PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY OF KASDI MERBAH OUARGLA

Faculty of New Technologies of Information and Communication

Department of Electronics and Telecommunications

MASTER

Field: Electronics

Specialty: Electronics of embedded systems

Presented by: Nada HIDEB & Sirine TALBI & Safa ZERGUI

Theme:

Personal Biometric Recognition-Based on Transfer Learning

Publicly Supported On: 24/06/2021

Before the jury:

Mrs Charif Fella	MCA	President	UKM Ouargla
Mr Korichi Maarouf	MCB	Examiner	UKM Ouargla
Mr Samai Djamel	MCA	Supervisor	UKM Ouargla
Miss Trabelsi Selma	PhD student	Co-Supervisor	UKM Ouargla

Acknowledgment

First of all, we thank ALLAH who gave us the strength and courage to complete our studies and develop this work.

We would like to thank the members of the jury for their careful reading of our thesis as well as for the comments they will address to us about this thesis in order to improve our work.

We would like to express our gratitude to our thesis supervisor, Mr SAMAI Djamel, for his patience, his availability and above all his judicious advice, which contributed to fuel our reflection, and whom without his methodology, his precious orientations, we could not have completed this thesis. We also would like to thank our co-supervisor Miss TRABELSI Selma for having had the patience to answer our countless questions.

Thanks for all the people who contributed directly or indirectly to the success of this project and who helped us in writing this thesis.

Dedication

We dedicate this thesis

To our fathrs Mr HIDEB Laid, TALBI El azhar, ZERGUI Brahim. No dedication can express the love, esteem, dedication and respect that we always have for you. Nothing in the world is worth the efforts made day and night for our education and well-being. This work is the fruit of the sacrifices you have made for our education and formation. Receive through this work, all our gratitude and our deep feelings. May Allah Almighty protect you and keep you in his vast paradise.

To our very dear mothers Mrs HAMDADACHE Ouardiya, ATLILI Malika, ABDOU Khadidja, for giving us life and the joy of living. Your good teaching, guidance and blessings have never been lacking, the source of tenderness and the example of dedication that has never ceased to encourage and pray for us, the light of our day, the source of our efforts, the flame of our hearts, our lives and our happiness, our mothers whom we adore. May Allah Almighty be by your side.

Thank you for you our dear parents, with great shame, because if we light our fingers like candles, we will not be able to reimburse you even if a little of what you gave us during our lifetime.

To our sisters and brothers for their great moral support throughout the journey we have taken to prepare this thesis.

Thank you everyone, from the oldest to the youngest, thank you very much. HIDEB Nada, TALBI Sirine, ZERGUI Safa.

Abstract

In recent years, biometrics has been included in all security systems with different forms: fingerprint, iris, signature, etc. Nowadays, biometrics is a solution to several problems of automatic identification of individuals. In this document, we gave a definition of the biometric system, showed its categories, the different processes that it can work in, and we also demonstrated its different composition modules and how can we evaluate its performances on several standards, furthermore we spoke about two biometric techniques, and those were Uni/Multimodal systems and revealed how is it that multimodal biometric system is better than the unimodal. We will experiment a biometric system based on palmprint and palmvein by using transfer learning which is used in deep learning applications. Transfer learning is that a preformed network has learned a rich set of characteristics that can be used as a starting point for learning a new task. In our work, the effectiveness of transfer learning was assessed on two publicly available databases **(PolyU and CASIA)** by comparing the spectral bands within each database while using one of the models of convolutional neural networks **(AlexNet)** and choosing the best band in the uni-modal system for each database. Finally, we will apply a multimodal system for each database using the fusion of the unimodal scores in order to improve the performance and get better results.

Keywords:

Biometrics.

Deep learning.

Transfer learning.

Palmprint.

Palmvein.

AlexNet.

PolyU and CASIA.

Résumé

Ces dernières années, la biométrie a été intégrée dans tous les systèmes de sécurité sous différentes formes: empreinte digitale, iris, signature, etc. De nos jours, la biométrie est une solution à plusieurs problèmes d'identification automatique des individus. Dans ce document, nous avons donné une définition du système biométrique, montré ses catégories, les différents processus dans lesquels il peut fonctionner, et nous avons également démontré ses différents modules de composition et comment évaluer ses performances sur plusieurs normes, de plus nous avons parlé de deux techniques biométriques, et celles-ci étaient des systèmes Uni/Multimodales et ont révélé comment se fait-il que le système biométrique multimodale soit meilleur que le système unimodale. Nous expérimenterons un système biométrique basé sur l'empreinte palmaire et la veine palmaire en utilisant l'apprentissage par transfert qui est utilisé dans les applications d'apprentissage en profondeur. L'apprentissage par transfert est qu'un réseau préformé a appris un riche ensemble de caractéristiques qui peuvent être utilisées comme point de départ pour apprendre une nouvelle tâche. Dans notre travail, l'efficacité de l'apprentissage par transfert a été évaluée sur deux bases de données accessibles au public **(PolyU et CASIA)** en comparant les bandes spectrales au sein de chaque base de données tout en utilisant l'un des modèles de réseaux de neurones convolutifs **(AlexNet)** et en choisissant la meilleure bande dans le système unimodale pour chaque base de données. Enfin, nous appliquerons un système multimodale pour chaque base de données en utilisant la fusion des scores unimodales afin d'améliorer les performances et d'obtenir de meilleurs résultats.

Mots clés:

La Biométrie.

L'apprentissage en profondeur.

Transférer l'apprentissage.

Empreinte palmaire.

La veine palmaire.

AlexNet.

PolyU et CASIA.

ملخص

في السنوات الأخيرة ، تم تضمين القياسات الحيوية في جميع أنظمة الأمان بأشكال مختلفة: بصمات الأصابع ، قزحية العين ، التوقيع ، إلخ. في الوقت الحاضر ، تعد القياسات الحيوية حلاً للعديد من مشاكل التعرف التلقائي على الأفراد . في هذا المستند ، قدمنا تعريفًا لنظام القياسات الحيوية ، وأظهرنا فئاته ، والعمليات المختلفة التي يمكن أن يعمل بها ، كما أوضحنا أيضًا وحدات تكوينه المختلفة وكيف يمكننا تقييم أدائه على عدة معايير ، علاوة على ذلك تحدثنا عن تقنيتي القياسات الحيوية ، وتلك كانت أنظمة (أحادية / متعددة) الوسائط, كما كشفنا كيف أن نظام القياسات الحيوية متعدد الوسائط أفضل من النظام أحادي الوسائط. سنقوم بتجربة نظام القياسات الحيوية على بصمة اليد و عروق اليد باستخدام تعلم النقل المستخدم في تطبيقات التعلم العميق. تعلم النقل هو أن شبكة سابقة التشكيل قد تعلمت مجموعة غنية من الخصائص التي يمكن استخدامها كنقطة بداية لتعلم مهمة جديدة. في عملنا تم تقييم فعالية تعلم النقل على قاعدتي بيانات متاحة للجمهور **(PolyU, CASIA)** من خلال مقارنة النطاقات الطيفية داخل كل قاعدة بيانات اثناء استخدام أحد نماذج الشبكات العصبية التلافيفية **(AlexNet)** واختيار أفضل نطاق في النظام أحادي الوسائط لكل من قاعدة البيانات. أخيرًا، سنطبق نظامًا متعدد الوسائط لكل قاعدة بيانات باستخدام دمج الدرجات أحادية الوسائط من أجل تحسين الأداء والحصول على نتائج أفضل.

الكلمات المفتاح:

القياسات الحيوية.

التعلم العميق.

تعلم النقل المستخدم.

بصمة اليد.

عروق اليد.

الشبكات العصبية التلافيفية **(AlexNet)**.

قاعدتي بيانات متاحة للجمهور **(PolyU, CASIA)**.

List of contents

List of figures

List of tables

List of abbreviations

PIN	Personal Identification Number
DNA	DeoxyriboNucleic Acid
FAR	False Accept Rate
FRR	False Reject Rate
GAR	Genuine Acceptance Rate
EER	Equal Error Rate
T_0	Threshold
ROC	Receiver Operating Characteristic
CMC	Cumulative Match Characteristic
UBS	Uni-modal Biometric System
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
ANN	Artificial Neural Networks
ConvNet or CNN	Convolutional Neural Network
GLM	Generalized Linear Model
RGB	Red Blue Green
ReLU	Rectification Linear Unit
LSVRC	Large-Scale Visual Recognition Challenge
NIR	Near-Infra Red
SUM	SUMmation, Addition
MUL	MULtiplication, Production
ROR	Rank One Recognition
RPR	Rank of Perfect Recognition
WHT SUM	WeigHTed SUMmation
WHT PROD	WeigHTed PRODuction
MIN	MINimum
MAX	MAXimum

General Introduction

We frequently need to authenticate our own/ others identities, or ascertain who someone is, as we go about our daily lives. Life is easier when you have reliable identification system that helps separate good people of the general population from known criminals and other risks, for example, to improve public safety. Financial and business transactions are also safer and more efficient with reliable identification. Automated authentication makes it possible to personalize the way that a device reacts to different persons and to ensure that it confidently responds to people in a correct manner. In fact, this includes two independent acts: an authentication system confirms the identity, and a second authorization system relates the necessary actions to a person's identification. Biometrics employs different physiological, behavioral and biological traits to identify an individual. It is one of the most reliable and convenient verification and identification tools, with growing worries about security breaches and transaction fraud [1].

Biometrics is the science of recognizing a person's identity through the components of their bodies based on their physiological, behavioral and biological characteristics. Fingerprints, palms, and feet are the most unique of these indicators, computers can match them with marks and distinguishing spots in seconds. Also, you can identify yourself using facial traits, voice, hand geometry, or the iris of the eye…etc. The biometric characteristics must at least guarantee the following conditions: universality, Permanence, acceptability... etc.

In this work, one of these biometric systems was chosen for the recognition of people which is by palm print. The palm print is an important biometric modality, with a distinctive, stable and highly differentiable nature. Its study has garnered a lot of interest in recent years. While various recognition methods based on palm prints have been developed, and effectively implemented for identity identification most previous research typically only employs images taken in natural light. We propose to employ efficient multispectral recognition for the palm print in this thesis, and to construct an identification system based on multispectral images, in order to achieve a high recognition rate with more discriminating information. The goal of our research is to develop a uni-modal and multimodal biometric systems using a deep learning-based transfer learning method.

We'll strive to accomplish our objectives in three chapters:

• **In the first chapter**, an overview of biometrics will be presented.

• **The second chapter** is devoted to deep learning, with a more in-depth look at convolutional neural networks (CNN) and transfer learning, the method we used in our research.

• **The third chapter** will contain a description of our proposed models as well as the outcomes.

• We'll wrap off this thesis with a basic conclusion, and discuss our work's perspectives.

Chapter I

Biometrics Generalities

I.1 Introduction

Person authentication and identification are becoming increasingly important in today's automated world. For these purposes, the latest technology of using a Personal Identification Number "PIN" or password barely meets the criteria of an identification system since one must remember too many passwords and the password or PIN are extremely vulnerable. As a result, identifying an individual remains a significant challenge in today's diverse, geographically mobile, and increasingly electronically wired information society [2]. In light of recent scientific and technological advancements in the field of security and individual recognition, the world has been trying to break away from the conventional situation in order to ensure the so-called mechanism and security, as well as the lack of penetration and falsification. This was the catalyst for the creation of the field of biometrics, which piqued the interest of many scientists and researchers. As a result, the subject of this chapter is biometrics, including basic principles, meanings, and everything else relevant to biometric systems.

I.2 Definition of biometrics

Biometrics is originated from a Greek term composed of two words "Bios" for "life" and "Metricos" for "measure", directly translates into "life measurement" [3]. One of the automated methods of identification and recognition is biometrics [4]. Biometrics is a relatively new technique that allows us to verify an individual's identity by using one or more of their personal characteristics. Its benefit is that, unlike passwords or personal identification numbers, it uses data as an identity that cannot be lost, stolen, or tampered with [5] (see Figure I.1).

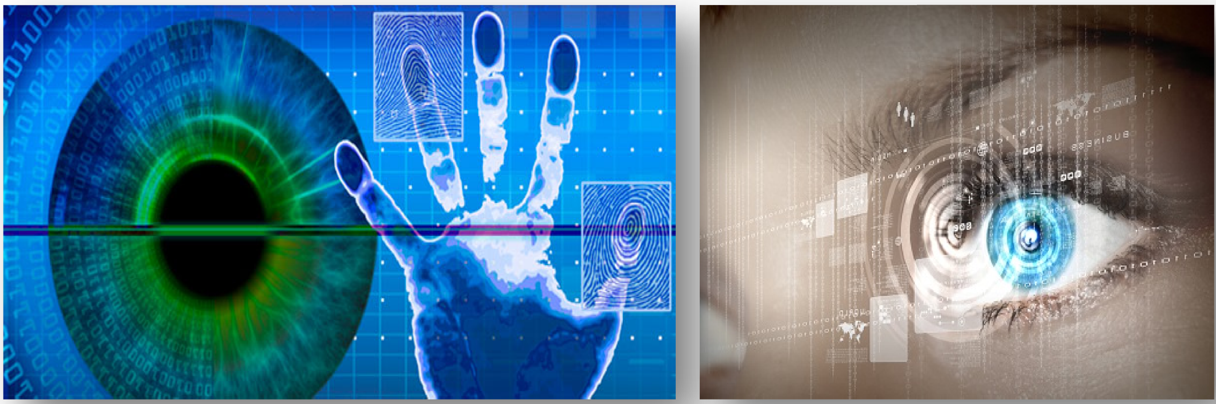


Figure I.1- Biometric technology.

I.3 Biometrics categories

Biometric methods are used in a variety of fields, with the most common biometric methods falling into the following categories (see Figure I.2):

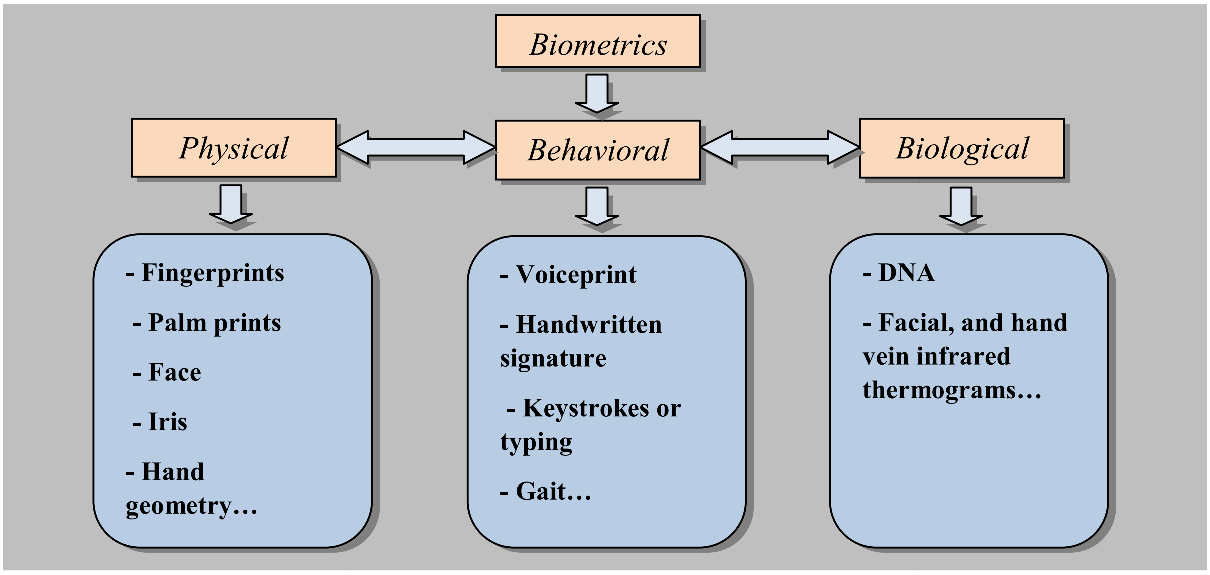


Figure I.2- Classification of a number of biometric modalities.

I.3.1 Physical characteristics: They are focused on everything related to the physical appearance of the various organs in the human body, including the following:

• **Fingerprint:** Fingerprints may be used to authenticate personal identity because the patterns of ridges and valleys on an individual's fingertips are special to that individual. Fingerprints are so different that even identical twins rarely have the same one [6] (see Figure I.3).

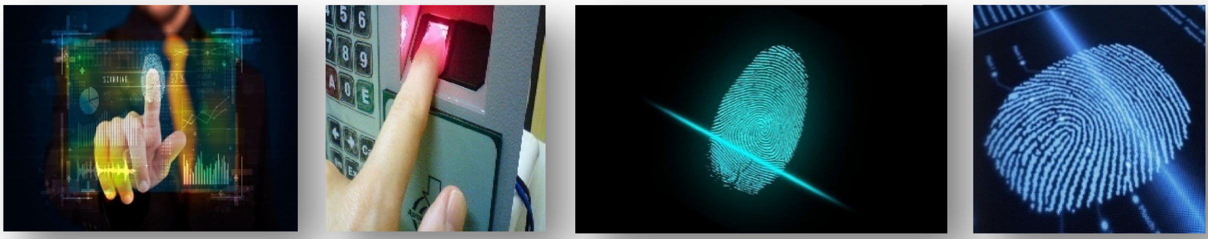


Figure I.3- Fingerprint images.

• **Palm print:** The inner surface of the palm contains a variety of unique identifying features that allow for accurate and reliable personal identification. Palm prints, like fingerprints, have permanent distinguishing features like ridges and valleys, minutiae, and even pores in high-resolution (>1000 dpi) images. Apart from these quasi-fingerprint characteristics, palm prints have other distinct characteristics such as principal lines and wrinkles. It is possible to retrieve all types of palm print features using a high-resolution capture device to create a highly accurate biometric system [6] (see Figure I.4).

Figure I.4- Palm prints images.

• **Iris:** Iris is the colored region that covers the pupil, is believed to have distinctive patterns. A video based image acquisition system can be used to acquire iris patterns. The iris based biometric devices proved their ability to work with individuals regardless of ethnicity or nationality [6] (see Figure I.5).

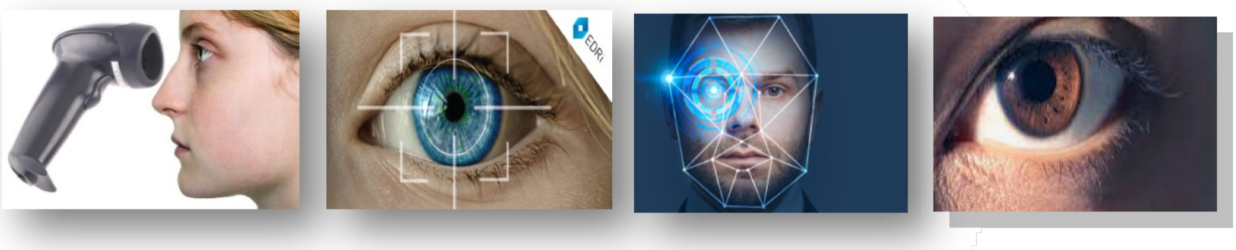


Figure I.5- Iris images.

• **Face:** Face is a natural human characteristic for automated biometric recognition. Face recognition systems typically utilize the spatial relationship among the locations of facial features like the eyes, nose, lips, chin, and the overall appearance of a face [7]; they consist of an images acquisition module with a camera which first detects the face in the obtained image, then there processes using algorithms to extract a signature from the face. Lastly, this signature is compared, through using a classifier, with the already existing signatures in the local database to identify the individual [8] (see Figure I.6).

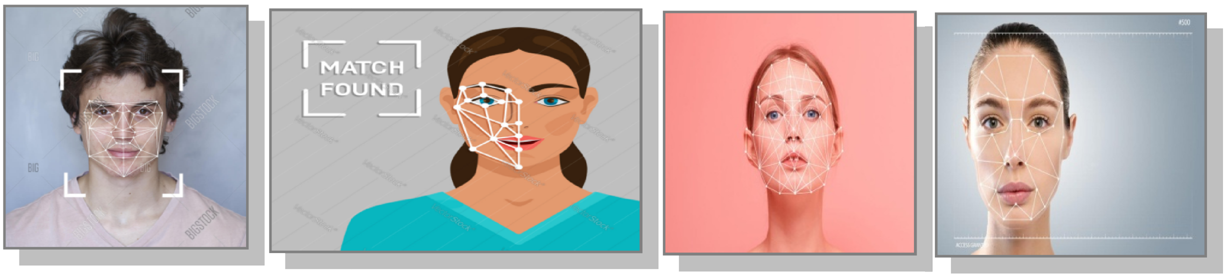


Figure I.6- Face images.

I.3.2 Behavioral characteristics: They are focused on an examination of a person's behavior.

• **Voice print:** Voice is a biometric mixture of physical and behavioral characteristics. The shape and size of the appendages used in sound synthesis (e.g., vocal tract, mouth, nasal cavities, and lips) determine the physical characteristics of an individual's speech. These physical features of human speech remain constant for a person, but the behavioral aspects of speech change over time as a result of age, medical conditions (such as common cold), emotional state, and other factors [9] (see Figure I.7).

Figure I.7- Voice print images.

• **Signature:** Signature authentication entails a dynamic examination of a signature in order to verify a person's identity. When an individual sign a document, a signature-based system can monitor their pace, pressure, and angle. This technology could be useful in e-business, where signatures could be used as a form of personal authentication [6] (see Figure I.8).

Figure I.8- Signature images.

• **Keystrokes:** Keystroke identification technology is a vital component of biometrics. Since it has the least effect on privacy, keystroke identification is thought to be the best way to authenticate an individual. The method of measuring and analyzing human typing rhythm on digital devices such as a computer keyboard, cell phone, or touch screen panel is known as keystroke identification technique. The form and manner in which we type on our computer keyboard differs from person to person and is considered a special behavioral biometric [10] (see Figure I.9).

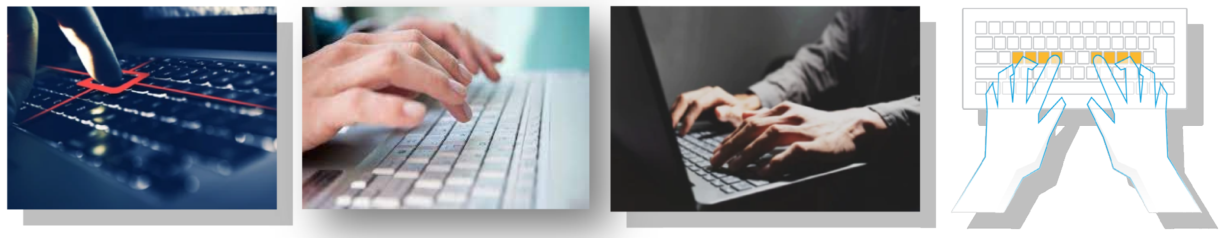


Figure I.9- Keystrokes or typing images.

• **Gait:** Gait is dynamic spatiotemporal behavioral biometrics that describes how one moves. Gait isn't supposed to be unique to each individual, but it should be distinctive enough to allow for identity verification. Gait may not stay invariant especially over a large period of time owing to significant changes in body weight (e.g., waddling gait during pregnancy), major accidents affecting joints or brain etc... As a result, gait could be a viable biometric. [11] (see Figure I.10).

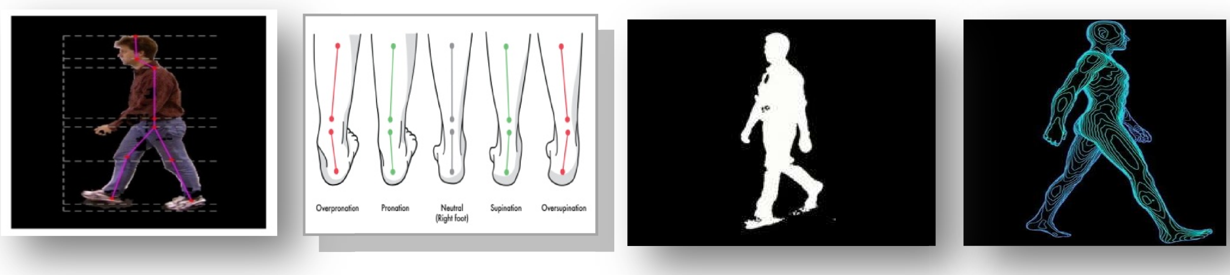


Figure I.10- Gait images.

I.3.3 Biological characteristics: These modalities are dependent on an individual's biological characteristics (DNA, saliva, Odor). Only in the most serious cases can this form of biometrics be used.

• **DNA:** Deoxyribonucleic Acid "DNA", which is represented by a one-dimensional code that is unique to each entity, is possibly the most accurate biometric, though it will fail to distinguish the identities of identical twins when used for personal authentication. Other issues with a biometric system based on DNA include privacy concerns, potential contamination, and poor DNA matching reliability [6] (see Figure I.11).

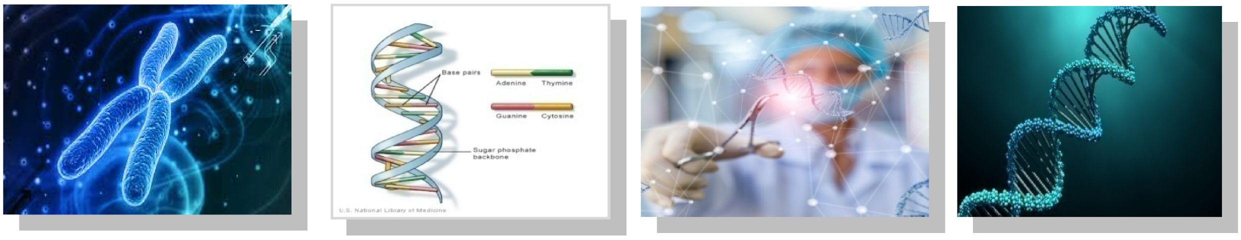


Figure I.11- DNA images.

• **Facial, and hand vein infrared thermograms:** The pattern of heat radiated by the human body is unique to each person and can be recorded by an infrared camera in a non-intrusive way, similar to a normal (visible spectrum) image. It's also possible that the technology could be used for covert identification [**9**] (see Figure I.12).

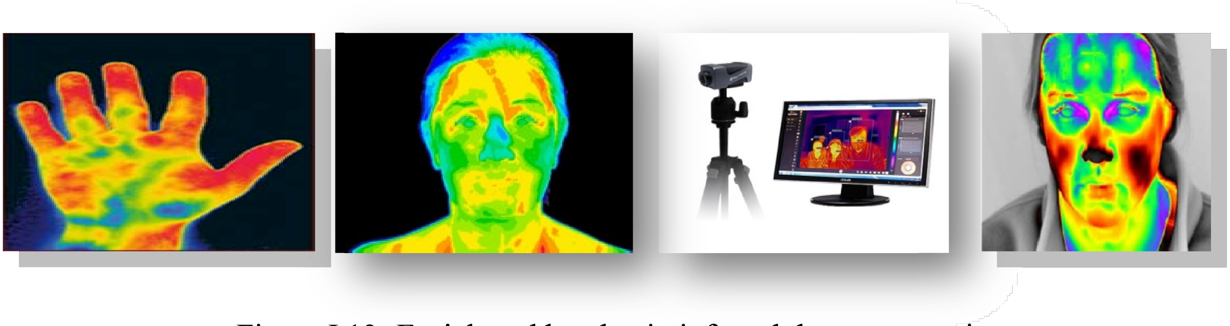


Figure I.12- Facial, and hand vein infrared thermograms images.

I.4 Biometric characteristics requirements

• **Universality:** Means that the trait should be present in everyone.

• **Uniqueness:** Means that any two people should be sufficiently different to differentiate themselves based on this attribute.

• **Permanence:** Means that the trait should be constant and not change dramatically over time or in response to the environment.

• **Collectability:** Refers to the ability to calculate a trait.

• **Acceptability:** Measure shows how likely people are to accept the biometrics scheme.

• **Performance:** Refers to the achievable identification accuracy, the resources needed to achieve appropriate identification accuracy, and the working or environment factors that influence identification accuracy.

• **Circumvention:** Refers to how simple it is to deceive the mechanism using deceptive methods. Only after examining vast amounts of samples can it be determined if a human trait is ideal for biometrics systems. The followings are some of the physical and behavioral traits that biometrics use in general [2]:

Table I.1- Biometric characteristics that are used to compare various biometric techniques.

Biometric	Universality	Uniqueness	Collectability	Permanence	Performance	Acceptability	Circumvention
Fingerprint	Medium	High	Medium	High	High	Medium	Medium
Iris	High	High	High	High	High	Medium	Low
DNA	High	High	Low	High	High	Low	Low
Palm print	Medium	High	Medium	High	High	Medium	Medium
Hand Vein	Medium	Medium	Medium	Medium	Medium	Medium	Low
Signature	Low	Low	High	Low	Medium	High	High
Keystroke	Low	Low	Medium	Low	Low	Medium	Medium
Voice	Medium	Low	Medium	Low	Low	High	High

I.5 Processes involved in biometric authentication system

“Enrollment”, “Identification”, and “Verification” are the three modes of operation for the biometric system.

I.5.1 Enrollment: It is the first phase of any biometric system where the user is registered in the system for the first time and where one or more biometric modalities are captured and registered in a database along with biographic information [8] (see Figure I.13).

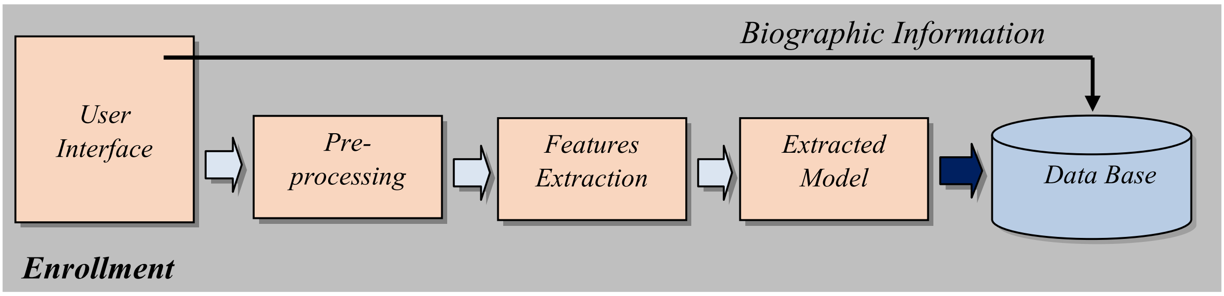


Figure I.13- Enrollment in a biometric system.

I.5.2 Verification: Verification, also known as authentication necessitates the assertion of the user’s identity (ID card, smart card, or PIN) and presenting a biometric modality, whereas the system checks if this assertion is valid or not, by comparing the captured biometric data only with the registered model corresponding to the user. This is referred to as a 1: 1 matching. The assertion is valid if the user's biometric input and the model registered in the database corresponding to the asserted identity have a high degree of similarity. Otherwise, the claim is rejected. The aim from verification is to prevent multiple people from using the same identity [8, 12] (see Figure I.14).

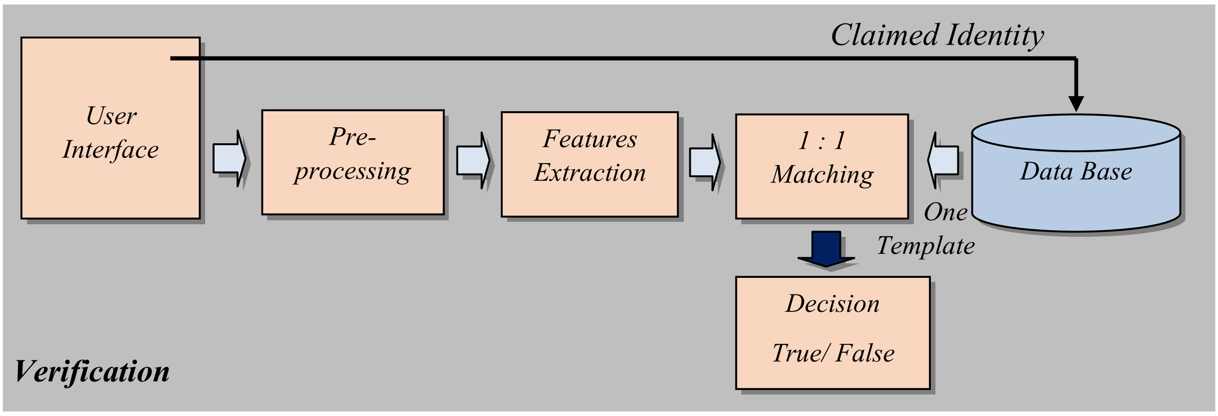


Figure I.14- Verification in a biometric system.

I.5.3 Identification: It is the process of identifying a user solely based on their biometric details, without any previous knowledge of their identity. The biometric system compares the biometric sample with the models of all the people in the database. This is referred to as a 1: N matching. The person is rejected if the greatest similarity between the sample and all models is less than a fixed minimum security threshold .Otherwise, the person is accepted [8, 12] (see Figure I.15). During this phase, there are two types of identification:

• **Closed set identification:** This type of identification is used to record the presence of people in certain company for example. If the sample has a certain degree of similarity with the samples in the system, the person will be accepted.

• **Open set identification:** If there is a great similarity between the tested biometric sample and all models pre-recorded and if this similarity is lower or higher than the safety threshold, this no one is rejected. This means that the person is not one of those registered by the system.

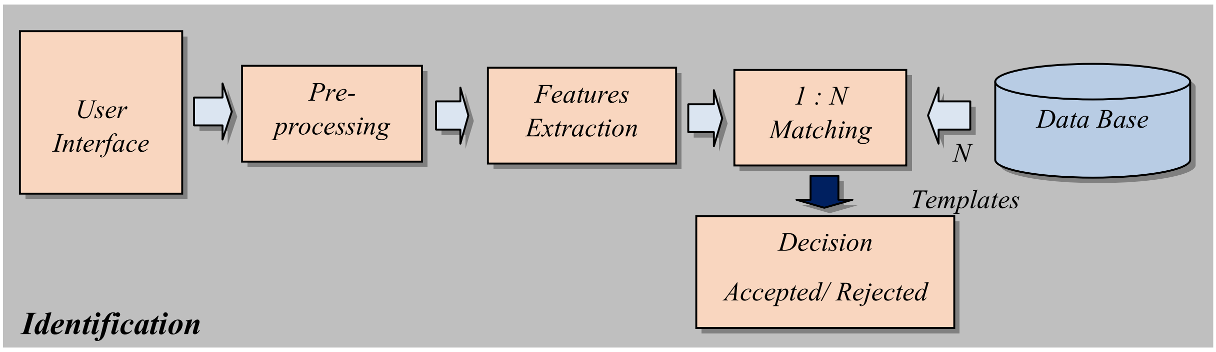


Figure I.15- Identification in a biometric system.

I.6 Basic modules in a biometric system

• **Sensor module:** Corresponds to measure or record the raw biometric data of the user in the form of video, audio and an image or some other signal [9, 13] by means of a suitable user interface incorporating a terminal biometric sensor or reader [9].

• **Pre-treatment module:** Corresponds to the selection of the region of interest for features extraction where the raw biometric data is pre-processed by 3 steps “quality assessment” to determine the suitability of the acquired biometric samples for further processing, “segmentation” to remove the background noise, and “enhancement” to improve its quality and further reduce the noise [9].

• **Features Extraction module:** Refers to an automated process of the biometric data so that a set of notable discriminatory features only are extracted which are essential for recognizing the person to represent the underlying trait, this feature set is commonly referred to as a “template” [14] which can be done by using Machine learning, Computer Vision and Pattern Recognition techniques [13].

• **Database module:** In which the template of enrolled users is stored in along with some biographic information characterizing the user [9].

• **Matching module:** Measures the similarity between the extracted features and the template stored in the database by comparing or matching the two, this measure is termed “matching score” [9].

• **Decision module:** Verifies a claimed identity or identifies the identity of the user based on the matching score [15].

I.7 Uni-modal and multimodal systems

Based on the number of patterns considered to be of a person's identity, biometric techniques can be divided into two categories [16].

I.7.1 Uni-modal Biometric Systems: A Uni-modal Biometric System “UBS” is a system that makes use of single biometric trait for identification or verification of an individual (palm print, palm veins, iris, or the fingerprint).

• **Limitation of the uni-modal biometric systems:** The error rates associated with uni-modal biometric systems are quite high, which reduces their effectiveness in security applications. Although these systems have many advantages, they also have to face a variety of problems like “Noisy Data” in which the sensibility of biometric sensors to noise leads to inaccurate matching, as noisy data may lead to false rejection. The biometric data acquired during verification will not be identical to the data used for generating template during enrolment for an individual. This is known as “Intra-class Variation”. Large intra-class variations increase the False Rejection Rate “FRR” of a biometric system. As well as “Interclass Similarities” those refer to the overlap of feature spaces corresponding to multiple individuals. Large inter-class similarities increase the False Acceptance Rate “FAR” of a biometric system, along with “Non universality” where some persons cannot provide the required standalone biometric, owing to illness or disabilities. And lastly “Spoofing”; Uni-modal biometrics is vulnerable to spoofing where the data can be imitated or forged. Each biometric trait has its advantages and limitations, and no single trait is anticipated to effectively meet all the requirements such as accuracy, practicality, and cost imposed by all applications. Therefore, there is no trait which is universally the best. Due to the issues discussed above, the performance of uni-modal biometric based system degrades; in order to improve it the usage of more than one biometric is preferred [17, 18].

I.7.2 Multimodal Biometric Systems: Multi modality is the combination of several biometric modalities with the aim of reducing the limitations of uni-modal systems. For example: face biometrics + fingerprint biometrics. In addition, multimodal biometric systems reduce the risk of impossibility of registration as well as the robustness to fraud.

I.7.2.1 Multi-biometric system categories: Recognition systems that use multiple biometric features work in one of the following integration scenarios:

• **Multi-sensor systems:** Information of the same biometrics obtained from different sensors is combined for all. For example, additional information about fingerprints can be captured with different types of sensors (such as optical and capacitive sensors). The information obtained is then integrated at the sensor level using fusion technology [19].

• **Multi-modal systems:** More than one biometric feature is used for user identification. For example, information obtained using the face and voice or other features can be integrated to establish the identity of the user. It can be more expensive; because it requires multiple sensors and each sensor recognizes different biometric properties. However, the improvement in performance is significant [19].

•**Multi-instance systems:** Several instances of a single biometric feature are recorded. For example, left and right iris images can be used for iris recognition. In addition, the fingerprints of two or more fingers of a person can be combined or an image of the same person can be combined. If a single sensor is used to capture these images one at a time, the system can be made really inexpensive as it does not require multiple sensors. In addition, no additional modules are integrated for the extraction and combination of functions [19].

• **Multi-sample systems:** Several samples with the same biometric character are used for registration and identification. For example, the left and right profiles are also recorded with the front side. Multiple prints of the same finger and multiple samples of a voice can be combined. Multiple samples can overcome poor performance. However, this will require multiple copies of sensors, or the user may wait longer to be detected, or a combination of both [19].

• **Multi-algorithm systems:** Different approaches to feature extraction and matching algorithms are applied to a single biometric feature. Final decision whether one of the matching fusion techniques can be applied to the results obtained using different matching algorithms. These systems are less expensive because no additional device is required to collect the data. However, these are more complex due to the use of different algorithms [19] (see Figure I.16).

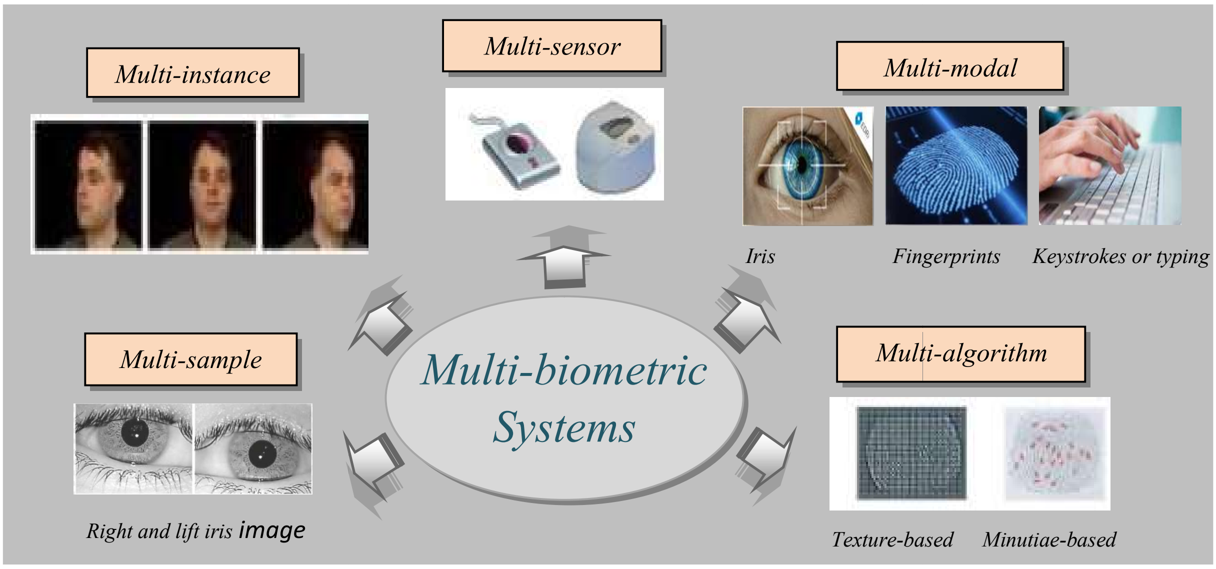


Figure I.16- Different types of multi-biometric systems.

I.7.2.2 Operating modes: A multimodal biometric system can operate in one of three modes: serial mode, parallel mode, or hierarchical mode. In the serial mode of operation, the output of a biometric feature is typically used to reduce the number of possible identities before the next feature is used. This serves as an indexing scheme in an identification system. For example, a multimodal biometric system that uses face and fingerprints may first use facial information to retrieve early matches and then use fingerprint information to converge on a single identity. This is in contrast to a parallel mode of operation, in which information from several features is used simultaneously to perform recognition. This difference is crucial. In the serial operating mode, the different biometric properties do not have to be recorded at the same time. A decision could also be made without acquiring all of the properties. This reduces the overall recognition time. In the hierarchical scheme, the individual classifiers are combined into a tree structure [16] (see Figure I.17).

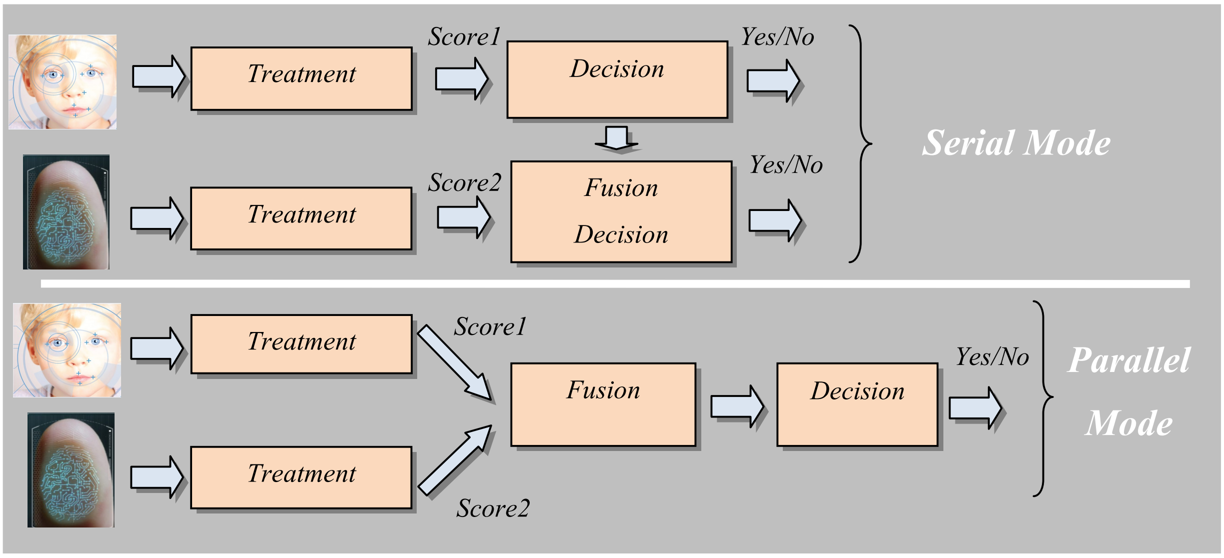


Figure I.17- Serial and Parallel Modes.

I.7.2.3 Fusion levels: Multimodal biometric systems integrate information represented by multiple biometric indicators. Information can be consolidated at different levels. It illustrates the four levels of fusion when combining two (or more) biometric systems [16].

• **Fusion at sensor level:** Many images can be combined into a single image. The image merging method seeks to solve the problem of combining information from multiple images captured of the same object to obtain a new merged image. The wavelet-based approach is widely used in image fusion [20] (see Figure I.18).

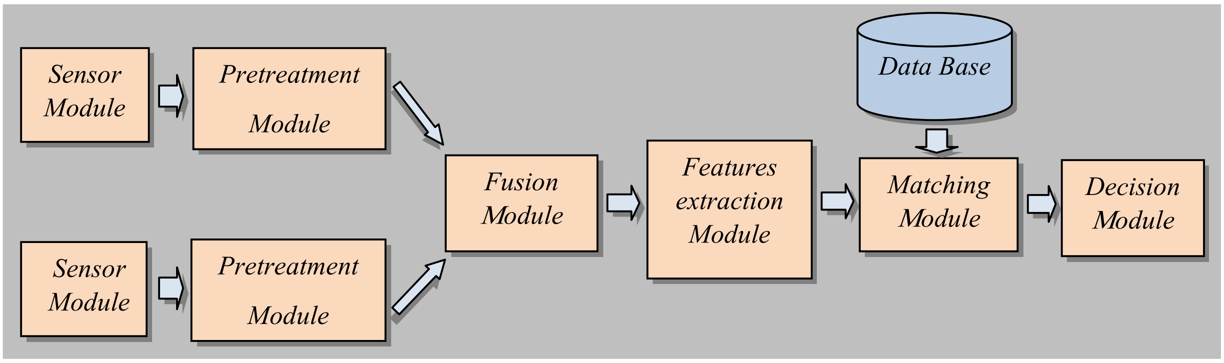


Figure I.18- Parallel mode fusion at sensor level.

• **Fusion at the feature extraction level:** The data obtained from each biometric modality is used to calculate a feature vector. If the features extracted from one biometric indicator are (somewhat) independent of those extracted from the other, it makes sense to concatenate the two vectors into a single new vector, provided that the features of the different biometric indicators relate to the same type of measurement Frame. The new feature vector has a higher dimensionality and represents the identity of a person in a different (and hopefully more discriminatory) feature space. Feature reduction techniques can be used to extract a small number of salient features from the larger set of features [16] (see Figure I.19).

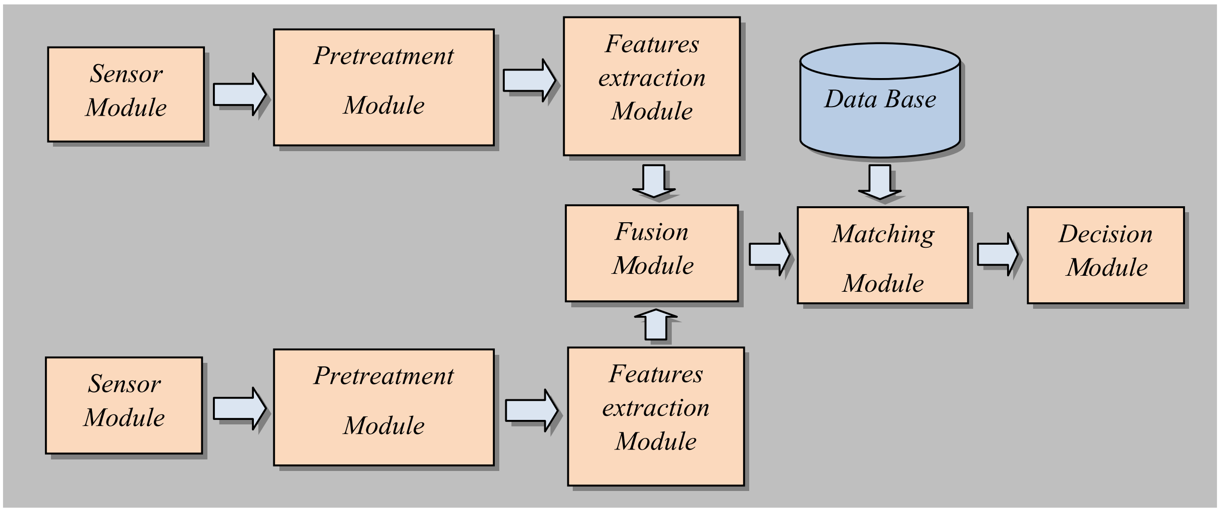


Figure I.19- Parallel mode fusion at the feature extraction level.

• **Fusion at the matching module level:** Each biometric comparator provides a similarity score that indicates how close the input feature vector is to the template feature vector. These ratings can be combined to confirm the correctness of the claimed identity. Techniques such as weighted average can be used to combine match results reported by multiple comparators [16] (see Figure I.20).

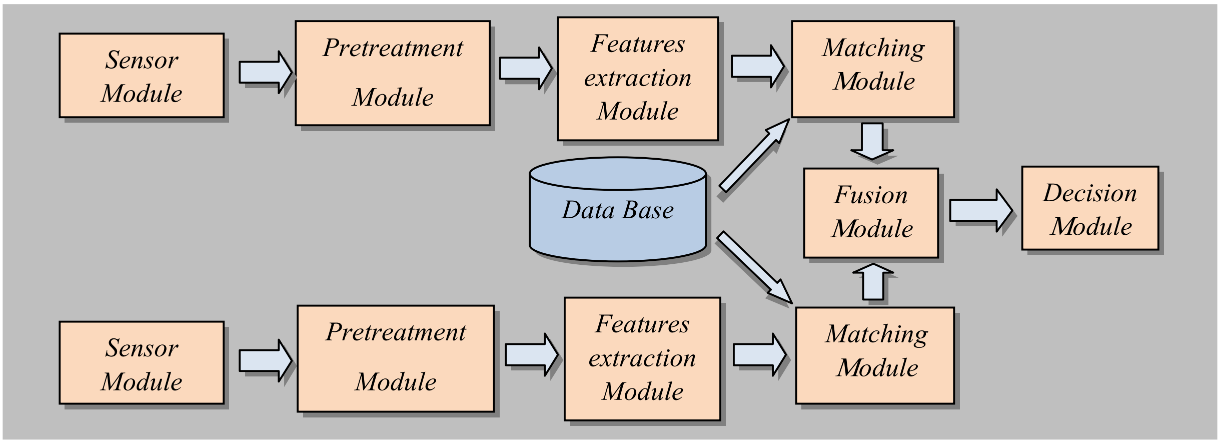


Figure I.20- Parallel mode fusion at matching level.

• **Fusion at the decision module level:** Each biometric system makes its own recognition decision based on its own feature vector. A majority voting system can be used to make the final recognition decision [16] (see Figure I.21).

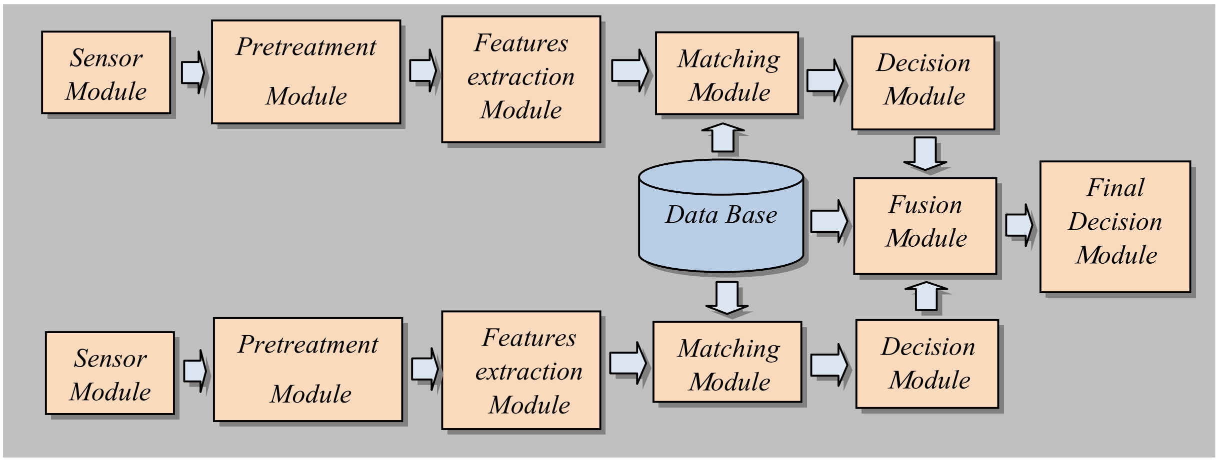


Figure I.21- Fusion in a parallel mode at decision level.

I.8 Biometric system performances measuring

Several quality performance metrics measure the performance of any biometric authentication technique. Helps you compare systems and motivates progress [19].

I.8.1 False Accept Rate (FAR): Confusing biometric measurements from two different people to give the impression that they are from the same person as the users are very similar. It measures the percentage of invalid matches [19].

$$\mathrm{FAR} = \frac{\text{Total number of forgeries accepted}}{\text{Total number of forgeries submitted to the system test}} \times 100$$

I.8.2 False Reject Rate (FRR): Confuse two biometric measurements of the same person to come from two different people due to large differences within the class. Measures the percentage of valid entries rejected [19].

$$\mathrm{FRR} = \frac{\text{Total number of genuine test pattern rejected}}{\text{Total number of genuine test submitted to the system}} \times 100$$

I.8.3 Genuine Acceptance Rate (GAR): Genuine Acceptance Rate (GAR) that is sometimes used. It is the percentage of the chances that a real person will be recognized as a match. The GAR of a valid user can be obtained by the equation [19].

$$\mathrm{GAR} = 1 - \mathrm{FRR}\ \%$$

I.8.4 Equal Error Rate (EER): It is used to summarize the performance of a biometric system defined at the point where (FRR) and (FAR) are equal [19] (see Figure I.22).

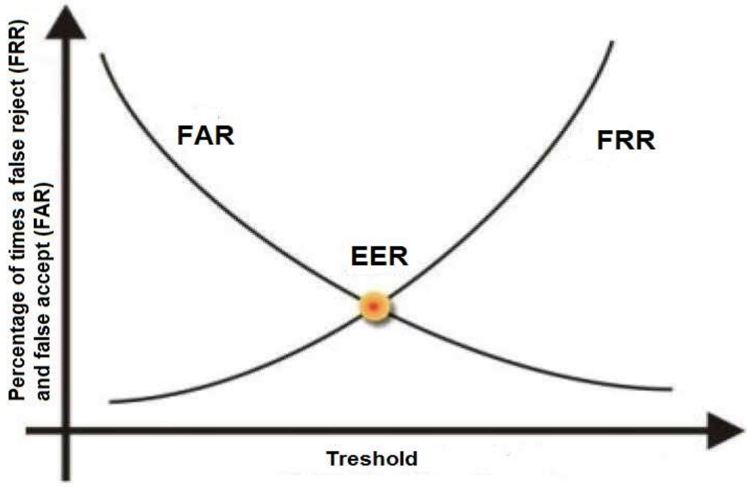


Figure I.22- Equal Error Rate point (EER).

I.8.5 Receiver Operating Characteristic (ROC): In every biometric system there is a trade off between FAR and FRR. In fact, both are functions of the system threshold (t); if it is rejected to allow the system to tolerate more input fluctuation and noise, the FAR will increase. On the other hand, if it is increased to make the system more secure, the FRR will increase accordingly. The ROC diagram is obtained by plotting the FAR values against the FRR at different operating points (threshold values) on a linear or logarithmic or semi-logarithmic curve [19] (see Figure I.23).

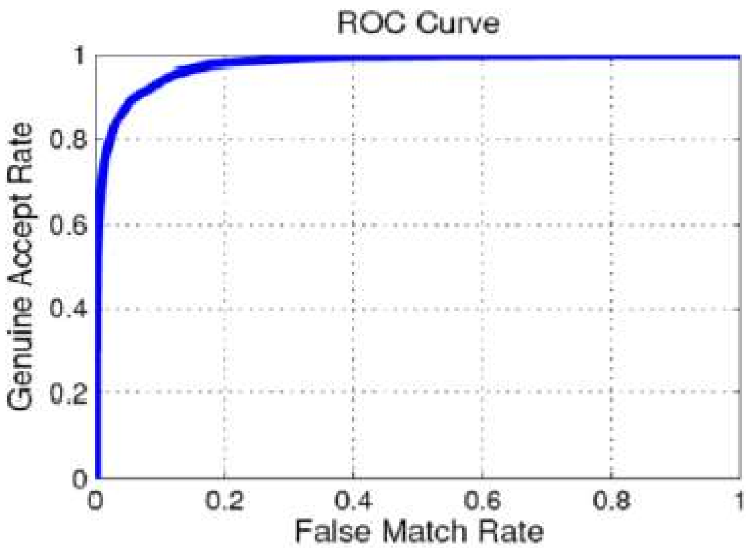


Figure I.23- Example of a ROC curve.

I.8.6 Cumulative Match Characteristic (CMC): Used in biometric identification to summarize the identification rate in different range values [19] (see Figure I.24).

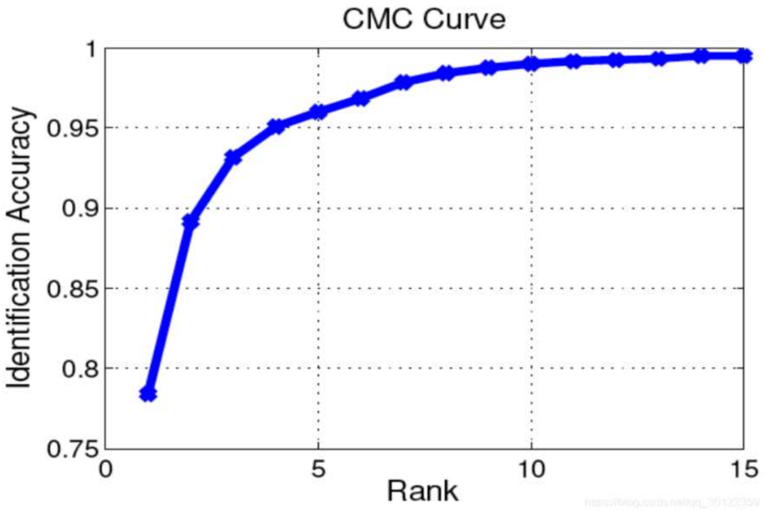


Figure I.24- Example of a CMC curve.

I.9 Conclusion

In this chapter, we have cited the main definitions of biometrics, performance, Uni-modal and multimodal systems and fusion. In the next chapter, we will focus on deep learning convolutional neural network.

Chapter II

Deep Learning and Transfer Learning

II. 1 Introduction

The term “Artificial Intelligence” or “AI” was originally coined in the 1950s as a simple theory of human intelligence being displayed by machines. In today’s era of rapid technological progress and massive increases in voluminous data sets (“Big Data”), AI has transitioned from absolute theory to concrete application on an unprecedented scale [21]. The AI is a system that generates adaptive behaviors to achieve complex goals in complex environments [22]. AI has become essentially ingrained in many aspects of our society and often operates reticently in the background of our personal electronic devices [21].

Considered a subset of AI, Machine Learning or “ML” is one of today's fastest growing technical fields, located at the intersection of computer science and statistics, as well as at the heart of artificial intelligence and data science. The development of new learning algorithms and theory, as well as the constantly increasing data volume, and low-cost computing, has fueled the recent advances in machine learning. Within AI, machine learning has emerged as the preferred method for developing practical software for computer vision, speech recognition, natural language processing, robot control, and other applications [23].

Deep Learning or “DL” in its turn is a subfield of ML that is concerned with algorithms inspired by the structure and function of the brain which can learn several levels of representation in order to model complex relationships between data. DL is also known as deep neural learning or deep neural network [24]. In this chapter we will present the concepts related to the deep learning (see Figure II.1).

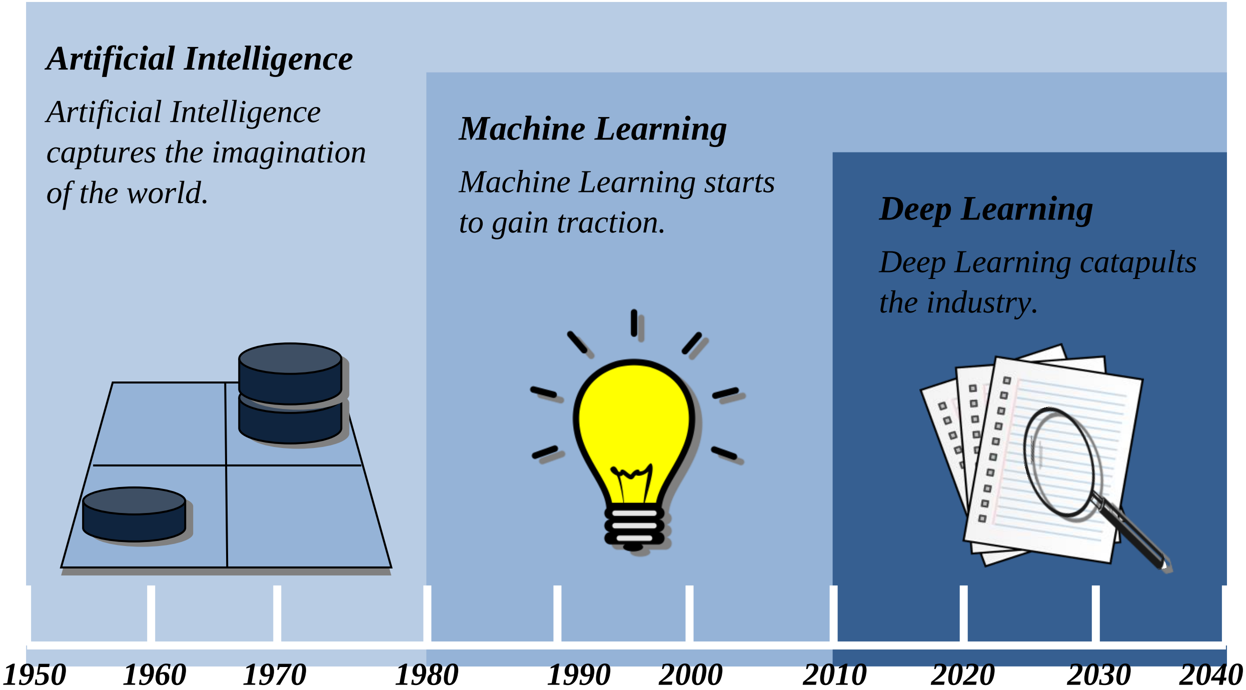


Figure II.1 - Relationship between AI, ML, and DL.

II.2 Machine Learning

Machine Learning is a branch of AI that gives machines the ability to “learn” without being “explicitly programmed” by employing a variety of techniques that can “learn” from and “predict” data. ML displays the experiential “learning” associated with human intelligence, while also having the capacity to learn and improve its analyses using computational algorithms. Large sets of data inputs and outputs are used by these algorithms to recognize patterns and effectively “learn” in order to train the machine to make independent recommendations or decisions. After enough repetitions and modification of the algorithm, the machine can take an input and predict an output. Outputs are then compared to a set of known outcomes to assess the algorithm's accuracy, which is then iteratively adjusted to master the ability to predict future outcomes [21]. Conceptually, ML algorithms can be thought of as searching through a large space of candidate programs, guided by training experience, to find a program that optimizes the performance metric, it is simple to predict future values once this program has been established [23] (see Figure II.2).

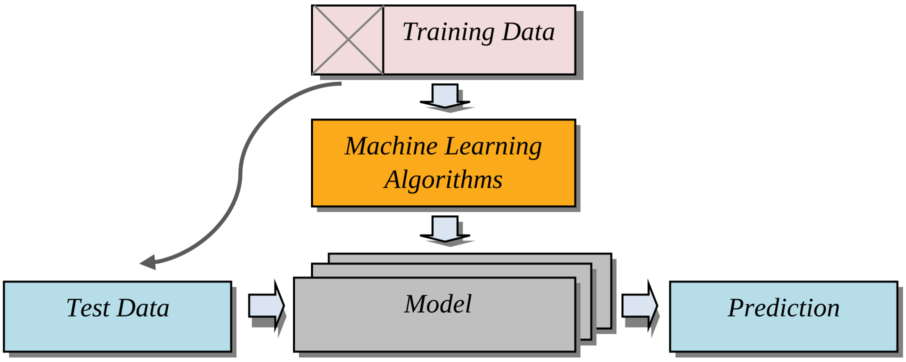


Figure II.2 - Typical ML process.

II.3 Deep learning

Deep Learning is a sub-category of ML that has aided major advancements in artificial intelligence in recent years. DL's advancement was made feasible due to machines' power increases and the development of large databases. The DL consists of building blocks known as "Artificial Neural Networks", made up of thousands of units that each performs small, simple operations. A DL model has the ability to extract characteristics from raw data through multiple layers of processing made up of multiple linear and non-linear transformations and learn about these characteristics that progressively detect features bit by bit through each layer with minimal human intervention. In initial layers it learns low-level features and as it moves up the hierarchy it starts to learn a more abstract representation of the data [24] (see Figure II.3).

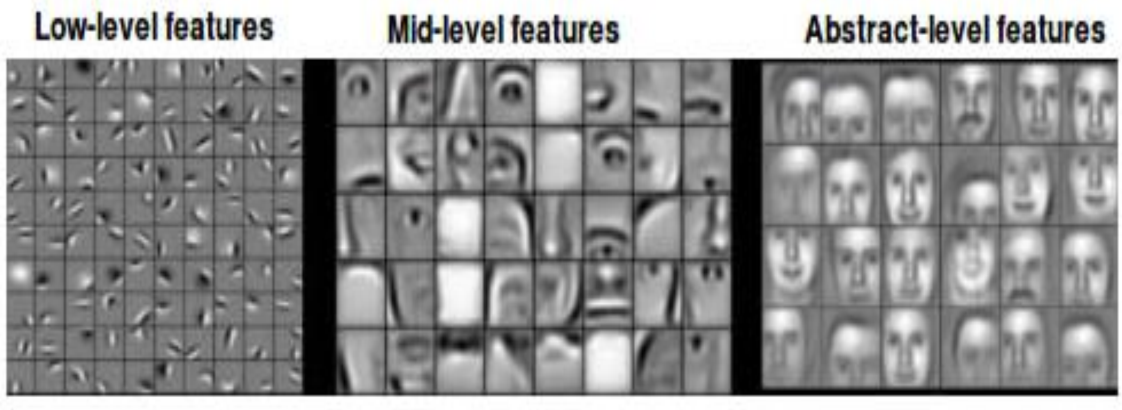


Figure II.3 - Layer-wise features learned by deep learning network.

The term "Deep" typically refers to the number of hidden layers in a neural network. Classic neural networks have only two or three hidden layers, while deep neural networks may have a hundred or more, indicating that they are built on a deep cascade of layers. DL is particularly suitable for image recognition, which is important for solving problems such as facial recognition, motion detection and many advanced driver assistance technologies such as autonomous driving, lane detection, pedestrian detection and autonomous parking (see Figure II.4).

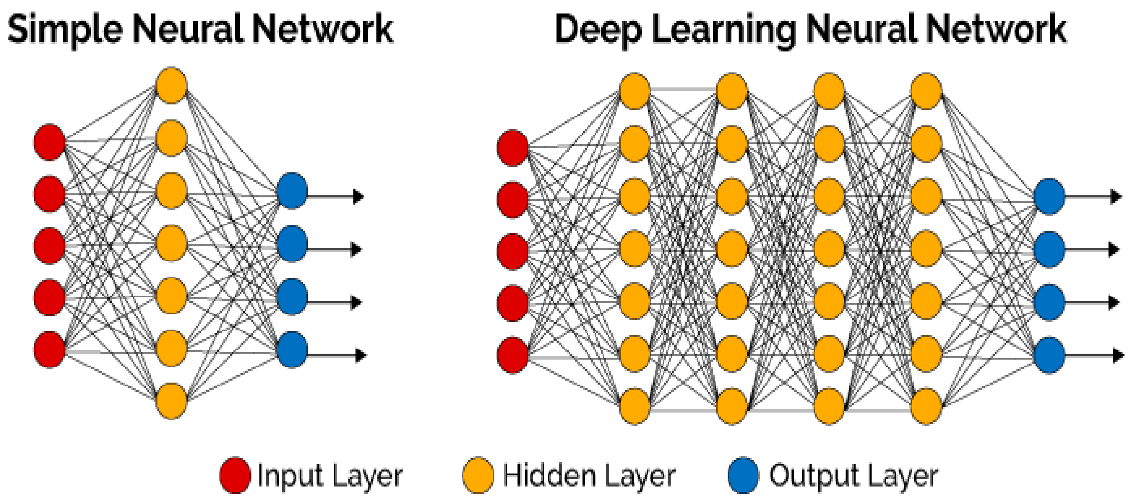


Figure II.4 - Difference between simple Neural Network and DLNN.

II.4 Difference between deep learning and machine learning

The main difference between DL and ML is due to the way data is presented in the system. ML algorithms almost always require “labeled/ structured” data, while DL networks rely on layers of ANNs. ML algorithms are designed to “learn” to act by understanding labeled data and then use it to produce new results with more datasets based on what it has learned while being trained. However, when the result is incorrect, there is a need to “teach them”. On the other hand DL networks do not require human intervention, as multilevel layers in neural networks place data in a hierarchy of different concepts, which ultimately learn from their own mistakes, and make intuitive and intelligent decisions. However, even they can be wrong if the data quality is not good enough. Data decides everything. It is the quality of the data that ultimately determines the quality of the result.

II.5 Why deep learning?

The development of DL was motivated in part by the failure of “traditional ML algorithms” to solve some major AI problems such as speech recognition and object recognition. We could only appreciate the real potential of DL once larger quantities of data became available and computing machines got more powerful. The most noticeable distinction between DL and traditional ML algorithms is that it adapts well, the greater the amount of data provided the better the performance of a DL algorithm is. Unlike many classic ML algorithms which have an upper bound on the amount of data they can receive, DL models do not have such limitations (theoretically) and they even went so far as to exceed human performance in areas such as image processing (see Figure II.5).

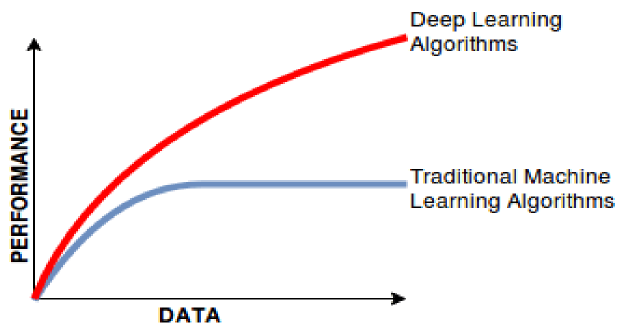


Figure II.5 - Performance of DL algorithms and traditional ML algorithms depending on the amount of data.

Another difference which is the feature extraction step, in traditional ML algorithms the feature extraction is done manually in order to reduce the complexity of the data and make patterns more visible to learning algorithms to perform well, it is a difficult and time consuming step and requires a specialist in the subject whereas in DL, the raw data is fed directly in a "deep neural network" that learns features automatically. DL often requires hundreds of thousands or millions of images for the best results. It is also very computationally intensive and requires a high performance GPU [25] (see Figure II.6).

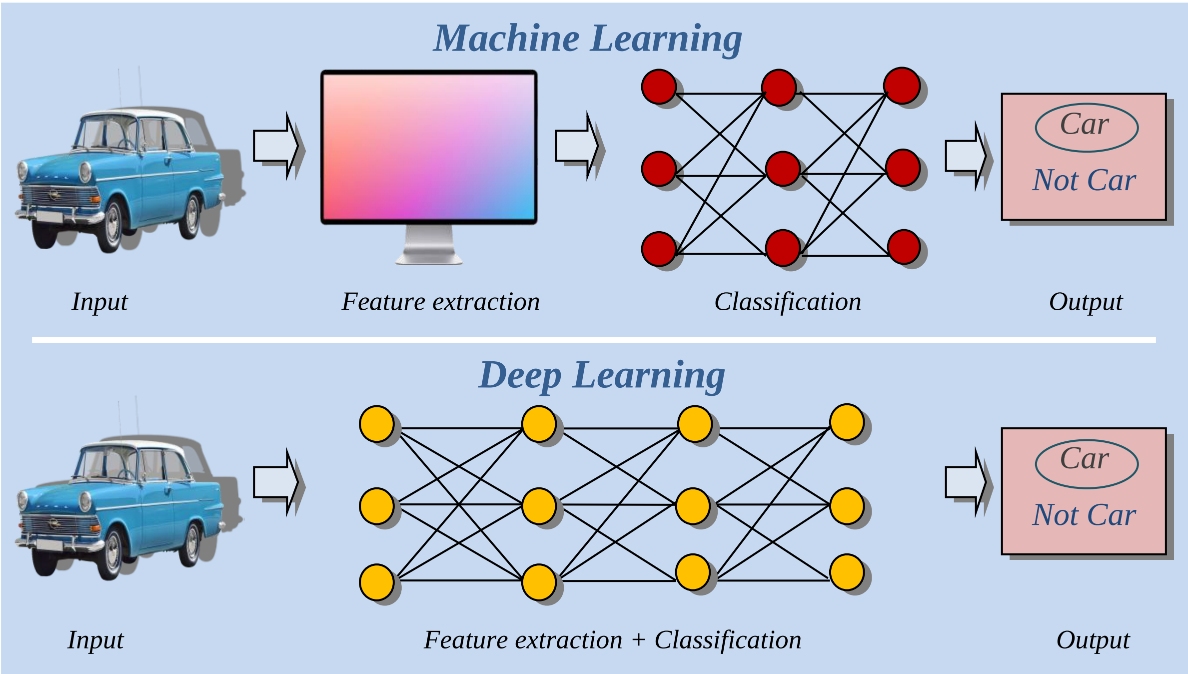


Figure II.6 - Difference between ML and DL process.

II.6 Artificial Neural Networks

Artificial Neural Networks or “ANN”s have been developed as computational models inspired by the function and structure of the biological nervous systems in the human brain [26]. ANNs are composed of a large number of interconnected processing elements called “artificial neurons”, or simply “neurons” which function in a highly parallel manner [26, 27]. These neurons analyze data patterns and learn to classify them [26]. An ANN usually consists of three types of neuron layers; input, hidden, and output layers [25]. The input layer receives various forms and structures of information, and the output layer predicts the final output, in between exist the hidden layers which perform most of the computations required by the network. Inputs are fed to each neuron of the first layer, neurons of one layer are connected to neurons of the next layer through channels, each of these channels is assigned with a numerical value known as “weight”, the inputs are multiplied to the corresponding weights and their “sum” is sent as input to the neurons in the first hidden layer, each of those neurons is also associated with a numerical value called the “bias” which is then added to the input sum, the resulted value is then passed through a non linear threshold function called “activation function” the result of this function determines if the particular neuron will get activated or not. An activated neuron transmits data to the neurons of the next layer over the channels, in this manner the data is propagated through the network, this is called “forward propagation”, in the output layer the neuron with the highest value “fires” and determines the output these values are basically a probability [27, 28]. The network compares its actual output produced with what it was meant to produce (“desired output”). The difference between both outcomes is adjusted using “back propagation” which is an abbreviation to “backward propagation”, this means that the network works “backward”, going from the output layer to the input layer to adjust the weight of the channels between the layers until the difference between the actual and desired outcome produces the lowest possible error [28].

II.7 Architectures of deep learning

There are a large number of deep architecture variables. Most of them are derived from certain original architectures. We will choose the Convolutional Neural Networks or “CNN”s in the remainder of this work.

II.7.1 Convolutional Neural Networks

Convolutional Neural Network (“ConvNet”) or CNN, it is a deep neural network that specializes in image recognition. This technique shows how important it is to improve deep layers for information processing (images). In fact, CNN is an old technique that was developed in the 1980s and 1990s. However, it was forgotten for some time because it was impractical for real-world applications with complicated images. Since 2012, when it was radically restarted, ConvNet has conquered most areas of image processing and is growing rapidly [29]. CNN is the class of Deep Neural Networks most commonly used for image processing tasks. You can achieve efficient representations of the original image so that visual patterns can be seen directly from raw pixels with little or no preprocessing. ConvNets are constructed with processing layers namely, Convolutional, pooling and fully connected or Dense. In CNN, each layer takes information from the previous layer and processes it to the next layer in particular CNN works best for image data, basically a large amount of image data. The design and arrangement of the CNN layers is determined by the model designer. The design depends on the type and complexity of the problem being solved, as well as the expected outcome of the network. The output of a convNet can be a trailing probability of a sample belonging to multiple classes, or differentiators that can be used with any other classifier.

II.7.2 Architecture of CNN

ConvNet is not just a Deep Neural Network with many hidden layers, it is a deep network that mimics the way the brain's visual cortex processes and recognizes images. As a result, even neural network experts often struggle to understand this concept when they first meet. This distinguishes ConvNet in concept and function from previous neural networks. This section briefly introduces the basic architecture of ConvNet. Basically, image recognition is a classification. For example, recognizing whether the image in an image is a cat or a dog corresponds to classifying the image as a kind of cat or dog. The same applies to letter recognition; recognizing the letter in an image is equivalent to classifying the image into one of the letter classes. Therefore, the ConvNet output layer typically uses the Multiclass Classification Neural Network. However, the direct use of the original images for image recognition leads to poor results regardless of the recognition achieved. For this reason, various techniques have been developed to extract features from images .Before ConvNet, the feature extractor was developed by experts in certain areas. As a result, it took a lot of time and money while achieving inconsistent levels of performance. These feature extractors were independent of machine learning [29] (see Figure II.7).

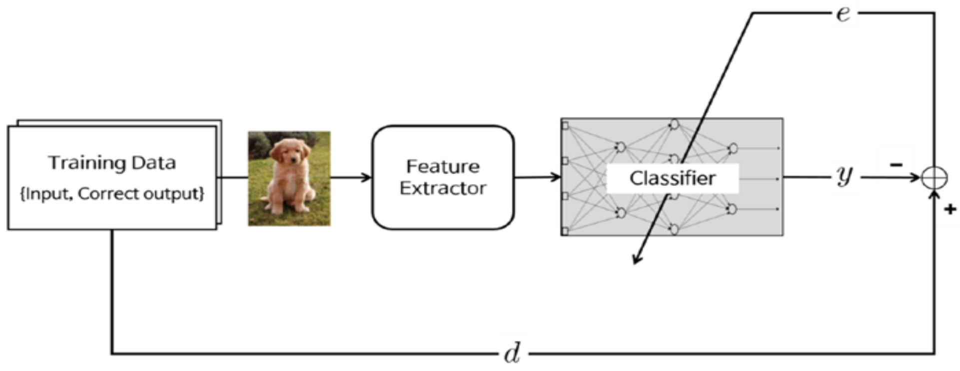


Figure II.7-Feature extractors used to be independent of ML.

ConvNet includes the feature extractor in the training process instead of designing it manually. The ConvNet feature extractor consists of special types of neural networks, the weighting of which is determined by the training process. That ConvNet turned the design of manual feature extraction into an automated process is the key feature and benefit [29] (see Figure II.8).

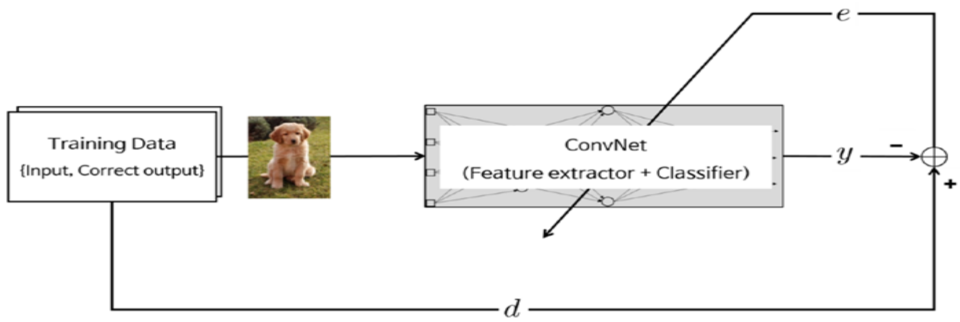


Figure II.8- CNN's feature extractor is composed of special kinds of neural networks.

II.7.3 CNNs parts

CNNs are divided into two parts:

II.7.3.1 Hidden layers/Feature extraction part: In this part, the network performs a series of convolution and grouping operations in which the features are recognized. If you had a picture of a zebra, this would be the part by which the web would recognize its stripes, two ears, and four legs.

II.7.3.2 Classification part: Here the fully connected planes serve as classifiers in addition to these extracted features. You assign a probability that the object in the image will match what the algorithm predicts (see Figure II.9).

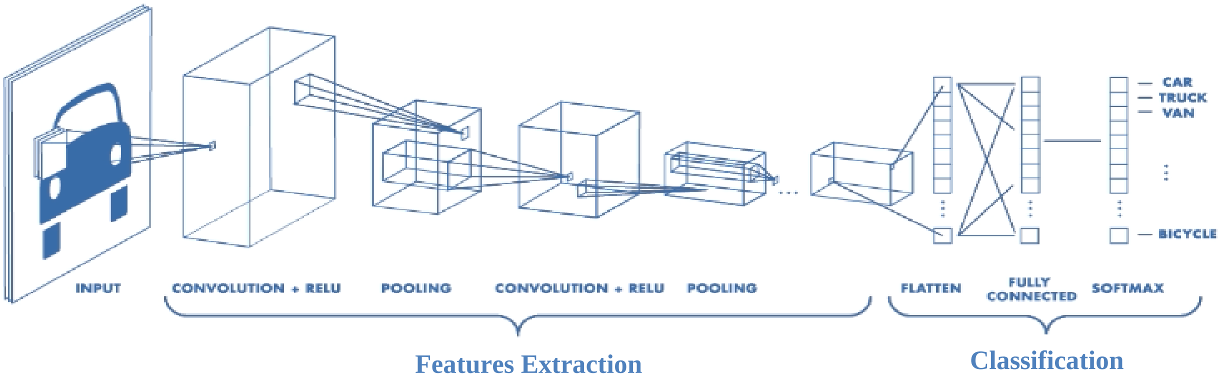


Figure II.9- CNNs parts.

II.7.4 CNN Layers

II.7.4.1 Convolution Layer

In image processing, images are read as pixels and displayed as a matrix of (N×N×3) - (height time's width time's depth). The colored images use three channels (RGB) so we have a depth of 3. The convolution filter in the base CNNs is a Generalized Linear Model "GLM" for the underlying local field of view. This works well for abstraction when the latent concept instances are linearly separable. The main processing component of this layer is a filter or mask, which is an array of weights. This mask is applied to a specific area of the feature map. The convolutional layer uses a number of filters that can be learned. A filter is used to detect the presence of certain features or patterns in the original image (input image). It is usually expressed as a matrix (MxMx3) with a smaller dimension but the same depth as the input file [30] (see Figure II.10).

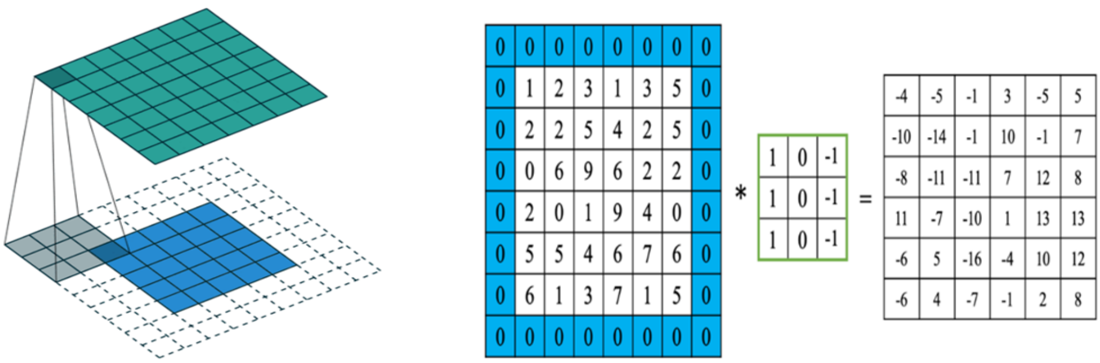


Figure II.10- Convolutional layer.

As we can see in the picture, the matrix can be configured to recognize the edges of the image. These matrices are also called filters because they act like conventional filters in image processing. In the convolution neural network, however, these filters are initialized, followed by the shape filters of the training method that are more suitable for the given task [31] (see Figure II.11).

Operation	*Filter*	*Convolved Image*
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (nourmalized)	$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
Gaussian blue (approximation)	$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	

Figure II.11- Effects of different convolution matrix.

■ **Hyper parameters of the folding layer**

• **Number of filters:** It is also necessary to determine the number of filters or neurons for each convolution layer. This hyper parameter is crucial to enable the transmission of information into the deeper layers, since too small a number of filters can lead to a considerable loss of information in this particular layer [32].

• **Kernel size:** We also have to define the size of the convolution or the kernel for each convolution layer. For the sake of consistency, all kernels in a layer are the same size. However, the optimal size depends on the problem and the structure of the rest of the network. This parameter is crucial so that the network can learn the patterns on the appropriate scales [32].

• **Depth:** The network depth is defined as the number of layers in the network. This has serious implications for the algorithm's ability to learn the task properly, as too shallow a grid cannot learn the complex patterns that the image categories represent, and too deep a grid has too many parameters and is easy to adjust [32].

• **Padding:** One of the disadvantages of the folding step is the loss of information that may be on the edge of the image. Since they are only captured when the filter is moved, they never have a chance to be seen. A very simple but effective method of solving the problem is to use zero padding. The other benefit of zero padding is the management of the output size [31] (see Figure II.12).

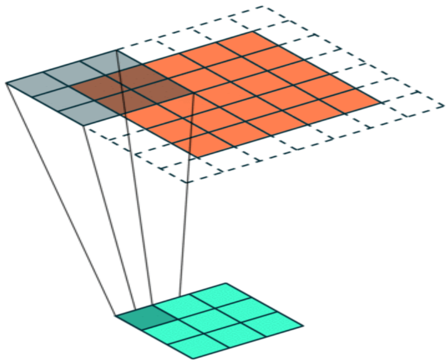


Figure II.12- Example of Padding.

• **Stride**: In fact, CNN has more options that offer numerous ways to lower the settings even further while reducing some of the side effects. One of those options is the step. We simply assume that when looking at the regions, the node of the next layer has a lot of overlap with its neighbours. We can manipulate the overlay by controlling step [31] (see Figure II.13).

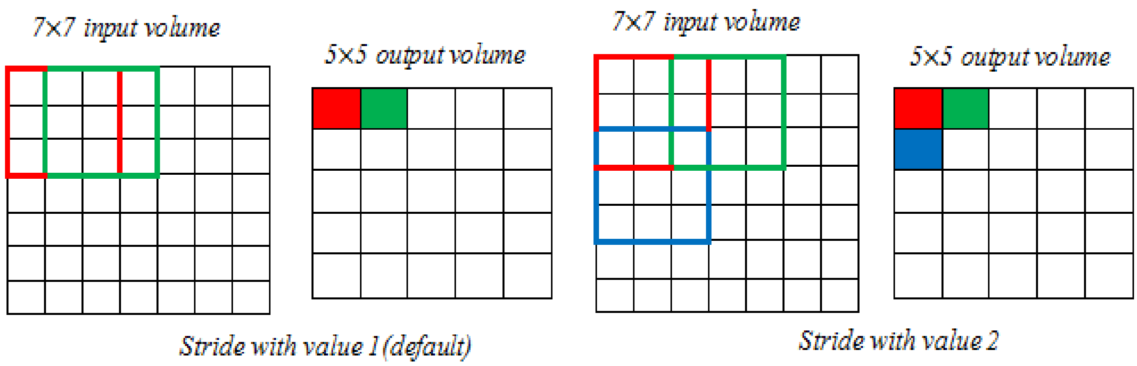


Figure II.13- Example of Stride.

II.7.4.2 Pooling Layer

The grouping layer reduces the size of the image by combining neighboring pixels of a particular image area into a single representative value. Bundling is a typical technique that many other image processing schemes have already used. To do the operations at the grouping level, we need to figure out how to select the grouping pixels in the image and set the representative value. Adjacent pixels are usually selected from the square matrix, and the number of pixels combined differs from problem to problem. The representative value is generally defined as the average or maximum of the selected pixels [29]. There are mainly two types of grouping: maximum and average grouping:

• **Max pooling:** as the name suggests, get the most out of a swimming pool. In fact, it does this with filters sliding through the entrance, and at each step the maximum parameter is removed and the rest is discarded [30].

• **Mean Pooling:** It is identical to the maximum grouping, except that the windows are averaged instead of choosing the maximum value (see Figure II.14).

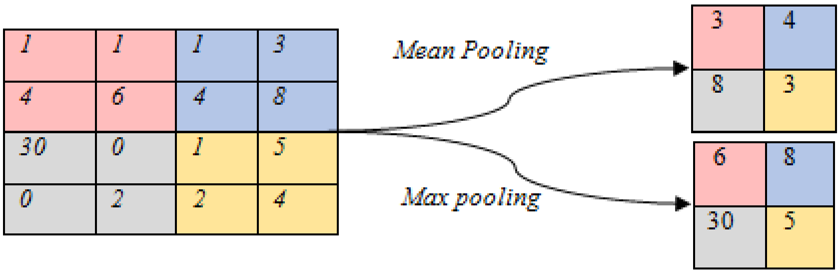


Figure II.14- Max and Mean pooling.

II.7.4.3 Fully Connected Layer

The last layers of the network are fully connected, which means that the neurons of the previous layers are connected to each neuron of the later layers. This mimics high-level reasoning that takes into account all possible paths from entry to exit. Linear or fully connected mathematically, a linear layer can be viewed as a function that applies a linear transform to an "I dimensional vector input" and produces an "O dimensional l vector". Usually the layer has a bias parameter. In this layer, neurons have a complete connection with all activations of the previous layers. Therefore, their activations can be calculated with a matrix multiplication followed by a polarization shift. This is the final phase of a CNN network.

The convolutional neural network is actually made up of hidden layers and fully connected layers [30] (see Figure II.15).

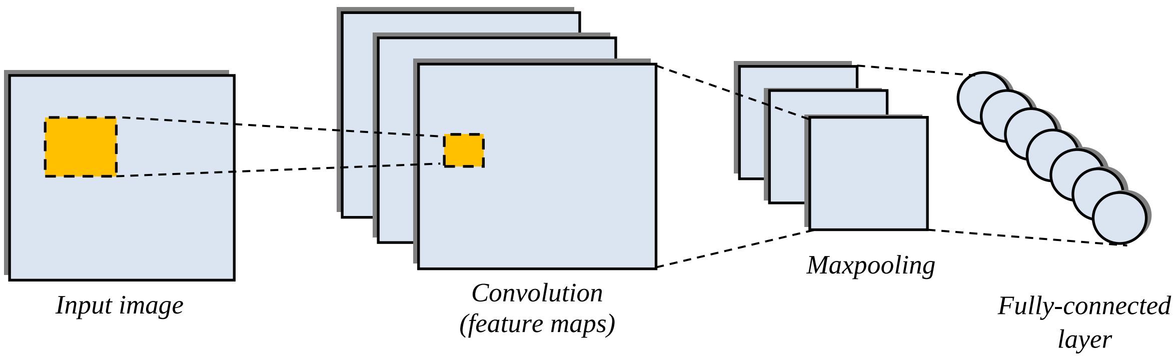


Figure II.15- Fully Connected Layer of CNN.

II.7.5 Activation functions

Activation functions are functions that are used in neural networks to calculate the weighted sum of inputs and distortions that are used to determine whether or not a neuron can fire. It manipulates the data represented by gradient processing, usually gradient descent, and then generates an output to the neural network that contains the parameters in the data. These activation functions are often referred to as the transfer function in some publications.

The activation function can be linear or non-linear depending on the function it represents and is used to control the outputs of the neural networks in various areas, from object recognition and classification to speech recognition, segmentation, understanding and description of the Scene, the machine. Translation testing for language systems, cancer detection systems, fingerprint recognition, weather forecasting, self-driving cars and other areas, to name a few, the first research results categorically confirm that an appropriate choice of function activation improves the results in the computerized neural network [33].

II.7.5.1 Rectification Linear Unit (ReLU)

ReLU stands for Rectified Sheath Unit and is a non-linear activation function that is widely used in neural networks. The advantage of using the ReLU function is that not all neurons fire at the same time. This implies that a neuron is only switched off when the output of the linear transformation is zero. It can be mathematically defined as: f(x) = max (0, x)**.** Neurons are not activated at the same time, but multiple neurons are activated at the same time. In some cases the gradient value is zero, so the weights and distortions are not updated during the back propagation step when training the neural network [34] (see Figure II.16).

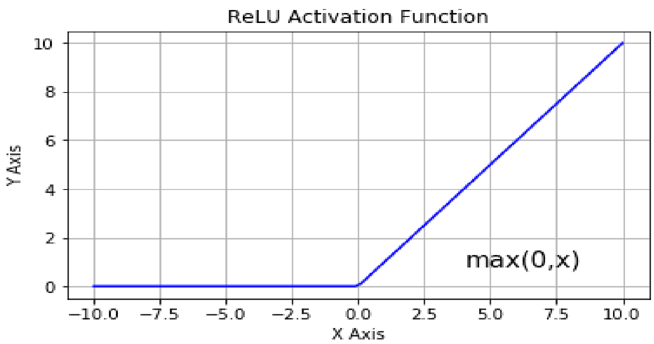


Figure II.16- ReLU Activation function plot.

II.7.5.2 Sigmoid function

This is the most common activation function because it is a non-linear function. The sigmoid function transforms values between 0 and 1. It can be defined as:

$$f(x) = \frac{1}{\mathrm{e}^{-\mathrm{x}}} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(II.1)}$$

The sigmoid function is continuously differentiable and an S-shaped function. The derivative of the function is: f'(x) = 1-sigmoid(x) ……………………………… (II.2)

In addition, the sigmoid function is not symmetric about zero, which means that the signs of all output values of the neurons are the same. This problem can be ameliorated by scaling the sigmoid function [34] (see Figure II.17).

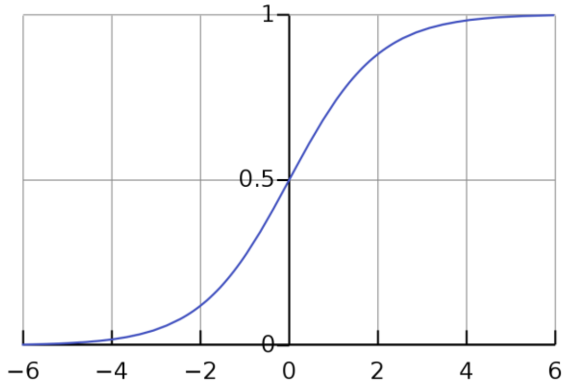


Figure II.17- sigmoid function.

II.7.5.3 Softmax activation function

The softmax function is a combination of several sigmoid functions. Since we know that a sigmoid function returns values between 0 and 1, these can be treated as probabilities of data points of a particular class. The Softmax function, unlike the sigmoid functions used for binary classification, can be used for classification problems with multiple classes. The function returns the probability for each data point of all individual classes. It can be expressed as: $\sigma(Z)j = \frac{e^{zj}}{\sum_{k=1}^{k} e^{zk}}$ $\text{forj} = 1, \ldots\ldots, k$ ……………………………… (II.3)

When we build a network or a model for classifying multiple classes, the output layer of the network will have the same number of neurons as the number of classes in the target [34].

II.8 Putting it all together

CNN is made up of various layers of convolution and pooling, as you can see. Let's take a look at the network:

• The first convolutional layer receives an image as input. An activation map is created from the convoluted output. The convolution layer's filters extract relevant features from the input image for further processing.

• Each filter must have a unique feature to help in accurate class prediction. We use same padding (zero padding) if we need to keep the image's size, otherwise, valid padding is used because it helps to minimize the number of features.

• After that, pooling layers are applied to reduce the number of parameters even more.
• Before making a prediction, several convolution and pooling layers are applied. The use of a convolutional layer aids in the extraction of features. When compared to a shallow network, where the features extracted are more general, as we go deeper in the network, more specific features are extracted.
• As previously mentioned, the output layer in a CNN is a completely connected layer that attenuates and sends the input from the other layers in order to turn the output into the desired number of groups.
• After that, the output layer generates the output, which is then compared to the output layer for error generation. To compute the mean square loss, a loss function is specified in the fully connected output layer. The error gradient is then estimated.
• Back propagation is then used to correct the filter (weights) and bias values.
• A single forward and backward pass complete one training period [35].

II.9 Transfer Learning

The aim of Neural Network is to use a learning-based approach inspired by human brain structure and processes to solve complex non-linear problems. In transfer learning, a deep neural network is regarded as an intelligent feature extraction module that provides great versatility in extracting high-level features [36]. Machine learning with an external source of knowledge in addition to the traditional training data is known as transfer learning, information gleaned from a single or multiple related tasks. The aim of transfer learning is to use information from the source task to enhance learning in the target task. There are three popular ways in which transfer can help students learn better. The first is the initial success achieved in the target task using only transferred information before any further learning, as opposed to an ignorant agent's initial performance. The second factor is the time it takes to completely understand the target task using transferred information versus learning it from scratch. The third factor is the difference between the final output level achieved in the goal task and the final level without transition. Transfer learning is significantly faster and easier than learning from beginning because the premade network has already learnt a rich set of features that can be used to a variety of other related jobs [37].

II.10 Alex Net network

Alex Net, which was proposed by Alex Krizhevsky et al., is made up of a series of cascaded phases, including convolution layers, pooling layers, rectified linear unit (ReLU) layers, and fully connected layers. Alex Net is a CNN architecture that consists of five convolutional layers and two fully connected layers. This was the first CNN architecture to demonstrate performance on the ImageNet classification challenge. There are 61 million parameters in this network. Alex Net is used in conjunction with batch normalization layers [38] (see Figure II.18).

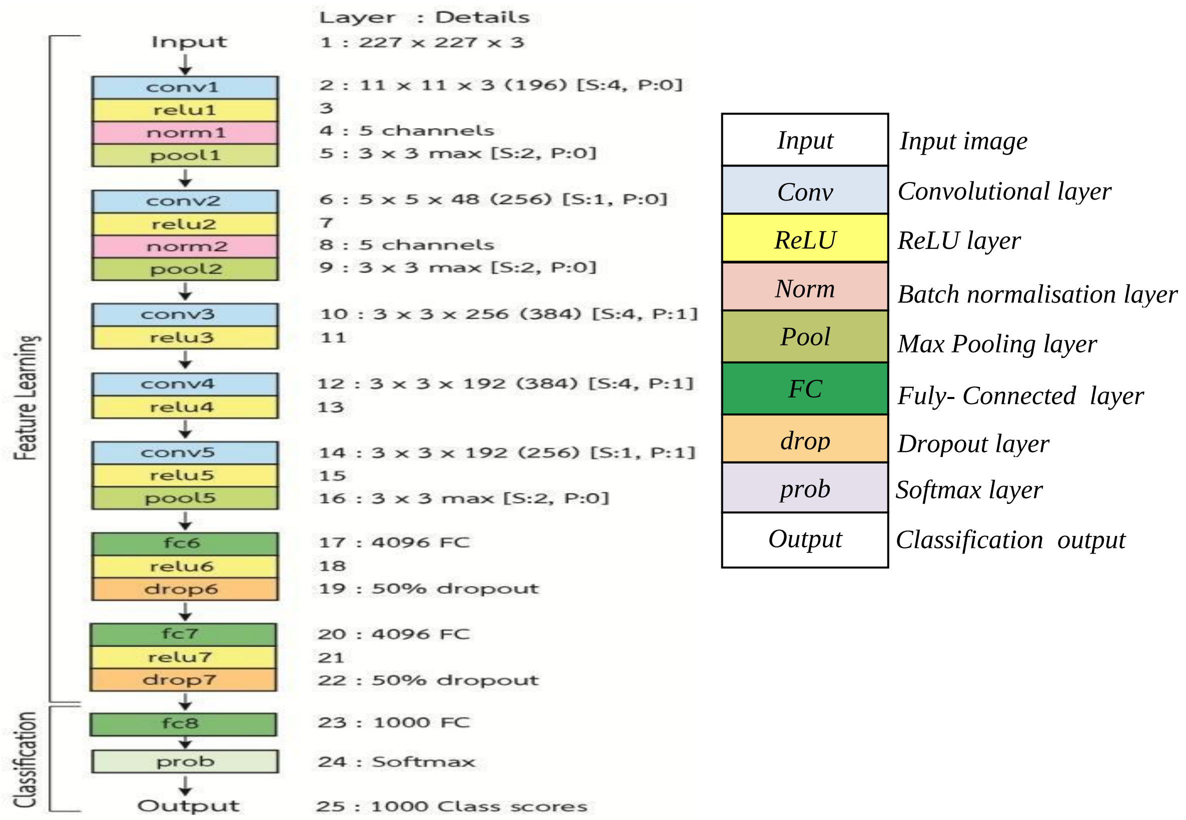


Figure II.18- Alex Net architecture.

II.11 Conclusion

In this chapter, we explored one of the world's greatest mysteries and depths of technology by walking through the path of deep learning to learn what it is and how it works, which varies dramatically from conventional deep learning. Every path has steps, as the first driving steps for this were the wave of artificial intelligence, whose creation led to the violation of machine learning veins and destabilization of them to include the deeper science, which is called deep learning. This is what took the place of the friendly encounter between us and the convolutional neural networks that were prevalent. We learned some fundamentals by deducing its detailed structure, as well as the layers of CNN, how it functions, and how to advance to the transfer learning process. In the next chapter, we will discuss the results and discussions.

Chapter III

Results and Discussion

III.1 Introduction

This chapter brings together the experimental results of recognizing palmprint images and then the identification of people, achieved by the transfer learning algorithm (AlexNet) to a database that contains several images of the palm prints from several people. The goal is to design a unimodal and a multimodal biometric identification system for individuals by recognizing palmprint based on transfer learning which is essentially used to extract the characteristics of images and their classification.

III.2 Palmprint and Palm vein

III.2.1 Palmprint

Palm print is a relatively new physiological biometric, serves as a reliable human identifier because the palmprint patterns are not duplicated in other people, including identical twins. Palm is the inner surface of the hand from the wrist to the root of fingers. A palmprint is an impression of the lines and wrinkles of the palm of a hand on a surface [39]. The Palmprint has a much larger surface area than a finger tip but is covered with the same type of skin, and contains a wealth of information that is useful for recognition [40]. Palmprints are rich in features such as principal lines, wrinkles, ridges, geometry, folds, delta point, minutiae, datum point features and texture benchmark [39] (see Figure III.1). In addition, the palmprint biometric system does not require a very high-resolution recording device because large lines and wrinkles can be seen on low resolution images (e.g., 100 dpi or less) [40].

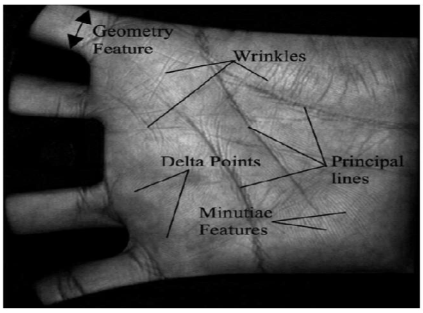
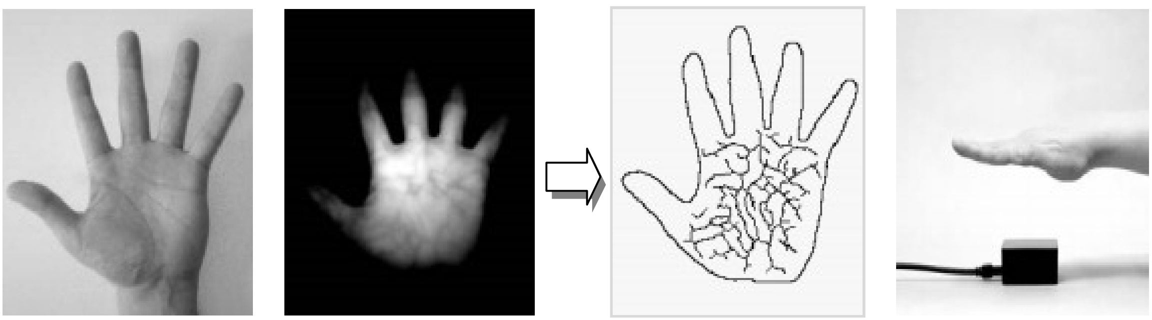


Figure III.1- Different features of a palmprint.

III.2.2 Palm vein

Palm vein authentication technology uses a person's palm vascular diagrams as personally identifiable data. The palm contains a large number of distinguishing features for personal identification, it is an ideal part of the body for this technology because such par tof the hand is easy to expose to a sensor [41]. Palm vein patterns are captured by a camera with near-infrared light. Deoxidized haemoglobin in the vessels of the veins absorbs the near-infrared light. When the infrared image is captured, only the pattern of blood vessels containing deoxidized haemoglobin is visible as a series of dark lines (see Figure III.2 (b)). Based on this property, the vein authentication device translates the black lines of the infrared ray image as the blood vessel model of the palm using image-processing technology (see Figure III.2 (c)) and then compares them with the recorded blood vessel model beforehand of the individual, through the use of vessel structure features such as directions and bifurcations, or by using the patterns themselves [42]. This technology has been spread due to its ease-of-use and affirmation that it has given users through its robust security (see Figure III.2).



(a) (b) (c) (d)

Figure III.2- Capture the Palm vein patterns.

(a): Visible ray image, (b): Infrared ray image, (c): Extracted vein pattern, (d): Palm vein sensor.

The palmprint and the palm vein can be complementary to each others. They can be acquired simultaneously using a specially designed system. For example, in PolyU database, which consists of both the RGB and the NIR illumination sources, extracts the palmprint and the palm vein features respectively from the RGB image and the NIR image almost at the same time. The two modalities are then fused for authentication performance enhancement [43].

III.3 Database of palmprint used

III.3.1 Palmprint PolyUdatabase

The palmprint image database was collected from 500people using an image capture device designed by university researchers Hong Kong Polytechnic. The images were taken at two different times, separated by a time interval of about two months. During each period,each individual had to take at least six photos of their handprints. Of more, in the second period, the light source and the camera (CCD) are adjusted by so that the images of the first and second period give the impression to have been taken by two different devices; palm print images were also taken in different light conditions to test the robustness of the recognition system. The image size is 128 × 128 with a resolution of 75 dpi. The system collects four images of four bands (Red, Green, Blue and NIR). We chose 500 different people. Each person has 12 images, 6 images arerandomly selected for training and the others for testing [44] (see Figure III.3).

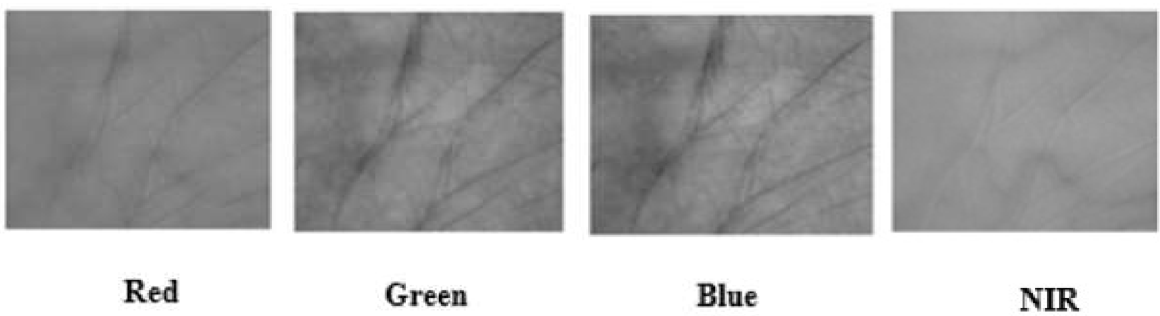


Figure III.3- Examples of ROI images from the PolyU database.

III.3.2 Palmprint CASIA database

CASIA's multispectral palmprint image database contains 7,200 images of palms taken by 100 different people using one device of multispectral imagery designed by the company itself. All pictures from palm are 8-bit grayscale JPEG files. For each hand we capture two sessions of palm pictures. The time interval between the two sessions is more than a month. In each session there are three samples. Each sample contains six palm images that are captured at the same time with six electromagnetic spectra different. The wavelengths of the illuminator corresponding to the six spectra are 460nm, 630nm, 700nm, 850nm, 940nm and white light respectively. Between two samples, we allow a certain degree of variation in the postures of the hands. Through this, we aim to increase the diversity of intra-class samples and to simulate practical use [45] (see Figure III.4).

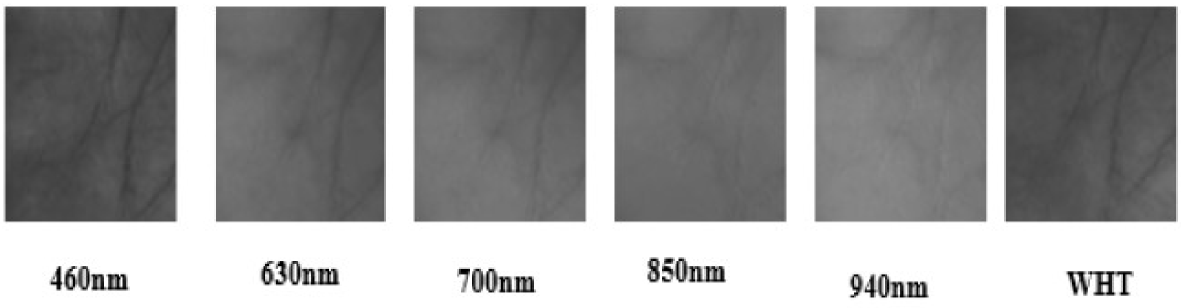


Figure III.4- Examples of ROI images from the CASIA database.

III.4 Proposed identification system

Our proposal is to make a multimodal biometric system based on a multi sensor category that consists of palmrint and palm vien as a type of recognition, and use the PolyU and CASIA databases for our experiments. As for the multimodal biometric system, we will do the fusion at the matching module level, and this is after having performed feature extraction and classification (fine tuning approach) of the uni-modal systems through the use of transfer learning technique by utilising the pre-trained AlexNet network (see Figure III.5).

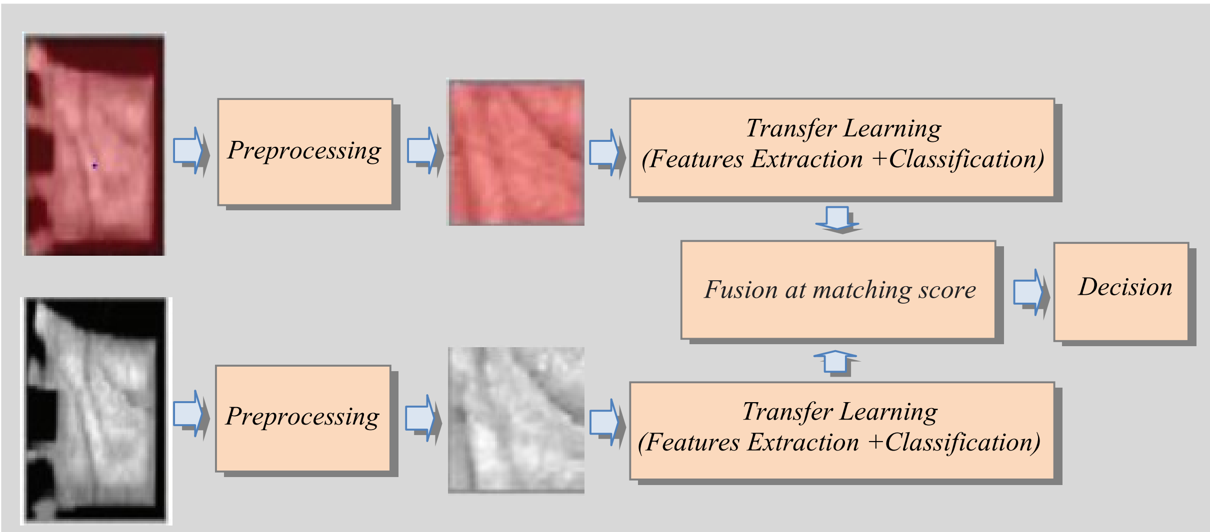


Figure III.5 - Architecture of the identification system.

III.5 Workprotocol

III.5.1 Separation of databases

• PolyU database

The palmprint image database was collected from 500 people, each person has 12 images, we selected 6 images for training [1 3 5 7 9 11] and the others ([2 4 6 8 10 12]) for testing.

• CASIA database

The palmprint image database was collected from 100 people, each person has 12 images, we selected 6 images for training [1 3 6 8 10 12] and the others ([2 4 5 7 9 11]) for testing.

III.5.2 Work environment

a) Physical environment

• Computer: HP Z8 G4 Workstation.
• Memory (RAM): 96. 00 Go.
• Processor: Intel(R) Xeon(R) Silver 4108 CPU @ 1. 80 GHz 1. 80 GHz.
• System type: 64-bit operating system, processor x64.

b) Software environments

• The software tool used by our approach is: Matlab R2021a.

III.5.3 Choosing the best parameters

Our study concentrated on a few parameters that we believe are important in our work, as well as the fact that other parameters adopted by default have delivered decent results, in order to create an effective biometric system with less complexity. The default settings of our basic model are as follows: Initial Learn Rate = 0.0001 and a fixed Validation Frequency to 100. To achieve the best parameters, the concept is to fix one parameter and focus on the other one at a time to achieve the best results. The mini-batch size and the maximum number of epochs are the parameters that we are modifying. The best parameter is the one that yielded the lowest EER.

For the PolyU database we got that the best one is EER = $2.7 \times 10^{-3}\%$ for the red spectral band, and by that the best experimental parameters are:

• The mini-batch size is 10 • The maximum number of epochs is 15

For the CASIA database we modified in the Initial Learn Rate alongside the previous two parameters and we got that the best one is EER = $2.36 \times 10^{-2}\%$ for the 460nm spectral band, and by that the best experimental parameters are:

• The mini-batch size is 64 • The maximum number of epochs is 28

After we found the best parameters, we will use them in the next step.

III.6 Experiments and results

III.6.1 Uni-modalidentification system

For the open set, we measured the equal error rate (EER) and the threshold (T_0). For the closed set, we measured the Rate Of Recognition (ROR) and the Rank of Perfect Recognition (RPR).

• Uni-modal system for the PolyU database

We compare all the identification metrics open set and closed set for the 4 bands (Red, Green, Blue, and NIR) to find the best band for the AlexNet network. Table III.1, shows the performance of the system resulting from the 4 spectral bands of open and closed identification modes.

Table III.1- Performance of a uni-modal system for AlexNet network in the PolyU database.

Spectral bands	Open set		Closed set	
	EER%	T_0	ROR%	RPR
Red	0.0027	0.0200	99.90	8
Green	0.0328	0.0030	99.80	75
Blue	0.0039	0.0530	99.83	18
NIR	0.0074	0.0090	99.80	3

From the comparison of the results of the 4 previous spectral bands in Table III.1, we can say that the best results were obtained with the Red spectral band where it offers a remarkable performance compared to the other bands. This band gives an ERR = 0.0027% with a T_0 = 0.0200 in open set and ROR = 99.90% and RPR = 8 in closed set in the PolyU database. The Blue and NIR bands give better results than the Green band which has the lowest performance where it achieved an ERR = 0.0328% and a ROR = 99.80% for the open set and the closed set respectively. Figures (III.6 and III.7) represent the performance comparison between the four spectral bands in open and closed set identification modes.

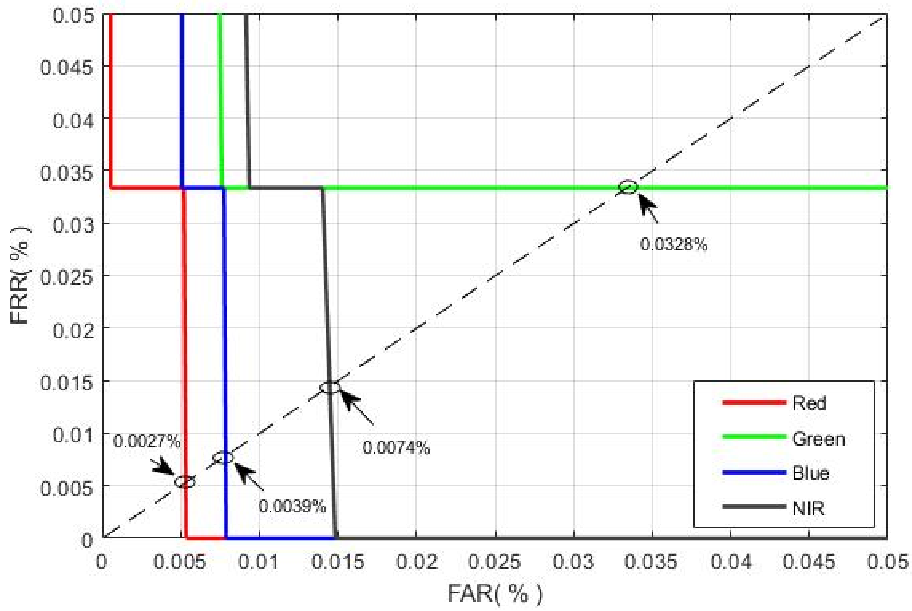


Figure III.6 - ROC curves (FRR as a function of FAR) for the four spectral bands.

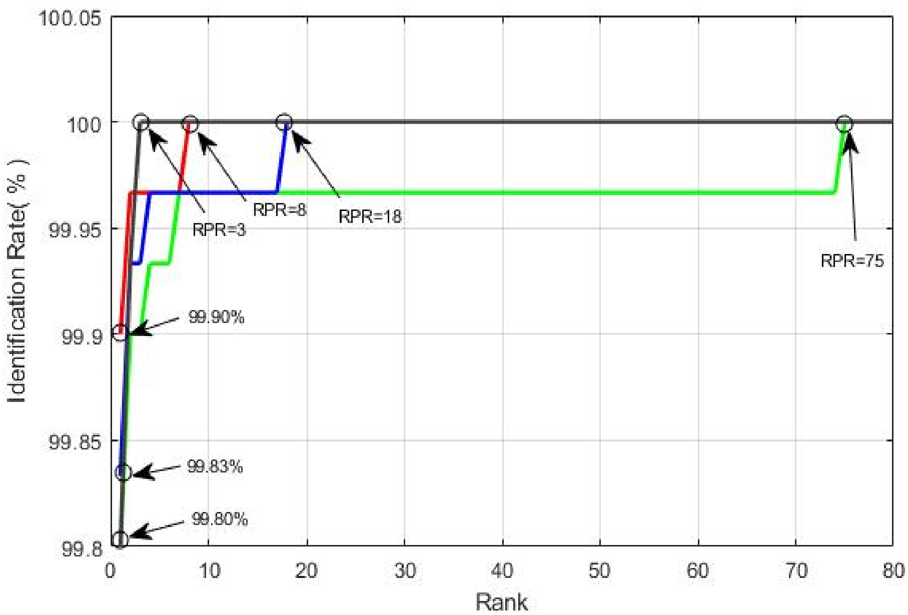


Figure III.7 - CMC curves (identification rate as a function of rank) for the spectral bands.

• Uni-modal system for the CASIA database

The following table shows the system performance resulting from the six bands of the open set and the closed set of AlexNet network in the CASIA database.

Table III.2- Performance of a uni-modal system for AlexNet network in the CASIA database.

Spectral bands	Open set		Closed set	
	EER%	T_0	ROR%	RPR
460	0.0236	0.1280	99.33	5
630	0.3316	0.0080	97.33	9
700	1.2306	0.0020	95.33	36
850	0.8291	0.0160	93.16	19
940	1.3165	0.0040	93.83	43
WHT	0.1658	0.1050	97.66	8

From the comparison of the results of the 6 previous spectral bands in Table III.2, we can say that the best results were obtained with the 460nm band where it offers a remarkable performance compared to the other bands. This band gives an ERR = 0.0236% with a T_0 = 0.1280 in open set and ROR = 99.33% and RPR = 5 in closed set in the PolyU database. The 630nm, 850 nm and WHT bands give better results than the700 nm and 940 nm bands.

The 940 nm spectral band obtained the poorest results it achieved an ERR = 1.3165% and a ROR = 93.83% in the open set and the closed set respectively. Figures (III.8 and III.9) represent the performance comparison between the six spectral bands in open and closed set identification modes.

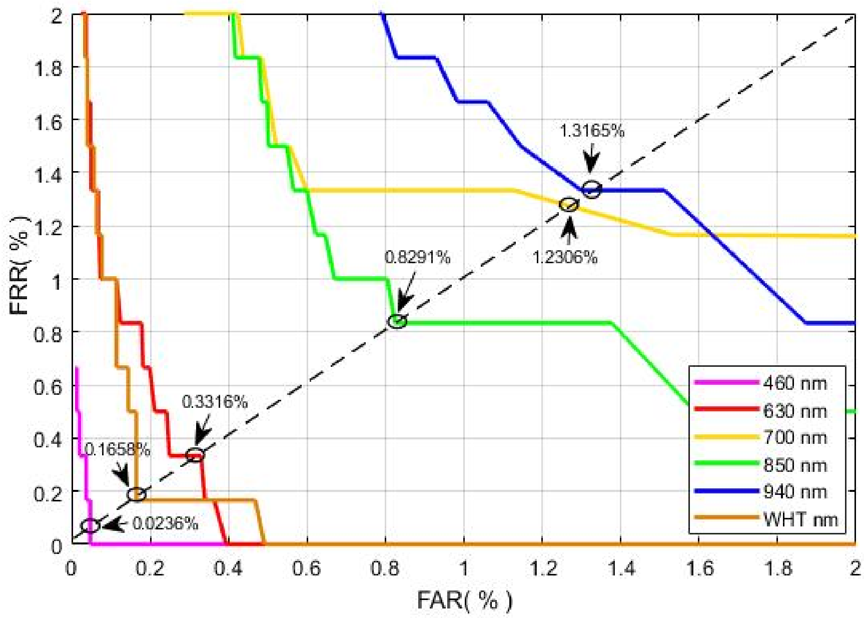


Figure III.8- ROC curves (FRR as a function of FAR) for the six spectral bands.

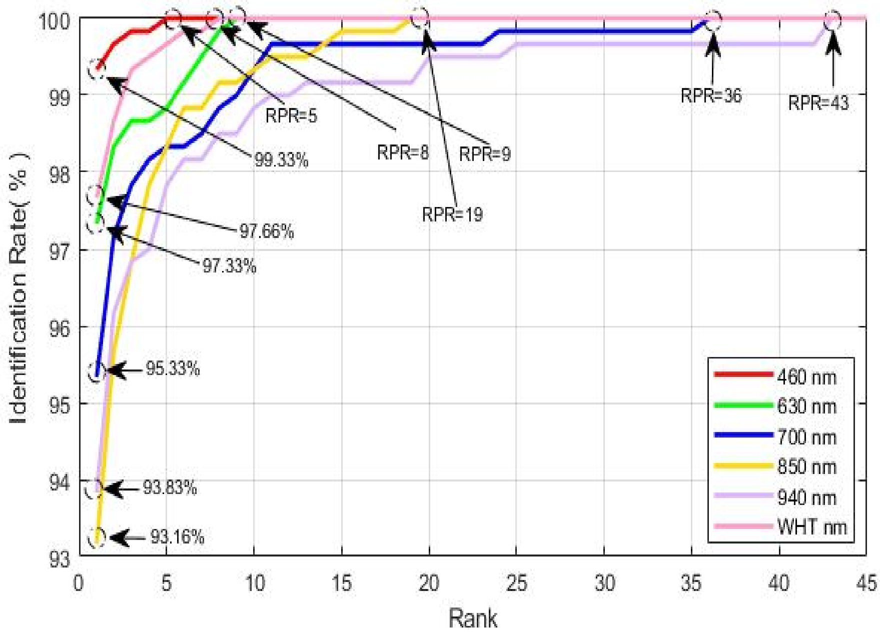


Figure III.9 - CMC curves (identification rate as a function of rank) for the six spectral bands.

III.6.2 Multimodal identification system

Multimodal biometrics refers to the identification of people using more than one biometric modality. As unimodal biometric systems can suffer from many limitations such as data noise, changes within the layer, and similarity between layers, this strategy can be used to reduce some of the problems and limitations associated with them. The goal of using a multimodal system is to improve the system's performance. The multimodal system can attain a very low EER value, which is required for an excellent biometric identification system. We merged the spectral bands in our multimodal system and experimented with the fusion using six rules: The simple sum rule (SUM), the product rule (PROD), the minimum rule (MIN), the weighted sum rule (WHT SUM), and the weighted product rule (WHT PROD).

• Multimodal system for the PolyU database

We use the multimodal system on the database (PolyU) and merge the three spectral bands of the database (Red, Green, and Blue) for a multispectral image, using the six fusion rules. The goal of fusion is to improve the system's performance.

Table III.3- Performance of a multimodal system for identification of AlexNet network in the PolyU database.

Fusion rules	Red Green Blue			
	Open set		Closed set	
	EER%	T_0	ROR%	RPR
SUM	0.0004	0.6940	99.96	5
PROD	0.0003	0.0090	99.96	5
MIN	0.0239	0.0010	99.96	65
MAX	0.0012	0.4460	99.96	6
WHT SUM	0.0006	0.5610	99.93	3
WHT PROD	0.0000	0.7280	100	1

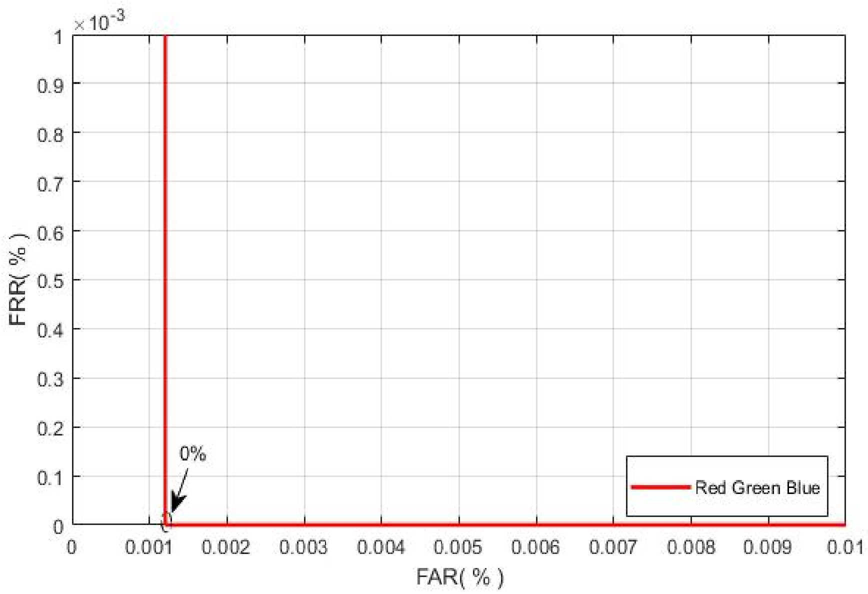


Figure III.10 - ROC curves (FRR as a function of FAR) for the three spectral bands.

Through the results of Table III.3, we see that multimodal systems are much better than uni-modal systems, and by merging the group (Red, Blue, and Green), we find that the group gives excellent results:

▪ **In the open set identification:** the table indicating that the combination (Red, Blue, and Green), can reduce the ERR to 0% (a GAR of 100%) when the case WHT PROD is used.

▪ **In the closed set identification:** a very significant improvement (100% of ROR) is observed compared to the best case in the uni-modal system. The best result of the ROR produces 100% accuracy in the case of the WHT PROD rule with (RPR = 1).

• Multimodal system for the CASIA database

We use the multimodal system on the database (CASIA) and merge the six spectral bands of the database using five of the six fusion rules only. Some spectra must be merged to increase the system's performance. As can be seen in the Tables (III.4, III.5), we limited the fusion to three groups (460nm, 630nm, 700nm) and (850nm, 940nm, WHT), and we combined all spectral bands (460nm, 630nm, 700nm, 850nm, 940nm, WHT) for a multispectral image.

Table III.4- Performance of a multimodal system for identification of AlexNet network in the CASIA database.

Fusion rules	**460- 630- 700**				**850- 940- WHT**			
	Open set		**Closed set**		**Open set**		**Closed set**	
	EER%	**T_0**	**ROR%**	**RPR**	**EER%**	**T0**	**ROR%**	**RPR**
SUM	0.0118	0.6060	99.50	2	0.0303	0.5160	99.00	4
PROD	0.0017	0.0130	99.66	3	0.0101	0.0380	99.66	4
MIN	0.1389	0.0020	99.33	9	0.1658	0.0290	99.16	16
WHT SUM	0.0152	0.1990	99.33	4	0.0152	0.4850	99.00	2
WHTPROD	0.0017	0.6810	99.66	2	0.0034	0.5620	99.83	2

Table III.5- Performance of a multimodal system of AlexNet network in the CASIA database.

Fusion rules	460- 630- 700 850- 940- WHT			
	Open set		Closed set	
	EER%	T_0	ROR%	RPR
SUM	0.0017	0.9830	99.66	2
PROD	0.0017	0.0080	97.83	2
MIN	0.0101	0.1660	95.66	7
WHT SUM	0.0000	0.6200	100	1
WHTPROD	0.0000	0.0010	100	1

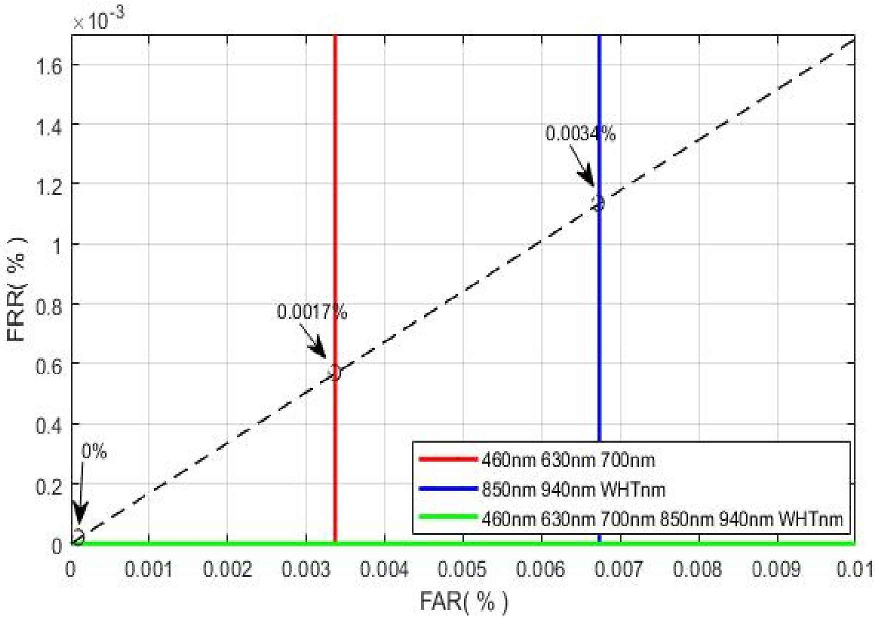


Figure III.11- ROC curves (FRR as a function of FAR) for the six spectral bands.

Through the results of Tables (III.4 and III.5), we see that multimodal systems are much better than uni-modal systems, and by comparing the three groups (460nm, 630nm, 700nm), (850nm, 940nm, WHT) and (460nm, 630nm, 700nm, 850nm, 940nm, WHT), we find that the third group gives excellent results:

▪ **In the open set identification**: the table indicating that the third combination (460nm, 630nm, 700nm, 850nm, 940nm, WHT) can reduce the ERR to 0% (a GAR of 100%) when the cases WHT SUM and WHT PROD are used. Compared to these groups, the average scores are obtained by the first and second combination, but even these combinations improve the results of uni-modal identification.

▪ **In the closed set identification:** in the third combination, a very significant improvement (100% of ROR) is observed compared to the best case in the uni-modal system. The best result of the ROR produces 100% accuracy in the case of the WHT SUM and WHT PROD rules with (RPR = 1, RPR = 1) respectively.

III.7 Conclusion

In this chapter, a person identification system based on palm prints was constructed using available biometric research. We suggested a unimodal and multimodal biometric systems based on transfer learning to extract and classify features, by employing the fine tuning approach to do this. We have witnessed a considerable and perfect improvement in the identification rate by validating these systems in databases (100 %).

General Conclusion

Identification of individuals by palmprints as a new member of the family of biometric modalities has become a very active field of research these last years. The work done so far has been based on palmprint image representation techniques for better classification. The works presented in this document fall within the context of the identification of persons from their biometric descriptors. We have proposed uni modal and multimodal biometric systems. After presenting the general concepts of biometrics, our main goal in this application was to create multimodal biometric systems.

Our main goal in this application was to create a multimodal biometric system based on the capture of multi-specters images of the hand palm (Red, Blue, Green) for the palmprint and (NIR) for the palmvein.We did our study using two multispectral palmprint databases (PolyU and CASIA), and we used the transfer learning convolutional neural network for feature extraction and classification, which is recognized for its robustness against errors and its ability to handle large databases. At the beginning, we started by comparing the spectral bands within each database while using the convolutional neural network model AlexNet, then we chose the optimal band based on the system's accuracy to get the greatest results. While utilizing the PolyU database, the best results were (EER= 0.0027 %, ROR= 99.90 %) in the RED band, and when using the CASIA database, the best results were (EER= 0.0236 %, ROR= 99.33 %) in the 460nm band. Finally, we combined the results obtained in the classification of these spectral bands at the level of the matching score using several fusion techniques (Sum, Prod, Min, Max, Weighted Sum and Weighted Prod). The outcome of the two databases (ERR = 0.00% and ROR = 100%) illustrates and confirms the efficiency of the transfer learning and the system's reliability, indicating that we have met the aim we set out to achieve at the start of this project.

As a perspective, we would like to use other methods of transfer learning with other databases such as contactless and hyperspectral palm prints, and maybe even try other modalities.

Bibliography

[1] John D. Woodward, Nicholas M. Orlans, Peter T. Higgins, Identity Assurance in the information age biometrics, McGrawHill Osborne (2003).

[2] David D. Zhang, Automated biometrics: Technologies and systems, Vol. 7, Springer Science & Business Media (2013).

[3] Claus Vielhauer, Biometric user authentication for IT security: from fundamentals to handwriting, Vol. 18, Springer Science & Business Media (2005).

[4] David Zhang, Yong Xu, Wangmeng Zuo, Discriminative learning in biometrics, Discriminative Learning in Biometrics, Springer (2016).

[5] Amine Naït-Ali, Regis Fournier, Signal and image processing for biometrics, John Wiley& Sons (2012).

[6] David Zhang, Fengxi Song, Yong Xu, Zhizhen Liang, Advanced pattern recognition technologies with applications to biometrics, IGI Global (2009).

[7] Anil K. Jain, Ajay Kumar, Biometrics of next generation: An overview, Second generation biometrics Vol. 12, Springer (2010).

[8] H. Ahmed, Identification des individus par la biométrie multimodale, Magister thesis in Simulation and artificial intelligence, University of Mohamed Boudiaf Oran (2014).

[9] Anil K. Jain, Arun A. Ross, Karthik Nandakumar, Introduction to biometrics, Springer Science & Business Media (2011).

[10] David Zhang, Guangming Lu, Lei Zhang, Advanced biometrics, Springer (2018).

[11] Anil K .Jain, Ruud Bolle, Sharath Pankanti, Biometrics: personal identification in networked society, Vol. 479, Springer Science & Business Media (2006).

[12] David D. Zhang, Palmprint authentication, Springer Science & Business Media (2004).

[13] T. Sabhanayagam, Dr. V. Prasanna Venkatesan, Dr. K. Senthamaraikannan, A comprehensive survey on various biometric systems, vol. 13, International Journal of Applied Engineering Research (2018).

[14] Anil K. Jain, Patrick Flynn, Arun A. Ross, Handbook of biometrics, Springer Science & Business Media (2007).

[15] K. Maarouf, Biometrics and Information Security for a Secure Person Identification, 3rd Cycle LMD Doctorat thesis in Communication and Signal Processing, University of Kasdi Merbah Ouargla (2019).

[16] Md. Morshedul Arefin, Md. Ekramul Hamid, A Comparative Study on Unimodal and Multimodal Biometric Recognition, Vol. 3, International Journal of Innovative Science and Modern Engineering (IJISME), Citeseer (2014).

[17] Ujwalla Gawande, Yogesh Golhar, Biometric security system: a rigorous review of unimodal and multimodal biometrics techniques, Vol. 10, International Journal of Biometrics, Inderscience Publishers (IEL) (2018).

[18] P. S .Sanjekar, J. B. Patil, , An overview of multimodal biometrics, Vol.4, Signal & Image Processing, Academy & Industry Research Collaboration Center (AIRCC) (2013).

[19] Ramadan Gad, Ayman EL-Sayed, Nawal El-Fishawy, M. Zorkany, Multi-biometric systems: a state of the art survey and research directions, Vol. 6, International Journal of Advanced Computer Science and Applications (IJACSA) (2015).

[20] Dr.R. Sridevi, Checking User Authentication by Biometric One Time Password Generation using Elliptic Curve Cryptography, Vol. 8, International Journal of Computer Science & Engineering Technology (IJCSET) (2017).

[21] J. Matthew Helm, Andrew M. Swiergosz, Heather S. Haeberle, Jaret M. Karnuta, Jonathan L. Schaffer, Viktor E. Krebs, Andrew I. Spitzer, Prem N. Ramkumar, Machine learning and artificial intelligence: Definitions, applications, and future directions, Vol. 13, Current reviews in musculoskeletal medicine, Springer (2020).

[22] D. B. Fogel, Review of computational intelligence: Imitating life, Vol. 83, Proceedings of the IEEE (1995).

[23] Michael I. Jordan, Tom M. Mitchell, Machine learning: Trends, perspectives, and prospects, Vol. 349, Science, American Association for the Advancement of Science (2015).

[24] M. Youcef, Deep Learning pour la classification des Images, Master thesis in Computer Science, Abou bekr Belkaid Tlemcen University (2017).

[25] Anil K. Jain, Ruud Bolle, Sharath Pankanti, BIOMETRICS Personal Identification in Networked Society, Kluwer Academic Publishers New York (2002).

[26] Ajith Abraham, Artificial neural networks, Handbook of measuring system design, John Wiley & Sons (2005).

[27] Daniel A. Zahner, Evangelia Micheli-Tzanakou, Artificial neural networks: definitions, methods, applications, Supervised and unsupervised pattern recognition: Feature extraction and computational intelligence, CRC Press (2017).

[28] Carlos Gershenson, Artificial neural networks for beginners, arXiv preprint cs/0308031 (2003).

[29] Phil Kim, Matlab deep learning, With machine learning, neural networks and artificial intelligence, Springer (2017).

[30] Dessalegn Mosisa, Applying a Deep Learning Approach for Wheat Rust Disease Detection, ASTU (2019).

[31] Saad Albawi, Oguz Bayat, Saad Al-Azawi, and Osman N. Ucan, Understanding of a convolutional neural network, 2017 International Conference on Engineering and Technology (ICET), IEEE (2017).

[32] Frankin Johnson, Alvaro Valderrama, Carlos Valle, Broderick Crawford, Ricardo Soto, and Ricardo Ñ anculef, Automating configuration of convolutional neural network hyperparameters using genetic algorithm, vol. 8, IEEE Access, IEEE (2020).

[33] Chigozie Enyinna Nwankpa, Winifred Ijomah, Anthony Gachagan, and Stephen Marshall, Activation functions: Comparison of trends in practice and research for deep learning, arXiv preprint arXiv: 1811.03378 (2018).

[34] Siddharth Sharma, Simone Sharma, Anidhya Athaiya, Activation functions in neural networks, vol. 6, Towards Data Science (2017).

[35] Ibtesam M. Dheir, Alaa Soliman Abu Mettleq, Abeer A. Elsharif, Samy S. Abu-Naser, Classifying nuts types using convolutional neural network, vol. 3, International Journal of Academic Information Systems Research (IJAISR) (2020).

[36] Jie Lu, Vahid Behbood, Peng Hao, Hua Zuo, Shan Xue, Guangquan Zhang, Transfer learning using computational intelligence: A survey, vol. 80, Knowledge-Based Systems, Elsevier (2015).

[37] Lisa Torrey, Jude Shavlik, Trevor Walker and Richard Maclin, Transfer learning via advice taking, Advances in Machine Learning I, Springer (2010).

[38] Mohammad Rastegari, Vicente Ordonez, Joseph Redmon, Ali Farhadi, Xnor-net: Imagenet classification using binary convolutional neural networks, European conference on computer vision, Springer (2016).

[39] K. Krishneswari, S. Arumugam, A review on palm print verification system, International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM) ISSN, Citeseer(2010).

[40] Goh Kah Ong Michael, Tee Connie, and Andrew Beng Jin Teoh, Touch-less palm print biometrics: Novel design and implementation, vol. 26, Image and vision computing, Elsevier (2008).

[41] Pallavi A. Patil, P. E. Ajmire, Survey: Human Identification Using Palm Vein Images, vol. 6, Int J Emerging Technologies in Engineering Research (2018).

[42] Masaki Watanabe, Toshio Endoh, Morito Shiohara, and Shigeru Sasaki, Palm vein authentication technology and its applications, Proceedings of the biometric consortium conference (2005).

[43] Sungchul Cho , Beom-Seok Oh , Kar-Ann Toh , and Zhiping Li Cho, Extraction and cross-matching of palm-vein and palmprint from the RGB and the NIR spectrums for identity verification, vol. 8, IEEE Access, IEEE (2019).

[44] https ://www4.comp.polyu.edu.hk/ biometrics/MultispectralPalmprint/MSP.htm.

[45] http ://www.cbsr.ia.ac.cn/english/MS PalmprintDatabases.asp.

Abstract

In recent years, biometrics has been included in all security systems with different forms: fingerprint, iris, signature, etc. Nowadays, biometrics is a solution to several problems of automatic identification of individuals. In this document, we gave a definition of the biometric system, showed its categories, the different processes that it can work in, and we also demonstrated its different composition modules and how can we evaluate its performances on several standards, furthermore we spoke about two biometric techniques, and those were Uni/Multimodal systems and revealed how is it that multimodal biometric system is better than the unimodal. We will experiment a biometric system based on palmprint and palmvein by using transfer learning which is used in deep learning applications. Transfer learning is that a preformed network has learned a rich set of characteristics that can be used as a starting point for learning a new task. In our work, the effectiveness of transfer learning was assessed on two publicly available databases **(PolyU and CASIA)** by comparing the spectral bands within each database while using one of the models of convolutional neural networks **(AlexNet)** and choosing the best band in the uni-modal system for each database. Finally, we will apply a multimodal system for each database using the fusion of the unimodal scores in order to improve the performance and get better results.

Keywords:

Biometrics.

Deep learning.

Transfer learning.

Palmprint.

Palmvein.

AlexNet.

PolyU and CASIA.

Résumé

Ces dernières années, la biométrie a été intégrée dans tous les systèmes de sécurité sous différentes formes: empreinte digitale, iris, signature, etc. De nos jours, la biométrie est une solution à plusieurs problèmes d'identification automatique des individus. Dans ce document, nous avons donné une définition du système biométrique, montré ses catégories, les différents processus dans lesquels il peut fonctionner, et nous avons également démontré ses différents modules de composition et comment évaluer ses performances sur plusieurs normes, de plus nous avons parlé de deux techniques biométriques, et celles-ci étaient des systèmes Uni/Multimodales et ont révélé comment se fait-il que le système biométrique multimodale soit meilleur que le système unimodale. Nous expérimenterons un système biométrique basé sur l'empreinte palmaire et la veine palmaire en utilisant l'apprentissage par transfert qui est utilisé dans les applications d'apprentissage en profondeur. L'apprentissage par transfert est qu'un réseau préformé a appris un riche ensemble de caractéristiques qui peuvent être utilisées comme point de départ pour apprendre une nouvelle tâche. Dans notre travail, l'efficacité de l'apprentissage par transfert a été évaluée sur deux bases de données accessibles au public **(PolyU et CASIA)** en comparant les bandes spectrales au sein de chaque base de données tout en utilisant l'un des modèles de réseaux de neurones convolutifs **(AlexNet)** et en choisissant la meilleure bande dans le système unimodale pour chaque base de données. Enfin, nous appliquerons un système multimodale pour chaque base de données en utilisant la fusion des scores unimodales afin d'améliorer les performances et d'obtenir de meilleurs résultats.

ملخص

في السنوات الأخيرة ، تم تضمين القياسات الحيوية في جميع أنظمة الأمان بأشكال مختلفة: بصمات الأصابع ، قزحية العين ، التوقيع ، إلخ. في الوقت الحاضر ، تعد القياسات الحيوية حلاً للعديد من مشاكل التعرف التلقائي على الأفراد . في هذا المستند ، قدمنا تعريفًا لنظام القياسات الحيوية ، وأظهرنا فئاته ، والعمليات المختلفة التي يمكن أن يعمل بها ، كما أوضحنا أيضًا وحدات تكوينه المختلفة وكيف يمكننا تقييم أدائه على عدة معايير ، علاوة على ذلك تحدثنا عن تقنيتي القياسات الحيوية ، وتلك كانت أنظمة (أحادية / متعددة) الوسائط, كما كشفنا كيف أن نظام القياسات الحيوية متعدد الوسائط أفضل من النظام أحادي الوسائط. سنقوم بتجربة نظام القياسات الحيوية على بصمة اليد و عروق اليد باستخدام تعلم النقل المستخدم في تطبيقات التعلم العميق. تعلم النقل هو أن شبكة سابقة التشكيل قد تعلمت مجموعة غنية من الخصائص التي يمكن استخدامها كنقطة بداية لتعلم مهمة جديدة. في عملنا تم تقييم فعالية تعلم النقل على قاعدتي بيانات متاحة للجمهور **(PolyU, CASIA)** من خلال مقارنة النطاقات الطيفية داخل كل قاعدة بيانات اثناء استخدام أحد نماذج الشبكات العصبية التلافيفية **(AlexNet)** واختيار أفضل نطاق في النظام أحادي الوسائط لكل من قاعدة البيانات. أخيرًا، سنطبق نظامًا متعدد الوسائط لكل قاعدة بيانات باستخدام دمج الدرجات أحادية الوسائط من أجل تحسين الأداء والحصول على نتائج أفضل.