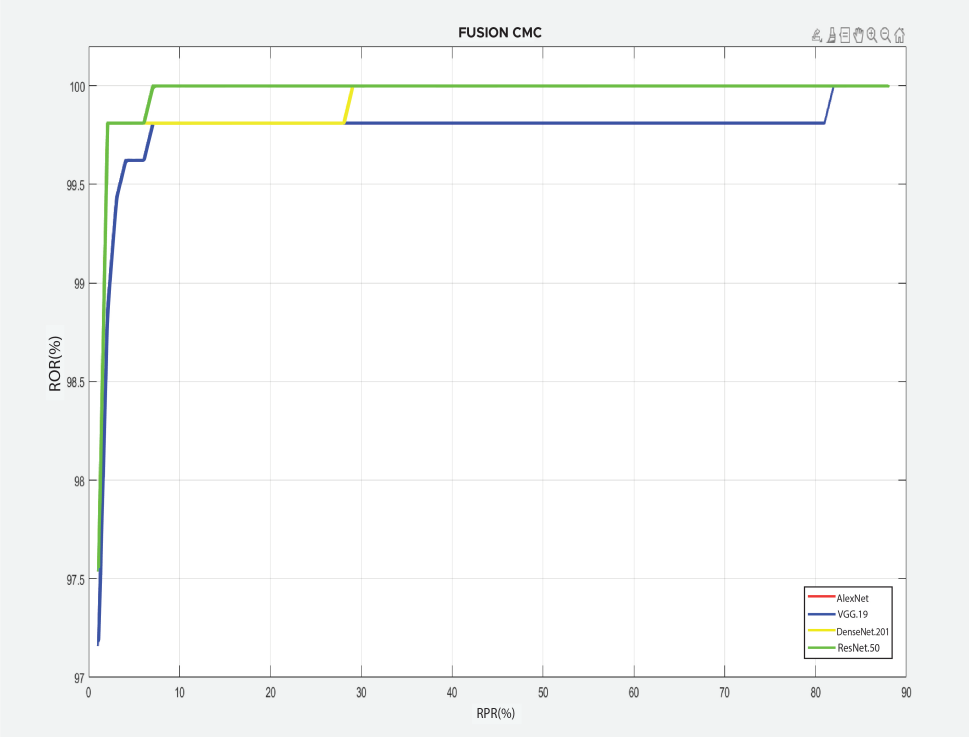



**Figure 4.6** – CMC curves for all CNN models (AlexNet, VGG19, ResNet-50, DenseNet-201)

#### 4.6.2.2 Separate Input of Finger Data with Post-Fusion

**Table 4.5** – Distinct Data Capture: Index, Middle, Ring Fingers Performance in Open Set and Closed Set for CNN Models.

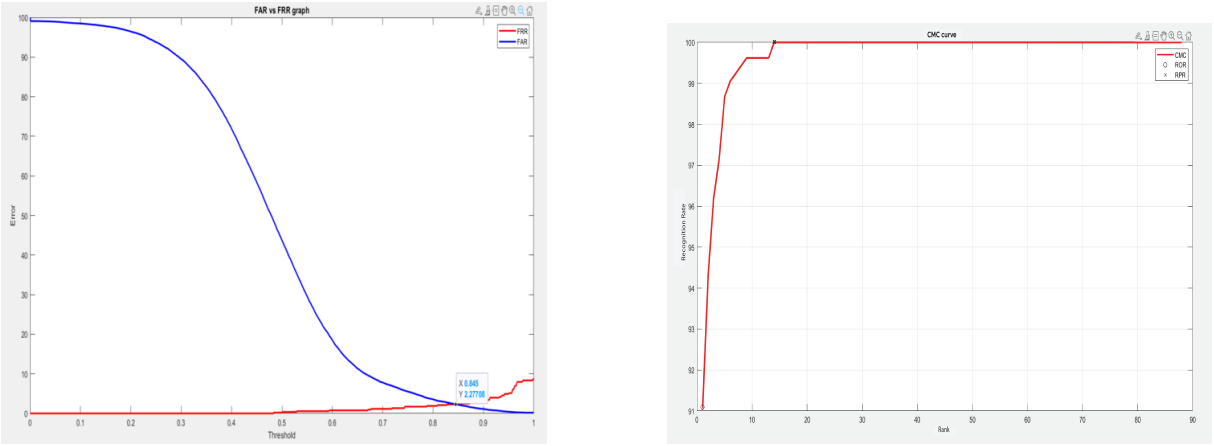
| Data Set | EER | $T_0$ | ROR | RPR |
|---|---|---|---|---|
| Index finger | 2.27708 | 0.845 | 91.0985 | 14 |
| Middle finger | 1.4542 | 0.76 | 91.0985 | 14 |
| Ring finger | 0.966562 | 0.752 | 91.0985 | 12 |

The table illustrates the individual performance metrics of custom CNN models for each finger type—Index, Middle, and Ring. For each finger, the data was processed through the CNN models(AlexNet, VGG19, ResNet-50, DenseNet-201) to evaluate the system's performance based on individual finger biometrics. The results were combined using the SUM fusion rule to enhance the identification system's accuracy and reliability.

- **Comparison between the three modalities**

Figure 4.7 shows the performance of the multi-modal system (open and closed set) using the index dataset based on the CNN models, as it is apparent that the index finger has the highest error In the open set identification mode. However, it gives better results than the index unimodal test. This method gives a $T_0 = 0.845$, significantly higher

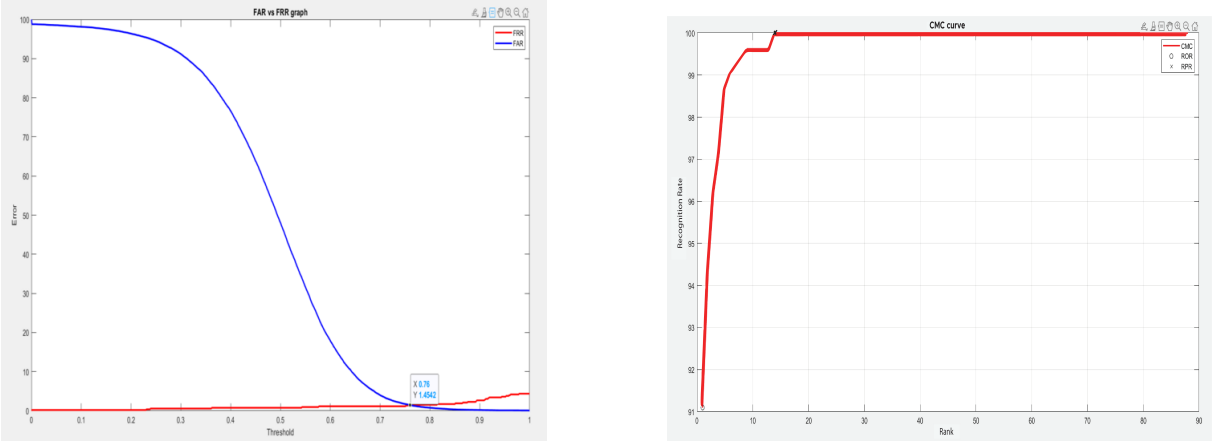
than the other data set modalities. All modalities perform more closely in closed set identification mode with an ROR = 91,09%




**(a)** **(b)**

**Figure 4.7** – Multi-modal biometric identification system performance (open/closed set) based on index data set,(a)accuracy of index data (b) CMC curve of middle data

Figure 4.8 shows the performance of the multimodal system (open and closed set) using the middle finger dataset based on the CNN models. As it is apparent, the middle finger has a lower error than the index finger in the open set identification mode, demonstrating better performance. This method provides a $T_0 = 0.76$, lower than the index finger dataset. All modalities perform similarly in the closed set identification mode, with an ROR = 91.0985% and an RPR = 14.

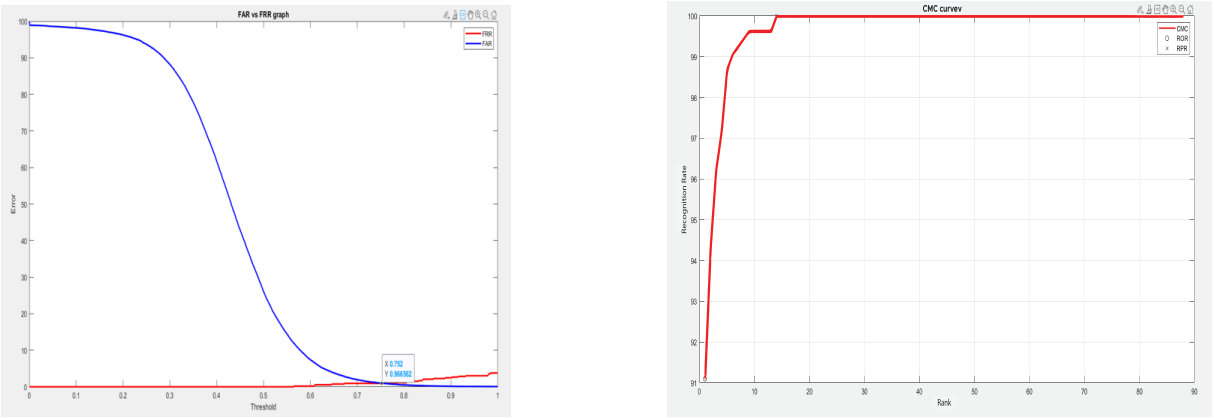



**(a)** **(b)**

**Figure 4.8** – Multi-modal biometric identification system performance (open/closed set) based on middle data set, (a)accuracy of middle data (b) CMC curve of middle data

Figure 4.9 shows the performance of the multimodal system (open and closed set) using the ring finger dataset based on the CNN models. The ring finger dataset yields

the lowest error in the open set identification mode among all three datasets, indicating the best performance. This method provides a $T_0 = 0.752$, slightly lower than the middle finger dataset. All modalities perform similarly in the closed set identification mode, with an ROR = 91.0985 % and an RPR = 12. The ring finger presents the best results among the three datasets, highlighting its effectiveness in biometric identification.




**(a)** **(b)**

**Figure 4.9** – Multi-modal biometric identification system performance (open/closed set) based on ring data set,(a)accuracy of ring data (b) CMC curve of ring data

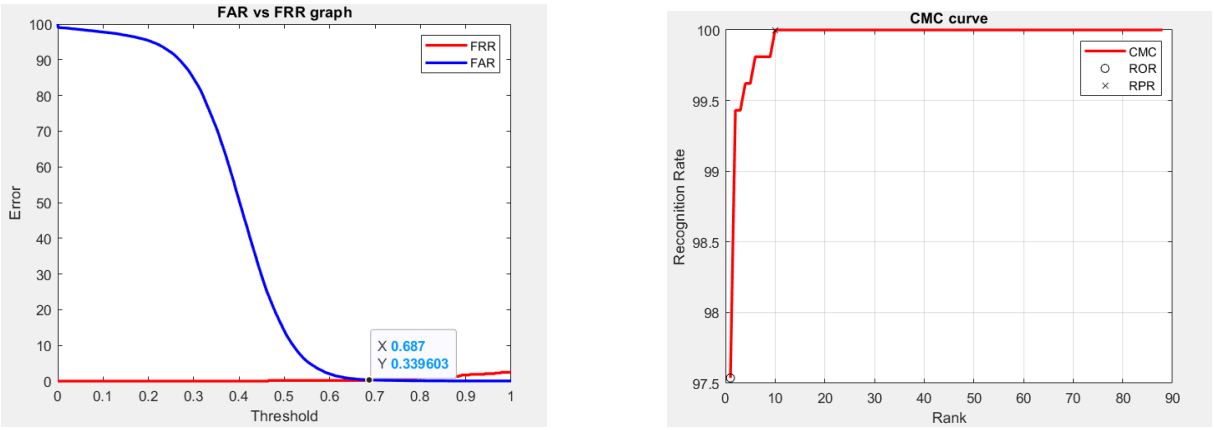
**Table 4.6** – Fusion results for the three modalities (index, middle, ring ) in open/closed sets

| test 3 | EER | THRESHOLD | ROR | RPR |
|---|---|---|---|---|
| **fusion of all dataset** | 0.339603 | 0.687 | 97.5379 | 10 |

The table shows the efficiency parameters of the multimodal biometric identification system at the end of Case 02 when the data from separate finger modalities (Index, Middle, and Ring) are fused.

Figure 4.10's results are really encouraging. The open set exhibited a notable decrease in error, as seen by the lowest Equal Error Rate (ERR) of 0.339603. This implies that integrating data from various finger modalities enhances the system's accuracy. Furthermore, the low threshold of 0.687 suggests that when the data is pooled, the system can distinguish between real and false matches more successfully.

The Rate of Recognition (ROR) in the closed set is 97.5379%, which aligns with the maximum ROR values noted for each of the individual CNN models in Case 01. Moreover, compared to earlier findings, the lowest Ranke of Perfect recognition (RPR) is 10. This suggests that there are fewer false positives with the fused technique.




**(a)** **(b)**

**Figure 4.10** – Multi-modal biometric identification system performance (open/closed set) based on fusion results, (a)accuracy of fusion results (b) CMC curve of fusion process

The lowest overall EER and the maximum precision were obtained in the final fusion stage by processing each finger modality independently and then fusing the results. This illustrates how combining the outputs can greatly improve the system's accuracy and precision, even though individual processing may have certain limitations.

To sum everything up, Case 02 demonstrates the effectiveness of multimodal fusion, which combines the advantages of many modalities to produce better biometric identification results.

## 4.7 Conclusion

This chapter describes the development of an accessible biometric research-based finger knuckle print (FKP) identification system. In addition to multiple pre-trained CNN models, the system used a custom-trained CNN to extract features from unimodal biometric systems. We focus on transfer learning architectures in multimodal biometric identification systems. Based on the results, Densenet201 came out on top, showing excellent results; however, the fusion process achieved the best results, with a significant improvement in the identification rate of 97.5379%based on a dataset of 44 individuals.

Chapter 5

# *General conclusion and future work*

In recent years, the recognition of Finger-Knuckle Prints (FKP) as a means of identifying individuals has become an important addition to biometric modalities. This research delves into identifying individuals using their biometric descriptors, specifically emphasising the innovative FKP modality.

Our study's introduction to the basic ideas of biometric systems and assessment procedures laid the foundation for this dissertation's main topic. We examined the distinct features of unimodal and multimodal biometric systems, such as their architecture, data sources, and various information processing layers.

To improve the outcomes of our research, we employed different descriptors to extract features from unimodal and multimodal biometric systems. These methods were examined to enhance the identification rate in open-set and closed-set scenarios. Specifically, we tested our algorithms on a database of 44 individuals, achieving a notable identification rate of 97.5379%.

This study utilized transfer learning models such as AlexNet, VGG19, ResNet-50, and DenseNet-201 and CNN from scratch also in uni-model system. These models were instrumental in obtaining the best results, significantly enhancing the performance of our biometric system. To put things into perspective, We suggest employing other descriptor techniques in the future to boost the functionality of our biometric system further and increasing the size of our dataset to enhance accuracy for future research.Other biometric modalities could also be added to improve the system's functionality and adaptability.

# *Bibliography*


[1] D. Zhang, Y. Xu, W. Zuo, D. Zhang, Y. Xu, and W. Zuo, *Discriminative learning in biometrics*. Springer, 2016.

[2] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM systems Journal*, vol. 40, no. 3, pp. 614–634, 2001.

[3] A. K. Jain, S. Prabhakar, and S. Pankanti, "On the similarity of identical twin fingerprints," *Pattern Recognition*, vol. 35, no. 11, pp. 2653–2663, 2002.

[4] S. Aoyama, K. Ito, and T. Aoki, "A finger-knuckle-print recognition algorithm using phase-based local block matching," *Information Sciences*, vol. 268, pp. 53–64, 2014.

[5] S. Z. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 627–639, 2007.

[6] S. Pankanti, S. Prabhakar, and A. K. Jain, "On the individuality of fingerprints," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1010–1025, 2002.

[7] J. Trader, "M2sys blog on biometric technology," *Delta ID*, vol. 11, 2012.

[8] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1148–1161, 1993.

[9] T. Chihaoui, "Système d'identification de personnes basé sur la rétine," Ph.D. dissertation, Université Paris-Est; Université de Tunis El Manar, 2018.

[10] A. K. Jain, A. A. Ross, K. Nandakumar, A. K. Jain, A. A. Ross, and K. Nandakumar, "Fingerprint recognition," *Introduction to Biometrics*, pp. 51–96, 2011.

[11] J. D. Woodward and J. D. Woodward, *Biometrics: A look at facial recognition*. RAND Santa Monica, 2003.

[12] S. Akrouf, "Une approche multimodale pour l'identification du locuteur," Ph.D. dissertation, 2014.

[13] J. L. Wayman, "Fundamentals of biometric authentication technologies," *International Journal of Image and Graphics*, vol. 1, no. 01, pp. 93–113, 2001.

[14] M. Belahcen, "Authentification et identification en biométrie," Ph.D. dissertation, Université Mohamed Khider Biskra, 2013.

[15] S. I. Safie, R. Ramli, and Z. Mohamad, "Deep learning evaluation using receiver operating curve (roc) for footprint biometric authentication," in *2022 IEEE 8th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*. IEEE, 2022, pp. 26–29.

[16] L. Osadciw, P. Varshney, and K. Veeramachaneni, "Improving personal identification accuracy using multisensor fusion for building access control applications," in *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002.(IEEE Cat. No. 02EX5997)*, vol. 2. IEEE, 2002, pp. 1176–1183.

[17] A. Dube, D. Singh, R. K. Asthana, and G. S. Walia, "A framework for evaluation of biometric-based authentication system," in *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, 2020, pp. 925–932.

[18] G. Finocchiaro, "The regulation of artificial intelligence," *AI & SOCIETY*, pp. 1–8, 2023.

[19] B. Samia, Z. Soraya, and M. Malika, "Fashion images classification using machine learning, deep learning and transfer learning models," in *2022 7th International Conference on Image and Signal Processing and their Applications (ISPA)*. IEEE, 2022, pp. 1–5.

[20] S. T. Krishna and H. K. Kalluri, "Deep learning and transfer learning approaches for image classification," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 7, no. 5S4, pp. 427–432, 2019.

[21] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, pp. 211–252, 2015.

[22] D. Paper, *TensorFlow 2. x in the Collaboratory Cloud: An Introduction to Deep Learning on Google {u2019} s Cloud Service.* Apress, 2021.

[23] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85–117, 2015.

[24] R. El Saleh, "Biometrics for face skin analysis using machine learning based approaches," Ph.D. dissertation, Paris 12, 2021.

[25] P. Suyal, A. K. Bhatt, J. Pant, and L. Mohan, "Comparative analysis between traditional learning and deep learning," in *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N).* IEEE, 2021, pp. 376–379.

[26] N. Dif, "L'apprentissage profond pour le traitement des images," Ph.D. dissertation, Djillali Liabes University, 2020.

[27] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, 1996.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[29] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, "A novel connectionist system for unconstrained handwriting recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 5, pp. 855–868, 2008.

[30] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in *Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 609–616.

[31] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th international conference on Machine learning*, 2008, pp. 1096–1103.

[32] M. Borda, R. Terebes, R. Malutan, I. Ilea, M. Cislariu, A. Miclea, and S. Barburiceanu, *Randomness and Elements of Decision Theory Applied to Signals*. Springer, 2021.

[33] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality-preserving dimensionality reduction and classification for hyperspectral image analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1185–1198, 2011.

[34] C. C. Aggarwal and C. C. Aggarwal, "Domain-specific neural architectures," *Artificial Intelligence: A Textbook*, 2021.

[35] M. A. Wani, F. A. Bhat, S. Afzal, and A. I. Khan, *Advances in deep learning*. Springer, 2020.

[36] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[37] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[38] N. Charfi, "Biometric recognition based on hand shape and palmprint modalities," Ph.D. dissertation, Ecole nationale supérieure Mines-Télécom Atlantique, 2017.