UNIVERSITY KASDI MERBAH OUARGLA

Faculty of New Information Technologies and Communication
Department of electronics and telecommunication

**Dissertation Submitted in Partial Fulfillment of the Requirements for a Master's Degree**

**Domain :electronic**

**Specialty :electronic of embedded systems**

**Directed by :** Sabah Benferdia and Faiza Khenfoussi

**Theme :**

# BIO-METRIC RECOGNITION BASED ON 3D PALM PRINT

**Committee Members:**

| | | | |
|---|---|---|---|
| Mrs. Fella Charif | President | MCA | Ouargla |
| Mr. Djamel Samai | Supervisor | MCA | Ouargla |
| Miss. Salma Trablsi | Co-Supervisor | PhD student | Ouargla |
| Mr. Khaled Bensid | Examiner | MCB | Ouargla |

**College year : 2020/2021**

## DEDICATION

*To my dear mother, no matter what I do or say, I cannot thank you properly. Your affection covers me, Your mercy guides me and your presence by my side has always been the source of my strength to overcome various obstacles.*
*To my dear father, you have always been by my side to support and encourage me. To all my family and everyone who holds a place in my heart.*
*To all the professors and teachers who followed me throughout my school career and who allowed me to succeed in my studies.*

*SABAH*

*I dedicate this work first and foremost to my mother's pure soul. You are here present in my heart * to my dear father, I thank him for his support and encouragement * for the source of my strength and success, my generous family, big and small, especially my grandfather, my grandmothers, my aunt and everyone who occupies a place in my heart * without forgetting my colleague Sabah who participated with me in this work. Thank you so much* and all my friends at the end, I just wanted to write a quick note to say thank you for your support and graduation gift, I am so fortunate to have so many wonderful people in my life.*

*FAIZA*

## ACKNOWLEDGMENTS

*First of all, we would like to thank Almighty and Merciful God, who gave us the health and the courage to accomplish this modest work. We would like to express our gratitude to our directors for this brief* ***Mr. SAMAI Djamel*** *, we thank him for our having supervised, oriented, helped, and advised. Thank you also to* ***Miss TRABELSI Selma*** *for having followed, advised, and guided us throughout the period of writing the thesis.*

*We would also like to thank the members of the jury for agreeing to review and evaluate this work.*

# List of Figures

# List of Tables

# Contents

## General Introduction

The world has known a great technological development and with the increasing need for effective security systems and with the inability of traditional digital security systems to protect information from piracy, the developers used the idea of bio-metrics-based on photos and palm-print as an alternative protection system and through this note we tried to address the bio-metric identification system based on the three-dimensional palm-print. Recent years have seen a growing interest in developing automatic palm-print recognition methods. Most of the previous work focused on two-dimensional (2D) palm-print recognition in the past decade.

However, shape information is lost in 2D images. Moreover, 2D palm-print recognition is not powerful enough in practice as its data can be easily falsified or polluted by noise. Thus, three-dimensional (3D) palm-print recognition is treated as an important alternative route to enhance the performance and robustness of currently available palm-print recognition systems.

The objective of our study is to realize efficient uni-modal and multi-modal Identification bio-metric systems using a transfer learning method based on deep learning. We will try to achieve our goal through three chapters:

•In the first chapter, we will study the main concepts related to bio-metrics systems, and their mode of operation, which reveals the two types of identification and authentication with the structure and performance of the system. An overview of the single-modal and multi-modal integration levels and the various possible integration levels will be provided.

•The second chapter: is devoted to deep learning, and is a more in-depth description of the evolving neural networks (CNN) and neuron networks, as well as the transfer of learning which is the method chosen in our project.

• The third chapter will describe our proposed models and explain the practical results of the various methods used in extracting and classifying traits. We will end this summary with a general conclusion. We also cite the perspectives of our work.

## Chapter 1

# Bio-metrics Generalities

## 1.1 Introduction

With the rapid technological development, companies, organizations and individuals needed greater protection of their information and security systems, especially with the increase in piracy operations. With the inability of traditional digital security systems to protect information from piracy, developers have used the idea of bio-metrics as an alternative protection system. The idea of bio-metrics depends on the physiological and behavioral characteristics of individuals, such as hand-print, facial image, iris, voice and signature, to identify and know people and distinguish them from others [1].

In this chapter, we start by presenting some generalities about bio-metrics such as: its definition, characteristics, and how they work and the types of bio-metric system and some basics.

## 1.2 Definition of bio-metrics

"Bio-metrics" is derived from the two Greek words "life" and "metric" (for measurement) which are the physical and behavioral features that uniquely characterize us and which can be practically sensed by devices and interpreted by computers so that they can be used as a substitute for our physical bodies in the world of digital technology. In this way, we can link digital data to our identity permanently, consistently and without ambiguity, and we can retrieve that data using computers in an automated way. It has a lot of promise over conventional password-based systems like PIN and password, and it has the potential to completely change the way authentication is done [2].

## 1.3 Bio-metrics categories

Bio-metric are divided into three categories: Physiological, Biological and Behavioral

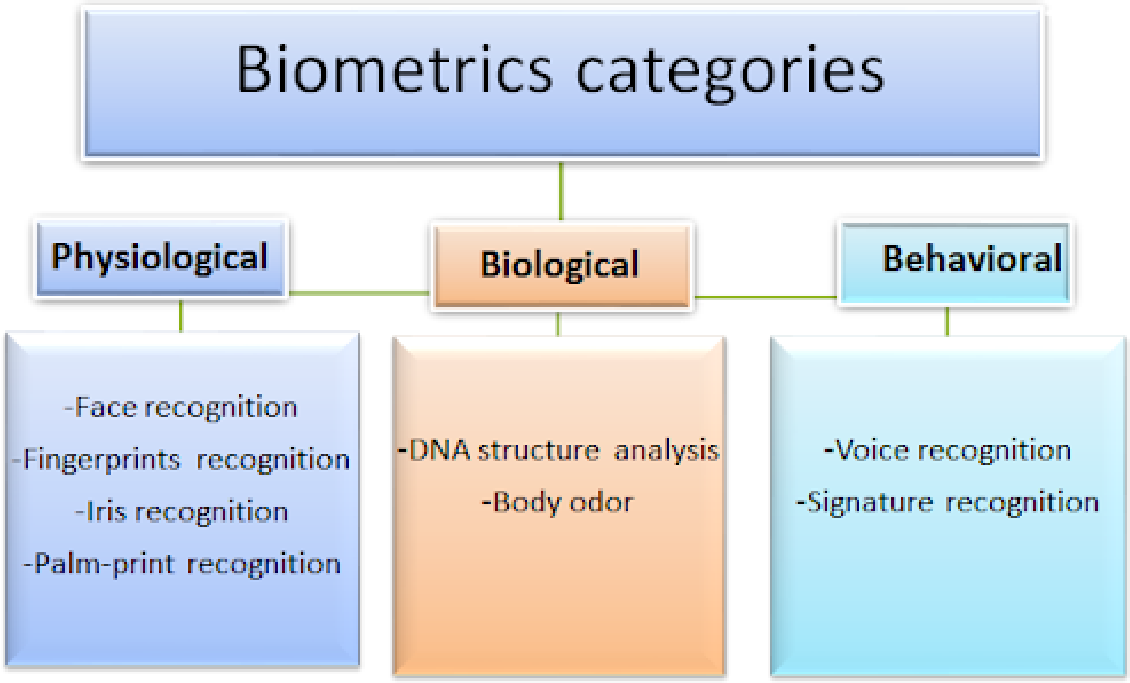


Figure 1.1: Categories of bio-metrics system

### 1.3.1 Physiological

Physiological category includes the features we are born with. This modality is based on the shape and size of the body. Examples are as follows:

**a)Face recognition:**

Face recognition techniques work by analyzing the human face in images, and then converting that into digital data according to features present in each face (such as the distance between the eyes, the length of the nose, the shape of the lip contour, the spacing of the ears, the width of the chin, etc.), and then matching the face with A picture of the device owner if he is an individual, or with pictures in a database of faces in other sectors. This technology can be used to identify the identity of the owner of the face or to ensure that he has sufficient powers to access a site or use a device. This process takes place in fractions of a second, down to a few seconds, depending on the degree of analysis performed by the system [3].

Figure 1.2: Face recognition example

**b)Fingerprint recognition:**

It is known that each person has his own fingerprints that distinguish him from the rest of humanity. It was used to be an alternative to the password for entering the personal computer or the workplace, and forensic medicine also used it to identify the perpetrators of the crimes, Fingerprints have three basic patterns called arcs, rings ,and swirls [4].

Figure 1.3: Fingerprints recognition example

**c)Iris recognition:**

The colored ring that surrounds the pupil called Iris. According to the ophthalmologist's research, every iris has a highly complex unique texture, which is unchanged over decades of life. For this reason, it can be used as bio-metrics recognition. These complex structures of the iris can be captured with a simple visible light camera; then can be digitally imaged. Iris recognition used for authentication or identification by Daugman algorithms, which used to abstract the iris data and compatriot previously enrolled irises. Iris recognition may be the most promising of all the physiological bio-metrics because of its very high accuracy [5].

It is suitable for high security such as airplane control rooms or nuclear plants [6].

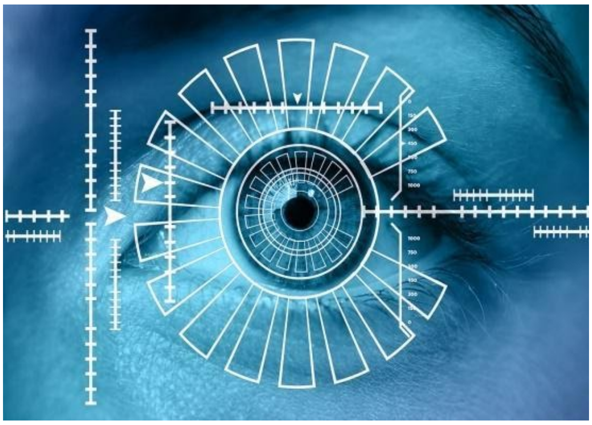

Figure 1.4: Some Fingerprints recognition example

**d)Palm-print recognition:**

The palm is the surface of the inner portion of the hand, from the wrist to the root of the fingers. The print is a mark made when some part of the body (finger; palm ... etc.) is pushed against some surface. So, the palm-print shows the physical features of palm patterns such as lines, points, minutiae, and texture. It can be used as a unique identification because it's varied from person to person [7]. In the palm, there are three types of lines: Principal Lines Wrinkles and ridges. There are six types of features on a palm: Geometry Features, Principal Line Features, Wrinkle Features, Datum Points, Delta Point Features, and Minutiae Features [8].

Figure 1.5: Palm-print recognition example

### 1.3.2 Biological

Biological bio-metrics use traits at a genetic and molecular level. These may include features like DNA or your blood, which might be accessed through a sample of your body's fluids.

**a)DNA structure analysis:**

They are found in the cells of the body; they are specific from one individual to another and make it possible to identify them with certainty through a small part of the skin, a trace of blood, or a drop of saliva. However, this bio-metric procedure is the subject of research. Intensive because it is an excellent identification technology with a much lower margin of error than other bio-metric methods [9].

Figure 1.6: DNA structure analysis example

**b)Body odor:**

It has long been known that people can be identified by their individual scents, and this fact has been exploited in the use of dogs to track people. . Recent advances in semiconductor-based chemical analyzes have led to the development of "electronic noses" that can measure spectrum concentrations (up to 32) of different chemicals. These sensors have neither range nor sensitivity to the human nose, and suffer from a range of problems, such as the need for calibration and the risk of permanent drift or "poisoning". It is also clear that subjective odor is susceptible to all kinds of influences, from diet and health status to the use of soaps, perfumes, and deodorants, and it is not yet clear whether these factors can be normalized far enough to allow reliable identification of individuals [10].

### 1.3.3 Behavioural

The behavioral category deals with the features we learn in our life as a result of our interaction with the environment and nature. This modality is related to change in human behavior over time. Examples of this category are:

**a)Voice recognition:**

voice recognition used for the identification and the verification of the voice. When a person speaks a waveform formed by the voice. The waveform is known as a voice pattern. Everyone has a different pitch, which can be considered physical features; the human voice is classified into a behavioral bio-metrics identifier. There are many factors ranging from the effect on the voice pattern like the qualities of the microphone, the communication channel, the aging, medical conditions, or even emotional states. Voice recognition methods are text-dependent and text-independent methods. The text-dependent methods distinguish the voice by the same utterance, such as specifically determined words, numbers, or phrases. But the text-independent methods identify the voice in whatever form of words or numbers [6].


Figure 1.7: Voice recognition

**b)Signature recognition:**

Belonging to the bio-metric family of products, Bio-metric Signature Verification authenticates the signers by measuring their handwritten signatures. The signature contains unique bio-metric data, such as the writing rhythm, acceleration, and pressure. Unlike other electronic signature capturing methods, Bio-metric Signature Verification does not treat the signature as a graphic image. With graphic images, such as the scanned-in signatures we often attach to our documents, it is impossible to detect the dynamics within each individual's signature. Hence, the signatures can easily be copied. By contrast, Bio-metric Signature Verification measures exactly how the signature is signed [11].

Figure 1.8: Signature recognition

## 1.4 Bio-metric characteristics requirements

Any human physiological and behavioral characteristic can be used as a bio-metric characteristic as long as it satisfies the following requirements [12]:
**a) Acceptability:**the degree to which people can accept a specific bio-metric identifier in their everyday lives.
**b)Performance:**which refers to a system's accuracy, pace, and robustness, as well as its resource requirements and operational or environmental factors that influence its accuracy and speed.
**c)Universality:**each person should have the characteristic.
**d)Uniqueness:**any two people must be sufficiently different in terms of the characteristic.
**e) Collectability:**the characteristic can be measured quantitatively.
**f) Circumvention:**Much traditional bio-metrics can be easily falsified or combined without anyone's consent. Fingerprints can be stolen from the user's cup and voices and faces can be secretly recorded.

## 1.5 The processes involved in bio-metric authentication system

**Enrollment(learning):**A bio-metric device is programmed to recognize a specific person. The person first provides an identifier, such as an identity card. The bio-metric is connected to the identity specified on the ID document. The bio-metric (fingertips, hand, or iris) is then presented to an acquisition system by He or she. One or more samples are selected, encoded, and saved as a reference guide for future comparisons after the distinctive features are found [13].

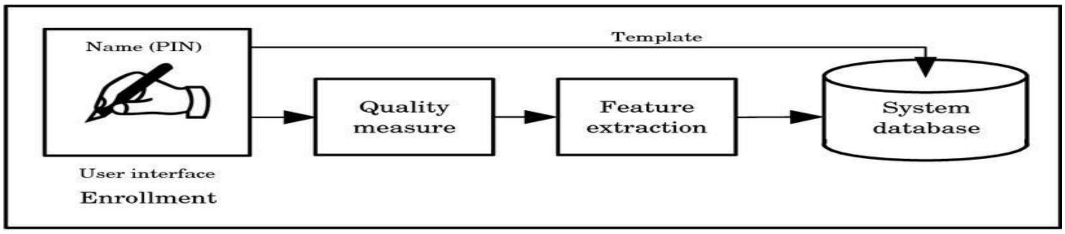


Figure 1.9: Enrollment in a bio-metric system

**Identification(1:N):**Identification systems are referred to as 1: N (one-to-N, or one-to-many) matching when an individual bio-metric is compared against multiple bio-metric templates in the system's database. Positive and negative identification systems exist. Positive identification systems are intended to ensure that a person bio-metric is accurate [14].

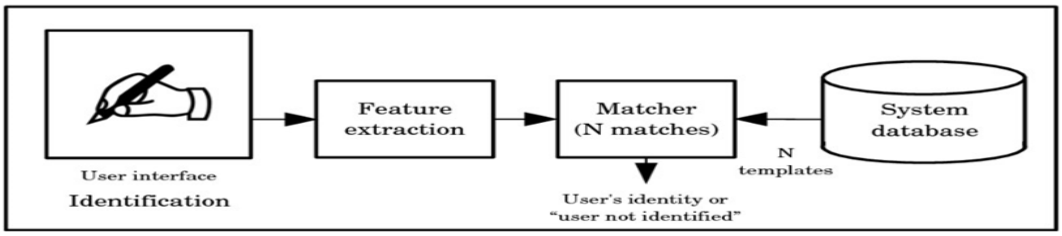


Figure 1.10: identification in a bio-metric system

**Verification(1:1):**In verification systems, the step after enrollment is to verify that the person is the person they claim to be (the person who enrolled) is this a term used to describe verification (one-to-one)? It is always based on matching the bio-metrics provided by the individual to their reference model. A system that requires employees to authenticate their purported identities before giving them access to secure buildings or computers [13].

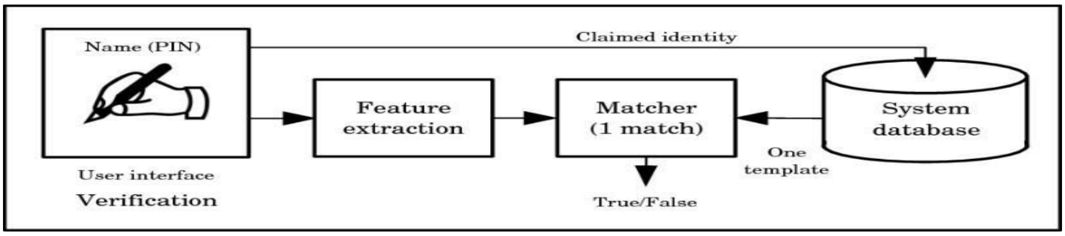


Figure 1.11: verification in a bio-metric system

## 1.6 Basic modules in bio-metric system

The sensor module, function extractor module, matcher module, and decision module are the four basic modules in bio-metrics systems. These four modules are required in any bio-metric system in order to obtain, process, and transform raw bio-metric data into useful information [15].
**Sensor module:**The sensor captures raw bio-metric data and scans the bio-metric trait to convert it to digital form in this type of module. This module sends the data to the function extraction module after it has been converted to digital form [15].
**Feature extraction module:**It takes the sensor's raw data and converts it into a bio-metric prototype. It extracts the required features from the raw data, which necessitates a great deal of care because important features must be extracted as efficiently as possible. It essentially takes out the noise from the input sample before sending it to the matcher module [15].
**Matcher module:**This module uses a matching algorithm to compare the input sample to the templates stored in the database, resulting in a match score. The decision module receives the corresponding match ranking [15].
**Decision module:**It compares the matching score to the predefined security threshold after accepting the match score from the matcher module. If the match score exceeds some specified security threshold, the individual will be accepted; otherwise, it will be rejected [15].

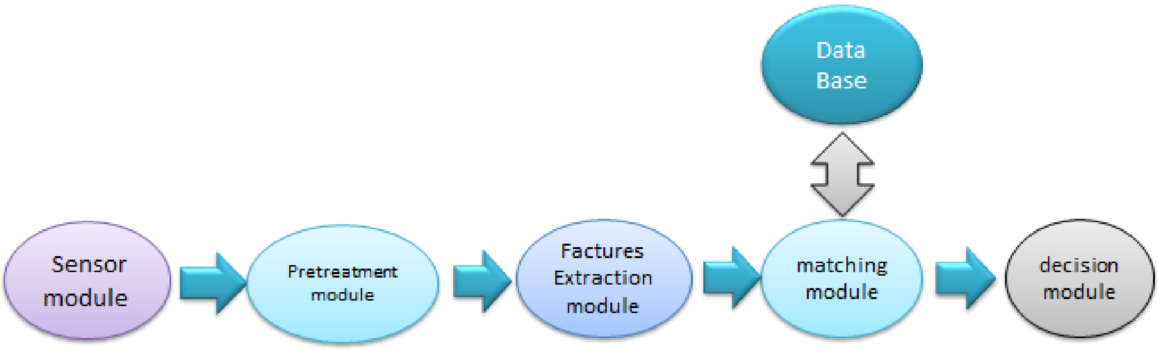


Figure 1.12: The basic contents of a bio-metrics system

## 1.7 Bio-metric system performances measuring

The performance and the accuracy of the bio-metric system are referring to the errors in the process of matching. The performance can be proved by the Receiver Operation Characteristics (ROC) curve, which shows the relation between the False Reject Rate (FRR) and False Accept Rate (FAR) at all thresholds. These systems sometimes falsely reject a genuine user (FRR) or falsely accept an impostor (FAR). FRR and FAR are indicators of an error rate of a bio-metric system, and they are the important ones. FRR is the proportion of the rejection of persons when it should have been accepted. And the FAR is the proportion of the acceptance of persons when it should have been rejected. When FRR and FAR have the same value, it becomes a new indicator which is called Equal Error Rate (ERR); and it is an important indicator to evaluate the performance of a bio-metric system. The FAR is more important than FRR in high-security domains like nuclear power stations. However, FRR tokens the same measure of convenience and availability [16].

$$FAR = \frac{the\ number\ of\ impostors\ accepted(FA)}{the\ total\ number\ of\ impostors\ access} \tag{1.1}$$

$$FRR = \frac{the\ number\ of\ rejected\ customers(FR)}{the\ total\ number\ of\ customer\ accesses} \tag{1.2}$$

$$ERR = \frac{Number\ of\ false\ acceptances(FA) + Number\ of\ false\ rejections(FR)}{the\ total\ number\ of\ accesses} \tag{1.3}$$

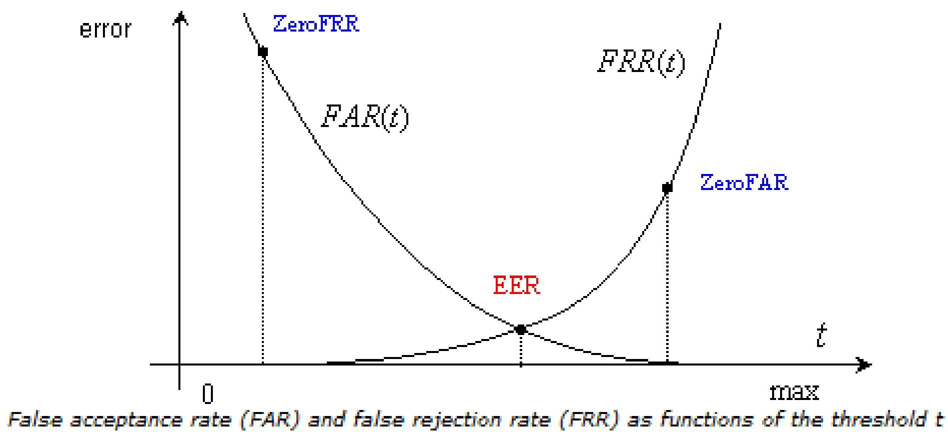


Figure 1.13: Equal Error Rate point (EER)

## 1.8 Uni-modal and multi-modal

Unimodal and multi-modal Bio-metric techniques are classified into two types based on the number of features used to determine a person's identity.

### 1.8.1 Uni-modal bio-metric technique

Uni-modal bio-metrics relies on the only obvious source of information for authentication (eg, a single fingerprint, a face). Typical bio-metric attributes may not meet the required performance requirements; because they have a lot of error rates [17].

### 1.8.2 Multi-modal bio-metric technique

Some of the problems that can affect uni-modal bio-metrics systems can be mitigated by using multi-modal bio-metrics systems. Systems that integrate signals obtained from two or more bio-metric sources for the purpose of identifying people are called multi-modal bio-metric systems. Multi-modal bio-metric systems have many advantages over uni-modal systems. Using an effective fusion technique to combine information from different modalities can considerably increase the overall accuracy of the bio-metric system [18].

### 1.8.3 Multi-bio-metric systems categories

We will describe the various scenarios in which multiple sources of evidence can be obtained. The first four scenarios described below use a single trait to achieve information fusion, whereas the fifth scenario uses multiple traits.

**Multi-sensor systems:** multiple sensor systems use multiple sensors to capture a single bio-metric trait for an individual. For example, a face recognition system may deploy multiple 2D cameras to obtain an image of a subject's face, and the use of multiple sensors can, in some cases, obtain supplementary information that can enhance the system's ability to recognize [17].

**Multi-algorithm systems:** The same bio-metric data is processed using multiple algorithms in these systems. A texture-based algorithm and a minutiae-based algorithm, for example, can both operate on the same fingerprint image to extract diverse feature sets that can improve system performance. This does not necessitate the use of new sensors. Furthermore, the user does not have to interact with multiple sensors, which increases user convenience [19].

**Multi-instance systems:** These systems, also known as multi-unit systems in the literature, use multiple instances of the same body trait. An individual's left and right index fingers, or left and right irises, for example, may be used to verify an individual's identity. In general, these systems do not necessitate the addition of sensors, the development of new feature extraction and matching algorithms, and are thus cost-effective [19].

**Multi-sample systems:** A single sensor can be used to collect multiple samples of the same bio-metric trait in order to account for trait variations or to obtain a more complete representation of the underlying trait. A face system, for example, may capture (and store) a person's frontal profile, as well as the left and right profiles, to account for variations in the facial pose. Similarly, a fingerprint system with a small sensor may acquire multiple dab prints of an individual's finger to obtain images of

different regions of the fingerprint [19].
**Multi-modal systems:** These systems combine evidence presented by various body traits to determine identity. Some of the first multi-modal bio-metric systems, for example, used face and voice features to determine an individual's identity. Physically uncorrected traits (for example, fingerprints and iris) are expected to improve performance more than correlated traits (e.g., voice and lip movement). The cost of deploying these systems is significantly higher due to the need for new sensors, and as a result, the development of appropriate user interfaces [19].

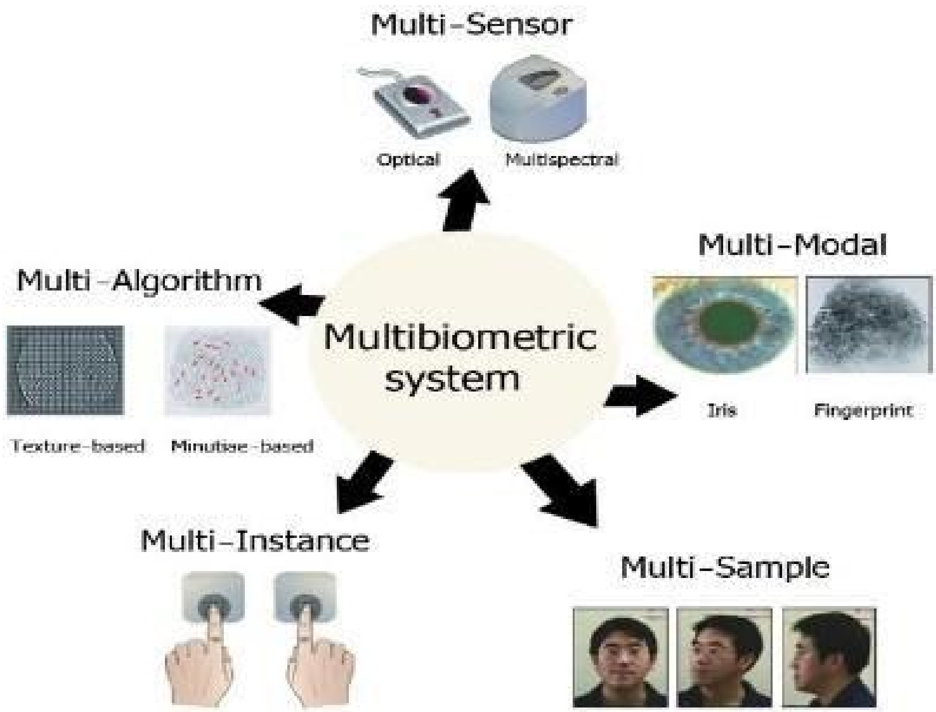


Figure 1.14: Multi-bio-metric systems categories

### 1.8.4 Fusion levels

Biometric fusion is a technique for combining data from different sources. Each bio-metric channel's classification results multi-modal bio-metric fusion incorporates aspects from various sources. A variety of bio-metric features to help you develop your strengths and weaknesses reduce the individual aspects 'limitations' the effectiveness of the fusion scheme has a major impact on the multi-modal bio-metric system's accuracy the

several degrees of fusion are as follows [20]:
**Fusion at the sensor stage:** Raw data from various sensors can be combined and exercised to create new bio-metric data from which a character can be derived. Bio-metric traits from various sensors, such as fingerprints, video cameras, iris scanners, and digital signatures, are combined to create a bio-metric trait that can be processed. Sensing a speech signal with two different microphones at the same time can be fused and then subjected to feature extraction [20].

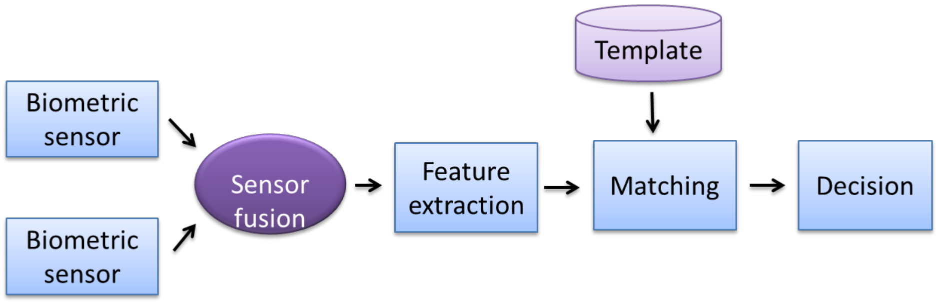


Figure 1.15: A parallel mode Fusion at the sensor stage

**Fusion of feature levels:** To create a composite feature set, feature sets extracted from different bio-metric channels can be fused using a particular fusion algorithm. Different modalities' feature collections agree to derive a minimum feature set from a high dimensional feature vector [20].

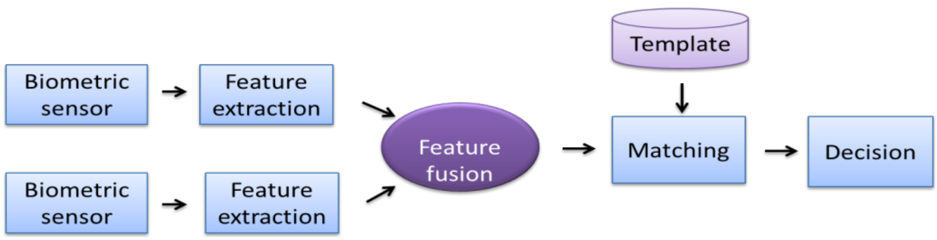


Figure 1.16: A parallel mode Fusion of feature levels

**Fusion of Level match score:** For each method, feature vectors are created separately. When comparing extracted feature vectors to models, they are uniquely stored in the database with all bio-metrics attributes in order to produce matching results. By exactitude Outputs a collection of matching results from each bio-metric channel, which are combined to produce a compound match score [20].

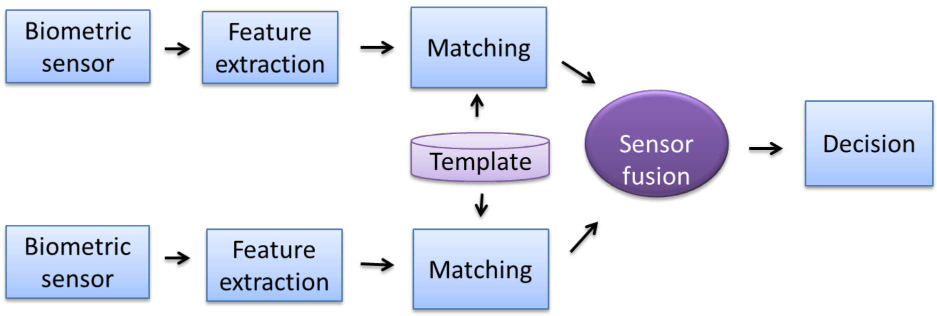


Figure 1.17: A parallel mode Fusion of Level match score

**Fusion at the rank level:** Rank-level fusion is a modern approach to fusion in which both the classifier and each registered identity are related. Fusion entails standardizing a rank's production to Subsystems for individual bio-metrics, as well as creating a new identification Rank to aid in making the final decision [20].
**Decision-making stage integration:** The final decision in a multi-modal bio-metrics system is based on separate decisions made by different modalities using techniques like majority voting, knowledge of behavior space, weighted vote, law, or rule degree of resolution [20].

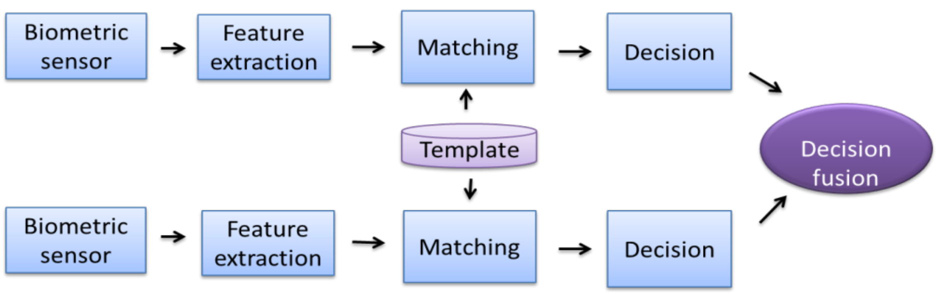


Figure 1.18: A parallel mode Decision-making stage integration

## 1.9 Conclusion

We discussed some definitions and concepts relevant to bio-metric systems in this chapter. We looked into the development of a multi-modal bio-metric system as well as the fusion process. We'll look at deep learning convolutional neural networks in the next chapter.

Chapter 2

# Deep learning and Transfer learning

## 2.1 Introduction

The researchers studied the reasoning/behavior of intelligent entities such as humans and created algorithms based on that behavior. These algorithms have been used to solve a wide range of interesting problems, including the creation of systems that behave intelligently in restricted domains such as medical diagnosis, software diagnostics, financial decisions, navigating difficult terrain, monitoring the potential failure of a space shuttle, speech and face recognition. AI arose as a new discipline to develop computer systems that could learn, react, and make decisions in a complex changing environment. The first AI efforts focused on modeling the neurons in the brain. An artificial neuron is regarded as a binary variable that can be turned on or off. Cognitive science is concerned with empirical studies of the mind, whereas AI is concerned with the creation of an artificial mind. However, because of their shared interests, each field can learn from the other. An artificial neural network is made up of a large number of neural units (artificial neurons) whose behavior is based on how real neurons communicate with one another in the brain. Each neural unit is linked to many other neural units[21]. Machine Learning is a branch of Artificial Intelligence. It is a technique for determining the "model" from the "data." In this context, data refers to information such as documents, audio, images, and so on. The final result of Machine Learning is the "model". The name alludes to the fact that the technique analyzes the data and finds the model on its own, rather than requiring a human to do so. Machine Learning is a technique for discovering (or learning) a model from data. It is appropriate for intelligent problems such as image recognition and speech recognition, where physical laws or mathematical equations fail to produce a model [22]. New neural network training algorithms and dramatically increased computer processing speed resulted in the re-emergence of neural networks in the field known as deep learning. Deep learning neural network architectures differ from traditional neural network architectures in that they frequently have more hidden layers. Deep learning has been used to solve computer vision and speech recognition tasks that were previously difficult to solve using other methods [21].

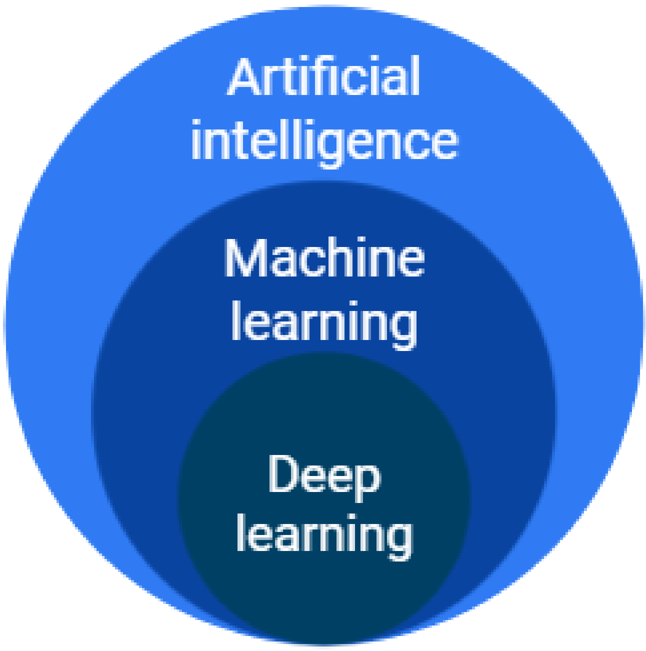


Figure 2.1: The relation between artificial intelligence, machine learning and deep learning

## 2.2 What's Deep learning?

Deep learning is a type of artificial intelligence derived from machine learning (machine learning) where the machine is able to learn by itself, his branch of machine learning. Unlike traditional machine learning algorithms, whose learning ability is limited regardless of the amount of data acquired, deep learning systems can improve their performance by accessing more data: a more experienced machine. Once machines have gained enough experience through deep learning, they can be used for specific tasks such as driving a car, spotting weeds in a field, detecting diseases, checking machines for malfunctions, etc. Deep learning is used in many fields: Image recognition, Facial recognition, Automatic natural language processing, Autonomous cars, Voice search and voice-activated assistants, Detection of brain cancer.Deep learning is a primary topic as well. It is a relatively new set of training techniques for multilayered neural networks. It includes several algorithms for training complex types of neural networks. With the advancement of deep learning, we now have efficient methods for training neural networks with multiple layers [24].

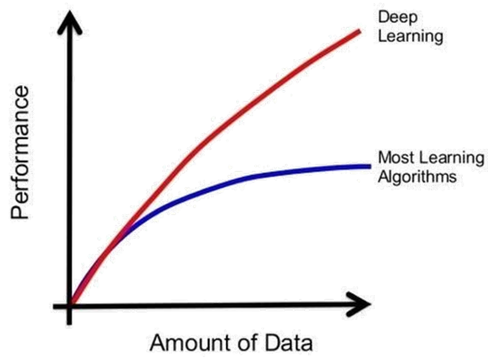


Figure 2.2: data science techniques scale with the amount of data

## 2.3 Neural Networks

An artificial neural network learning algorithm, or neural network, or just neural net , is a computational learning system that uses a network of functions to understand and translate a data input of one form into a desired output, usually in another form. The concept of the artificial neural network was inspired by human biology and the way neurons of the human brain function together to understand inputs from human senses. Neural networks are just one of many tools and approaches used in machine learning algorithms. The neural network itself may be used as a piece in many different machine learning algorithms to process complex data inputs into a space that computers can understand. Neural networks are being applied to many real-life problems today, including speech and image recognition, spam email filtering, finance, and medical diagnosis, to name a few [24].

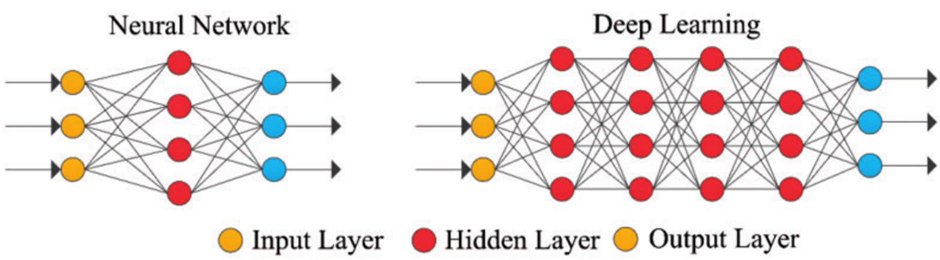


Figure 2.3: Architecture of neural network and deep learning

## 2.4 The basic Architecture of Neural Networks

The neural network evolved from a simple architecture to a more complex structure. The first neural network pioneers used a very simple architecture with only input and output layers, known as single layer neural networks. When hidden layers are added to a single-layer neural network, a multi-layer neural network is formed.

### 2.4.1 Single Layer Neural Networks:

is the most basic type of neural network. This neural network has only one input layer, one hidden layer and one output node [23].

### 2.4.2 Multi-layer Neural Networks:

A feed-forward neural network with fully connected layers and at least one hidden layer [25]. This type of architecture is also known as a multi-layer perceptron (MLP). The input of data features is referred to as an input layer, the prediction is referred to as an output layer, and others are referred to as hidden layers [26].

## 2.5 Convolutional Neural Networks

When we see an image or a face, we can recognize it right away. It is one of our fundamental abilities. This identification process is a conglomeration of many small processes and coordination between various vital components of our visual system. Using Convolutional Neural Network, or CNN, can replicate this astounding capability [27]. Convolutional Neural Network, also known as ConvNet or CNN, is a deep learning technique with many layers [29]. Because of their feature detection power, convolutional neural networks (CNNs) have become a widely representative deep model, significantly improving object classification and detection [30]. Convolutional layers apply a filter to the input image (or sound) by sliding the filter across the incoming signal to create an n-dimensional activation map. Filters are one of the main components of CNNs. They are square matrices with dimensions K x K, where K is an integer and is usually a small number, such as 3 or 5. Filters are sometimes referred to as kernels. The use of kernels is derived from traditional image processing techniques [31]. ConvNets have demonstrated exceptional performance in a variety of applications, including image classification, object detection, speech recognition, natural language processing, and medical image analysis. Convolutional neural networks are so effective at visual tasks that they outperform almost all traditional methods [31].

## 2.6 Architecture of CNN

Convents are made up of a series of different types of layers that work together to accomplish various tasks. The layers of a typical convolutional neural network are as follows:

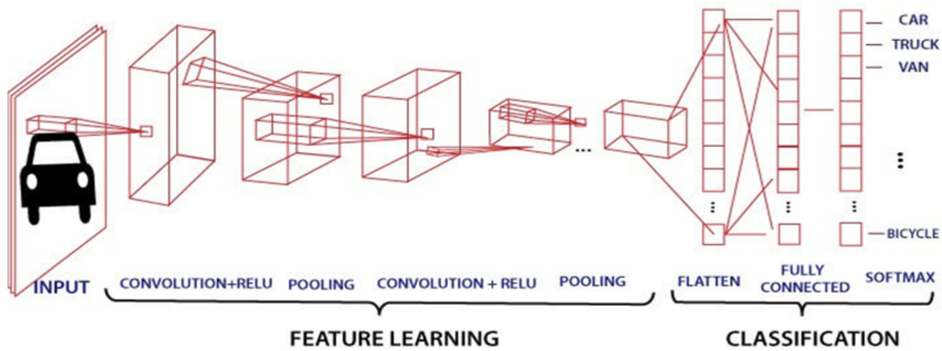


Figure 2.4: : CNNs Parts.

### 2.6.1 Convolutional layer

The core building block of a convolutional neural network is the convolution layer, which uses convolution operations instead of general matrix multiplication. Its parameters are made up of a collection of learnable filters known as kernels. The convolutional layer's main task is to detect features found within local regions of the input image that are common across the dataset and map their appearance to a feature map. A feature map is generated for each filter in the layer by applying the filter repeatedly across sub-regions of the entire image. The input area to which a filter is applied is referred to as the local receptive field. The receptive field is the same size as the filter [29]. The weights in each convolutional layer specify the convolution filters, and each convolutional layer may contain multiple filters. Every filter has some feature, such as an edge or a corner, and during the forward pass, each filter is slid across the width and height of the input, generating a feature map for that filter [29].

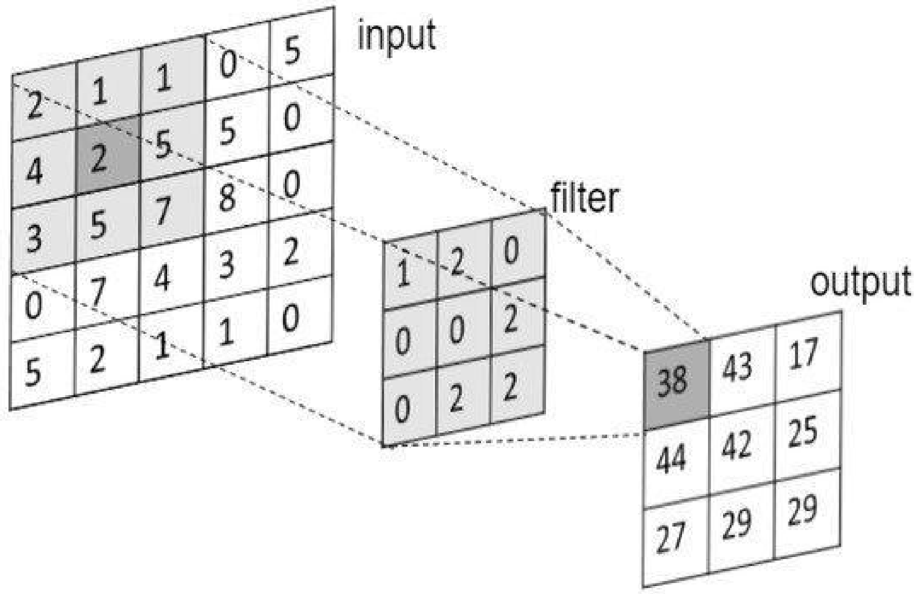


Figure 2.5: Convolution Operation

### 2.6.2 Activation function layer(ReLU)

To create a nonlinear relationship between the inputs and the outputs, an activation function is typically applied to the outputs of each hidden layer. There are two activation functions to choose from: ReLU or sigmoid. ReLU is a thresholding function that always returns the same input if it is greater than zero. Otherwise, it returns a value of zero. Each feature map returned by the convolutinal layer is activated by the ReLU layer using the ReLU activation function. It is extremely simple. Simply loop through each element in the feature map and return the feature map's original value if it is greater than 0. Otherwise, return the value [32].The output layer's activation function is determined by the task [33].

### 2.6.3 Pooling layer

That technique is frequently used in CNNs. By pooling, we reduce the entropy of the data by reducing the size of the activations. Reducing the size of activation is often beneficial because, while we lose some spatial information and spatial frequency in the activation itself, we gain a lot more activation responses through the filter bank of a layer. The deeper we go in a neural network, the more activation we will see and the larger the sizes of these activations will be. This will become computationally intractable very quickly. Pooling aids in the preservation of tractability [34]. The pooling layer employs a fixed operator, which is typically either max-pooling or average-pooling, as these are the most commonly used operations [35]. A pooling layer reduces the size of these feature maps by $\frac{1}{4}$ [36]. Pooling, like convolution, has hyper parameters: the size of the filter and the stride [33]. A pooling layer is typically added after a convolution layer, and it receives the output of the convolution as input. When a volume is pooled, each matrix in the volume is processed independently of the others. As a result, the output of a pooling layer applied to a volume is the same depth as the input. Typically, pooling contributes to the model's increased accuracy. It also increases training speed by reducing the number of neural network parameters [33].

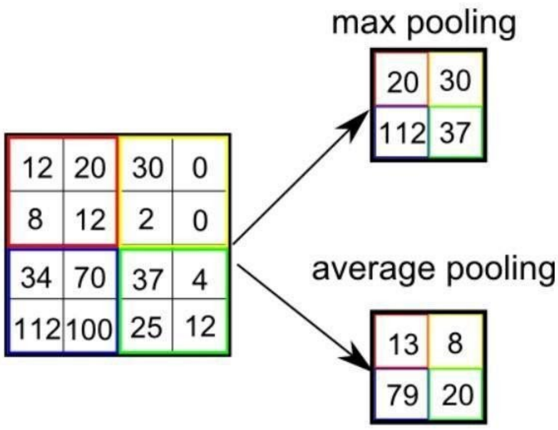


Figure 2.6: The resultant of max and average pooling

### 2.6.4 Fully connected layer

Convolutional neural networks are divided into two stages: feature extraction and classification. In ConvNets, the feature extraction stage is made up of a stack of convolution and pooling layers, while the classification stage is made up of one or more fully connected layers followed by a softmax function layer. The output of the final fully connected layer is fed into a classifier, which produces class scores. The two main classifiers used in ConvNets are Softmax and Support Vector Machines (SVMs). Softmax classifier generates probabilities for each class with a total probability of one, and SVM generates class scores, with the class with the highest score being considered the correct class [29]. The fully connected layer receives inputs from the outputs of the previous layer (activation maps of high-level features) and outputs a vector of n dimensions. In this case, n indicates the number of distinct classes. A Fully Connected layer will look at the features that are most closely related to a specific class and have specific weights. When we get the product of the weights and the previous layer, we get the correct probabilities for different classes [27]. In general, an FC layer simply converts the output of the previous layer into a 1D vector [32].

### 2.6.5 Dropout layer

More layers are available in CNN, and they can be easily added to the preceding layers. A dropout layer, for example, could be implemented by removing a percentage of the neurons in the final layer [36]. This layer can be linked to any other layer, including convolutional or fully connected layers. This layer has the same number of elements as the previous layer, which has the same number of outputs. In this layer, there is only one parameter called dropout ratio, which is denoted by p and is defined by the user when designing the network.

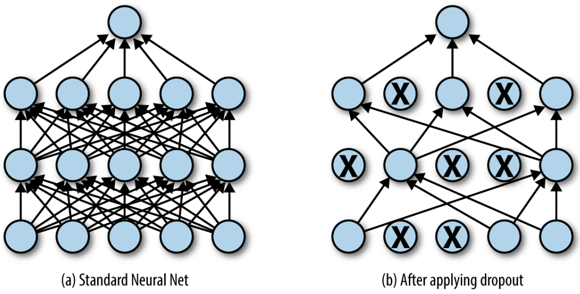


Figure 2.7: Dropout Neural Net Model.

## 2.7 Transfer learning

The goal of transfer learning is to improve learning in the target task by leveraging knowledge from the source task which was developed for re-used as the starting point. Transfer learning employs a model that was originally trained on a base dataset and is now being re-purposed to learn features on a new dataset instead of beginning the learning process on the data from scratch with random weight initialization. These models act as a feature extractor, the weights of all the layers in the network remain the same, but the last layer in the network is removed and replaced by your own classifier. Transfer learning has been applied and studied in the context of inductive learners, such as neural networks, and Bayesian networks [39]. In transfer learning, domain and task are two basic concepts. Domain D consists of two components: a feature space X and a marginal probability distribution P(X), where: X= x1,x2,. . . ,xnD Here are two basic domains: source domain (Ds) and target domain(DT). Domain with a lot of labeled data and knowledge is source domain, which is used to train base models.The target domain is the object to be transferred and given knowledge. A domain is defined as a pair D= X,P(X) Task T consists of two components: a label space Y and an objective predictive function(·). Correspondingly, the target task (TT) in the target domain (DT) using knowledge from the source task (TS) in source domain (Ds). A task is defined as a pair T=Y,P(Y—X) .

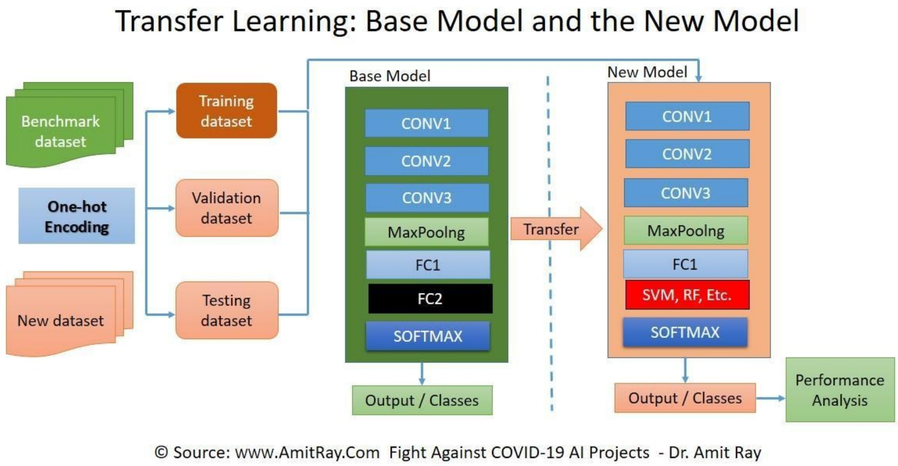


Figure 2.8: Transfer learning base model and new model

## 2.8 Transfer learning methods

Can be categorized based on the type of traditional ML algorithms involved, such as:
**Inductive transfer learning:** the source and target domains are the same, yet the source and target tasks are different from each other. The algorithms try to utilize the inductive biases of the source domain to help improve the target task.
**Transductive transfer learning:** there are similarities between the source and target tasks but the corresponding domains are different. In this setting, the source domain has a lot of labeled data while the target domain has none.
**Unsupervised transfer learning:** This setting is similar to inductive transfer itself, with a focus on unsupervised tasks in the target domain. In this scenario, labeled data is unavailable in either of the domains.

## 2.9 Pre-trained Neural Networks

At present, many models pre-trained on the Image Net challenge data are opened to the public and readily accessible, along with their learned kernels and weights, such as AlexNet, VGG, ResNet, GoogleNet [38].
**a) AlexNet**

AlexNet outperformed all previous rivals and won the ILSVRC in 2012 by lowering the top-5error to 15.3 percent, compared to the runner-26 up's percent. This paper popularized the use of CNNs in computer vision. Alex Net employs stacked convolution rather than alternative convolution pooling. A stack of small convolutions is preferable to a single wide receptive area of convolution layers since it introduces more nonlinearities and has less parameters.
Assume there are three 3 x 3 convolution layers stacked on top of each other (with non-linearity or pooling layer in between). Each neuron on the first convolution layer has a 3x3 view of the input volume in this case. A neuron in the second convolution layer has a 3x3 view of the first convolution layer and, as a result,a5 x5 view of the input amount. A neuron on the third convolution layer, for example, has a 3 x3 view of the second convolution layer and hence a 7 x 7 view of the input amount. Clearly, the number of parameters for a 7x7 receptive field is multiplied by 49, while three stacked 3x3 convolution shave factor of 3x(3x3)=27 parameters[39].

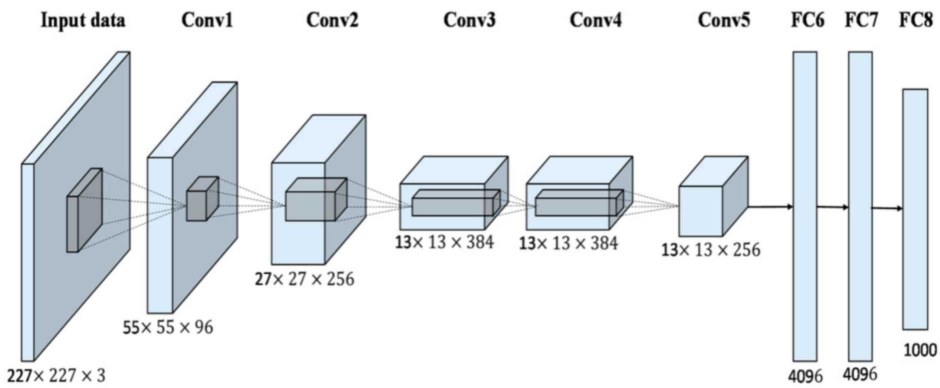


Figure 2.9: AlexNet graphic

## b) VGG

TheVGG network was created by researchers from the Oxford Visual Geometry Group, or the VGG for short. It is distinguished by its simplicity, with only three x three convolutional layers stacked on top of each other in increasing depth. Max pooling handles volume size reduction .After that,two completely connected layers with 4, 096 nodes each are accompanied by a softmax layer. The only input preprocessing is the subtraction of the mean RGB value computed on the training set from each pixel. Pooling is done by max-pooling layers, which are placed after some of the convolution layers. Max pooling is not applied on all convolution layers. Max pooling is done with a stride of 2 over a 2 x 2-pixelwindow. In each of the secret layers, ReLU activation is used. In most VGG variants, the number of filters increases with depth. The 16-layered architecture VGG-16 is depicted in the diagram below. In the following segment, the 19-layered architecture with uniform 3x3 convolutions(VGG-19)is shown alongside Res Net. The performance of VGG models demonstrates the value of depth in image representation [39]:

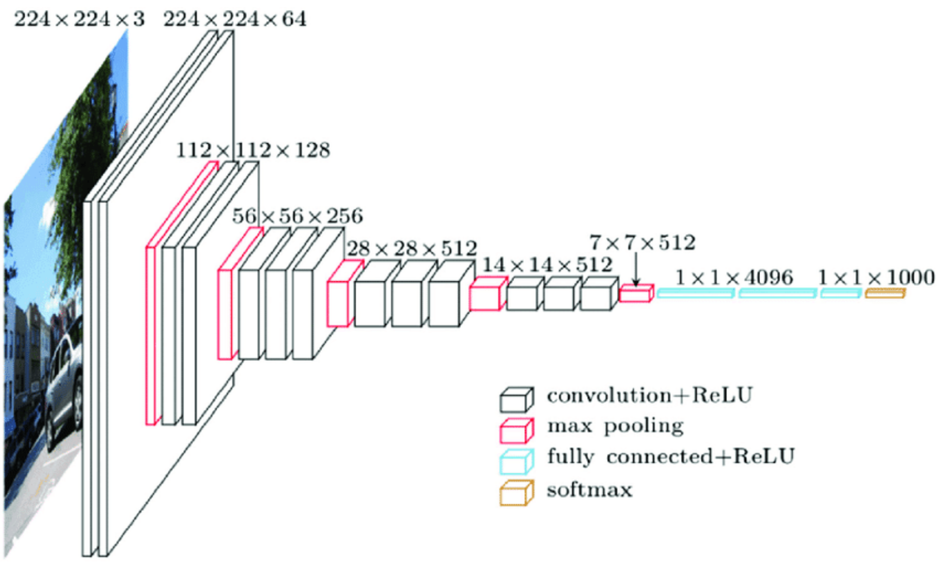


Figure 2.10: VGG Net graphic

### c) ResNet

The Res Net model includes a residual learning system to make training deeper networks easier. The architecture is focused on recasting network layers as learning residual functions with respect to layer inputs. Although the depth of the residual network is eight times, its complexity is lower,depicting the architecture[40].

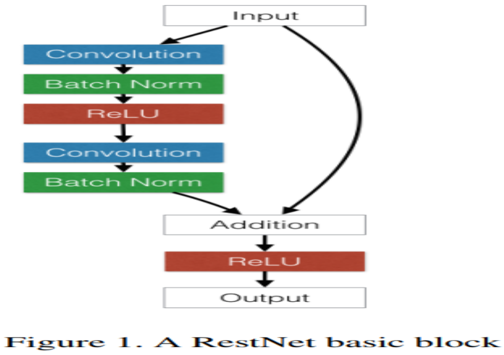


Figure 2.11: A ResNet basic block

### d) GoogleNet

GoogLe Net convolutional network was the ILSVRC 2014 winner. It had a top-five error rate of 6.67 percent! This performance was very similar to that of a human. The inception layer is a new architectural component introduced by Google Net that employs a CNN. The inception layer's intuition is to use larger convolutions while maintaining a fine resolution for smaller details on the images. As a result, we can convolve in parallel with different kernel sizes, ranging from 1x1 to 5x5, and the outputs are concatenated to create the next layer. Adding more layers clearly expands the parameter space. A dimensionality-reduction trick is used to manage this. It is important to note that the 1 x 1 convolution does not minimize the image's spatial dimensions. However, we can minimize the number of feature maps with 1x1 filter sand the depth of the convolution layer, as shown in the diagram below [39]:

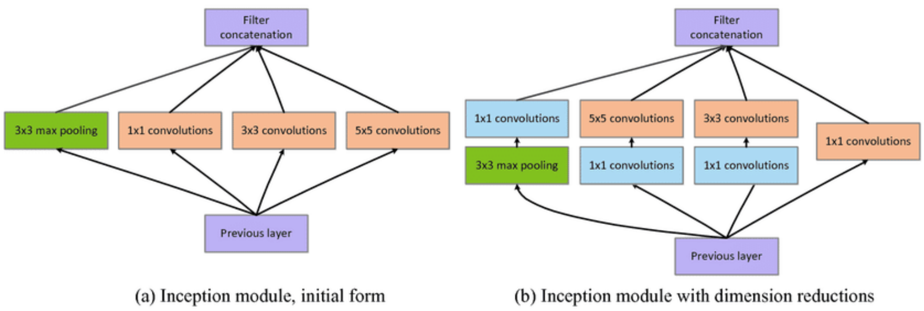


Figure 2.12: Inception module

## 2.10 Conclusion

We learned in this chapter the key concepts related to deep learning (definition, process, etc.)and we learned the Architecture of CNN and his layers .and that the deep learning convolutional neural network is a great method to classify images, it's easy to extract feature from the image because it doesn't need a domain expert to extract features and reduce the images size into a form which is easier to process, without losing features which are critical for getting a good prediction, as well as some ideas on transfer learning.

# Chapter 3

# Results and discussion

## 3.1 Introduction

Bio-metric identification of individuals by means of palm-print recognition based on transfer learning which is mainly used to extract and classify image attributes. palm-print verification is one of the most important and popular methods of personal authentication despite its high accuracy and efficiency. This chapter collects empirical results to identify Image the palm-print and then identify people using transformed learning applied to the PolyU database, which contains two-dimensional (2D) images and three-dimensional (3D) depth information. We combined the points to obtain a multi-modal identification system. We leveraged CNN to create our bio-metrics system and combine two or more bio-metrics patterns to create a multi-modal bio-metrics system.

## 3.2 Definition of Palm-Print

As it munched in chapter 2, the palm-print has many features that can offer a great way of personal recognition. A new personal authentication system presented by Zhang et al. that simultaneously exploits 2D and 3D palm-print features. An active stereo technique, structured light used in this developed system to capture 3D image or range data and a registered intensity image of the palm. The integration of 3D features gives an important development in performance compared to the 2D palm-print features alone. And also, it presents results to illustrate that the bio-metrics system is difficult to avoid [41].

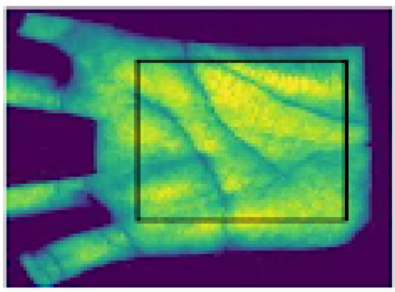

Figure 3.1: Palm-Print example

### 3.3 Palm-print database used

The PolyU Palm-print image database 2D/3D, that we used in our experiments contains 8000 samples collected from 400 various palm trees. Twenty samples were collected from each of these palms in two separate sessions, each with 20 palm images collected from the right and left hand in a separate session, where 10 samples were taken in each session, and the average time between the two sessions is 1 month. Each sample contains a 2D ROI and a 3D ROI. All 2D images in this database are 8-bit gray-scale BMP (bitmap) and their sizes are 128 x 128 pixels. The PolyU 2D/3D database was created by the Bio-Research Center (UGC/CRC) of Hong Kong Polytechnic University. The following figure shows some images from this database.

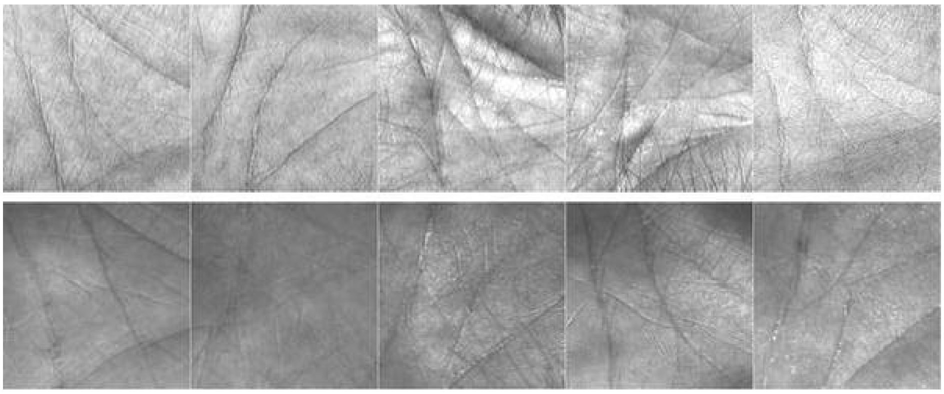

Figure 3.2: Examples of 3D and2D palm-print database

### 3.4 Extraction of ROI (Region Of Interest)

An ROI is a region of interest in an image and can be used as a starting point for many image processing algorithms. In the field of bio-metrics, the meaning of ROIs depends on the type of bio-metric modality. A novel approach is proposed where convolutional neural networks (CNNs) along with transfer learning are exploited. The extracted palm-print ROIs are fed to the final verification system, which is composed of two modules. These modules are (i) a pre-trained CNN architecture as a feature extractor and (ii) a machine learning classifier. To evaluate our proposed model, we computed the intersection over union (IoU) metric for ROI extraction along with accuracy, receiver operating characteristic (ROC) curves, and equal error rate (EER) for the verification task. The experiments demonstrated that the ROI extraction module could significantly find the appropriate palm-print ROIs, and the verification results were crucially precise. Therefore, the quality of the algorithm used to detect ROIs often conditions the quality of the result of the entire processing chain that we want to apply to an image. Also, the fact that the same ROIs (or more or less) can be detected on two different images but representing the same scene, is an important property and generally required for all ROI detection algorithms.

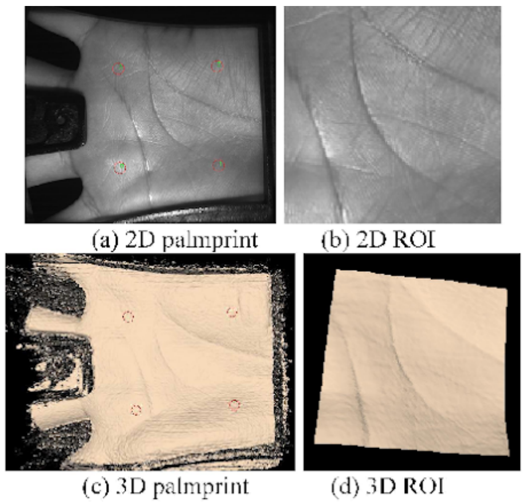


Figure 3.3: The ROI extraction of 3D palmprint from its 2D counter

## 3.5 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.

## 3.6 The architecture of our system

In this section, we describe our system. It is used to extract the features of the images of the PolyU database and classification them, by applying transfer learning. The system has two-stage. The first is a fine-tuning stage, in this stage, the image fine-tuned the network. and the second stage is knowledge transfer with soft labels, and in this stage, the network is trained by images of [42]. After getting the results of the uni-modal system ( 2D or 3D-GC or 3D-MC ), we will use the multi-modal system ( 2D with 3D-MC, or 2D with 3D-GC, or 3D-MC with 3D-GC, or 2D with 3D-MC with 3D-GC ) to enhance the performance of the system.

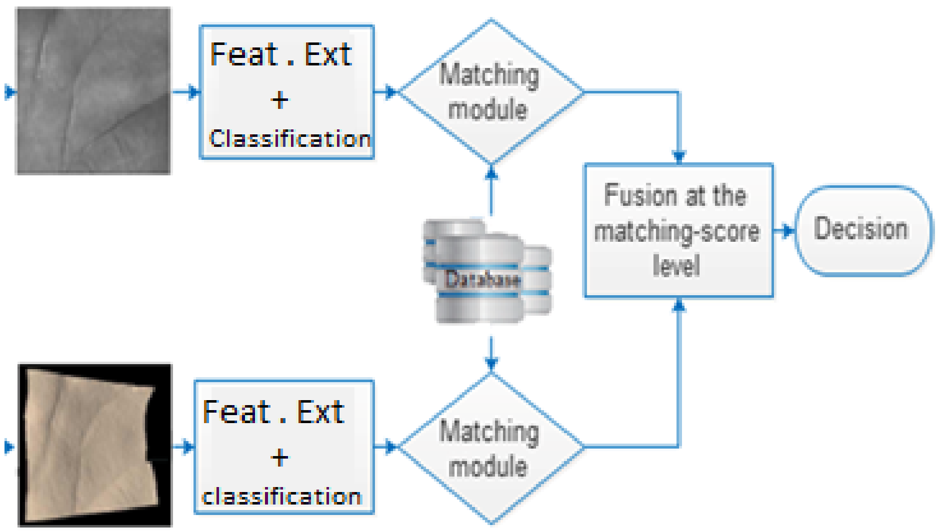


Figure 3.4: Architecture of our system

## 3.7 Work protocol

We make a system uni-modal for the identification of the person by the image of his palm. In this study, we use a database that contains palm print images for 400 persons; every person has 20 images. The twenty (20) palm print images are divided into two groups: one for training and the other for testing techniques and determining their performance. There are no rules for defining this post in quantitative terms. In our

series of tests, the database was divided as follows: Ten (10) pictures were used for the learning stage.
Palm Train = [1, 3, 5, 7.9, 11, 13, 15, 17, 19]
The remaining ten (10) images are used for various tests.
Palm test = [2, 4, 6, 8, 10, 12, 14, 16, 18, 20]
We use the PolyU database, which contains two types of images, two-dimensional (2D) images and three-dimensional (3D) images. The surface curvature feature-based method is investigated for 3D palm-print feature extraction and represents the information in the captured 3D palm-print images. In our work, we used two approaches of surface curvature, mean and Gaussian curvatures are used to classify points on a surface into different classes. The mean curvature (MC) and Gaussian curvature (GC) of the smoothed 3D ROI are calculated using partial differential estimation images:

$$MC(i,j) = \frac{M_1(i,j) + M_2(i,j) - 2M_2(i,j)}{2\sqrt{([1 + d_u^2(i,j) + d_u^2(i,j)]^3)}} \tag{3.1}$$

$$GC(i,j) = \frac{d_{uu}(i,j)d_{vv}(i,j) - d_{uv}^2(i,j)}{([1 + d_u^2(i,j) + d_v^2(i,j)]^2} \tag{3.2}$$

$$M1(i,j) = [1 + d^2u(i,j)]duu(i,j) \tag{3.3}$$

$$M2(i,j) = [1 + d^2u(i,j)]dvv(i,j) \tag{3.4}$$

The first, second hybrid partial derivatives of "d" are: duu, dvv, du, dv and duv Thus

$$d_u(i,j) = D_u * \bar{d}(i,j) \tag{3.5}$$

$$d_v(i,j) = D_v * \bar{d}(i,j) \tag{3.6}$$

$$d_{uu}(i,j) = D_{uu} * \bar{d}(i,j) \tag{3.7}$$

$$d_{vv}(i,j) = D_{vv} * \bar{d}(i,j) \tag{3.8}$$

$$d_{uv}(i,j) = D_{uv} * \bar{d}(i,j) \tag{3.9}$$

Derived estimation window masks are defined as follows:

$$\mathrm{D}_u = \vec{d_0}\vec{d_1}^T,\ D_v = \vec{d_1}\vec{d_0}^T,\ D_{uu} = \vec{d_0}\vec{d_2}^T,\ D_{vv} = \vec{d_2}\vec{d_0}^T,\ D_{uv} = \vec{d_1}\vec{d_1}^T$$

Where the column vectors $\vec{d_0}$, $\vec{d_1}$, $\vec{d_2}$ are given with:

$$\vec{d_0} = \frac{1}{7}[1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1]^T \tag{3.10}$$

$$\vec{d_1} = \frac{1}{28}[-3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3]^T \tag{3.11}$$

$$\vec{d_2} = \frac{1}{84}[5 \quad 0 \quad -3 \quad -4 \quad -3 \quad 0 \quad 5]^T \tag{3.12}$$

In order to compute more efficiently and take advantage of the curvature information, we convert it to gray scale images. We first normalize the Gaussian curvature GC or the mean curvature MC in $\bar{c}$ by the following:

$$\bar{C}(i,j) = \frac{0.5[c(i,j) - \mu]}{2\sigma} + 0.5 \tag{3.13}$$

Where ($\mu$) and ($\sigma$) are the mean and the standard deviation of the curvature value. With $\bar{c}$ most curvature values will be normalized in the interval[0, 1].

| | |
|---|---|
| {0 | $\bar{c}(x,y) \leqslant 0\}$ |
| $\{raund\ 255 \times \bar{c}(x,y)$ | 0 ¡ $\bar{c}$(x,y) ¡ 1} |
| {255 | $\bar{c}(x,y) \geqslant 0\}$ |

After that, we use transfer learning to extract the feature and classify these images. We use a pre-trained model of transfer learning which is “AlexNet”. After that, we will use a uni-modal system to find the best results for the two identification modes, open-set, and closed-set. Finally, to enhance the performance and get better results of a uni-modal system, we will use the multi-modal system by combination of the types.

## 3.8 Results and discussion

In this study, we used the model pre-trained Neural Networks alexNet, for treating the dataset PolyU 2D/3D . And there are two parties, the first is a uni-modal system, and the second is a multi-modal system.

### 3.8.1 Uni-modal systems

We change some parameters and other parameters taken by default for getting the best EER and ROR of each modal in our work. Our basic model took the following default settings: rmsprop optimizer, the learning rate is 0.0001. The parameters that we modify are the batch size and the epochs. After several experiments, the best parameters chosen of batch size and epochs are:

Table 3.1: The best parameters

| | Batch Size | Epoch |
|---|---|---|
| 2D | 22 | 18 |
| 3D-GC | 38 | 48 |
| 3D-MC | 42 | 41 |

The best parameter is the parameter that gives the smallest EER. After we find the best parameters we will use them in the two identification modes, open-set, and closed-set. The following table shows the best results we get:

Table 3.2: Uni-modal system results

| SPECTRAL BANDS | Open-set | | Closed-set | |
|---|---|---|---|---|
| | EER (%) | $T_0$ | ROR (%) | RPR |
| 2D | $5 \times 10^{-4}$ | 0.1100 | 99.92 | 16 |
| 3D-GC | 0.1045 | 0.0100 | 98.87 | 294 |
| 3D-MC | 0.0550 | 0.0280 | 99.70 | 099 |

Table 3.2 shows that the best results of each type, the best results in spectral band 2D are EER = 5 × 10-4 (%) in open-set mode and ROR = 99.92 (%) in closed-set mode. And for 3D-GC are EER = 0.1045 (%) in open-set mode and ROR = 98.87 (%) in closed-set mode. And at least for 3D-MC, we note an EER = 5.5 × 10-2 (%) in open-set mode and ROR = 99.70 (%) in closed-set mode. Figure 3.5 shows the FAR and FRR plots for 2D and 3D (MC and GC)

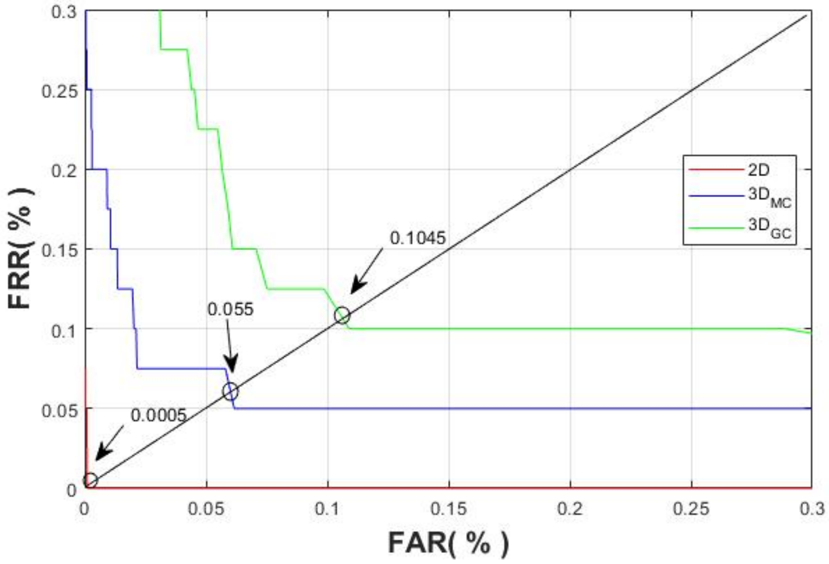


Figure 3.5: The curve of ROC (FRR in the function of FAR) of the system uni-modal

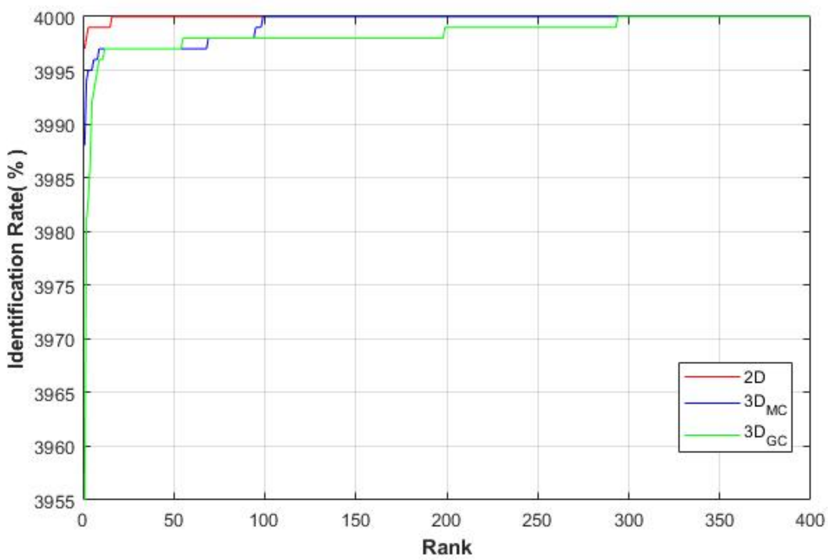


Figure 3.6: The curve of CMC of the system uni-modal

### 3.8.2 Multi-modal systems

We use the multi-modal system to enhance the performance and get better results of a uni-modal system. In our multi-modal system, we do the fusion at the matching score level and was carried out with six Rules of fusion which are: Rule Sum fusion(SUM), Rule of the product (PRO), Min Rule (MIN), Max Rule (MAX), Rule of the weighted Sum (WHT-SUM), and Rule of the weighted product (WHT-PRO).
The following tables present the results obtained with the different fusion rules that we have mentioned:

## Fusion 2D with 3D-GC

Table 3.3: Multi-modal system results (fusion between $2D$ and $3D - GC$ )

| Rules of fusion | Open-set | | Closed-set | |
|---|---|---|---|---|
| | EER (%) | $T_0$ | ROR (%) | RPR |
| SUM | 0.01 | 0.1294 | 100 | 1 |
| PRO | $3.7 \times 10^{-4}$ | 0.0190 | 100 | 1 |
| MIN | 0 | 0.1260 | 100 | 1 |
| MAX | 0.0154 | 0.0154 | 99.5 | 4 |
| WHT-SUM | 1.0025 | 0.1100 | 99.975 | 2 |
| WHT-PRO | $8.772 \times 10^{-4}$ | 0.1160 | 99.975 | 2 |

It can be observed from this Table 3.3 that the proposed multilevel approach for a combination of 2D and 3D-GC features achieves the best performance.Equal Error Rate (EER) of 0(%) and ROR of 100(%), and RPR=1. for the Min Rule fusion. This performance is significantly higher as compared to the case when either 2D or 3D-GC palm-print features alone are used.

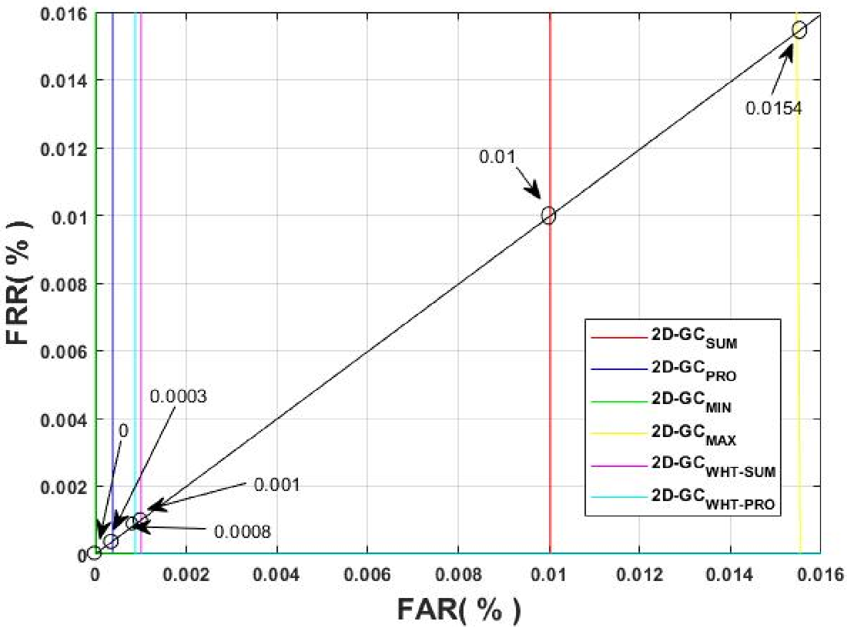


Figure 3.7: The curve of ROC (FRR in the function of FAR) of the system multi-modal 2D with 3D-GC

## Fusion 2D with 3D-MC

Table 3.4: Multi-modal system results (fusion between $2D$ and $3D - MC$ )

| Rules of fusion | Open-set | | Closed-set | |
|---|---|---|---|---|
| | EER (%) | $T_0$ | ROR (%) | RPR |
| SUM | $3.88 \times 10^{-3}$ | 0.1391 | 100 | 1 |
| PRO | $2.506 \times 10^{-4}$ | 0.0360 | 100 | 1 |
| MIN | $3.75 \times 10^{-4}$ | 0.1790 | 100 | 1 |
| MAX | 0.0101 | 0.1242 | 99.875 | 2 |
| WHT-SUM | $1.005 \times 10^{-3}$ | 0.11104 | 99.975 | 2 |
| WHT-PRO | $8.772 \times 10^{-3}$ | 0.1160 | 99.975 | 2 |

The experimental results in Table 3.4(Multi-modal system 2D with 3D-MC) show good results but the SUM and PRO and MIN give ROR=100(%)

and RPR=1 and with EER so small. The minimum EER in this fusion is the Rule of the product (PRO) which EER=2.506 x 10-4.

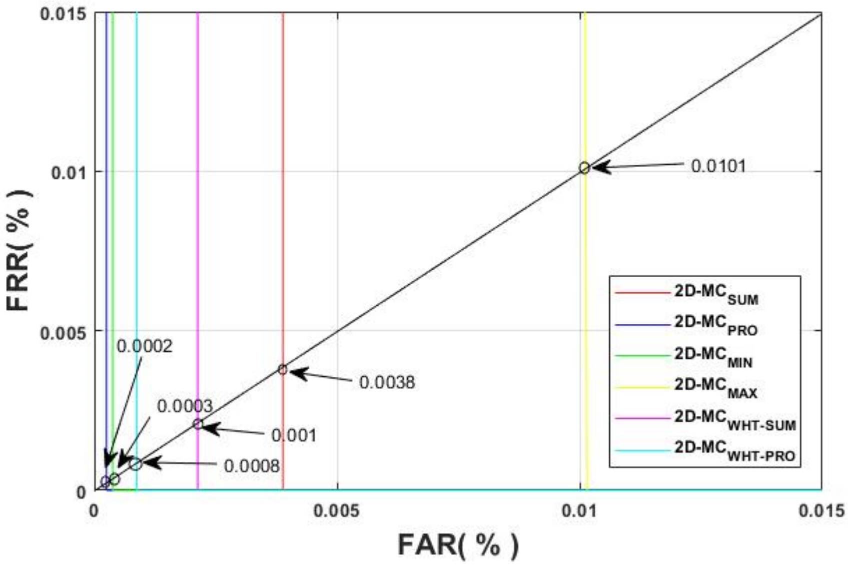


Figure 3.8: The curve of ROC (FRR in the function of FAR) of the system multi-modal 2D with 3D-MC

## Fusion 3D-MC with 3D-GC

Table 3.5: Multi-modal system results (fusion between $3D - MC$ and $3D - GC$ )

| Rules of fusion | Open-set | | Closed-set | |
|---|---|---|---|---|
| | EER (%) | $T_0$ | ROR (%) | RPR |
| SUM | 0.0251 | 0.084 | 99.92 | 56 |
| PRO | 0.024 | 0.002 | 99.92 | 46 |
| MIN | 0.025 | 0.045 | 99.85 | 59 |
| MAX | 0.025 | 0.126 | 99.42 | 56 |
| WHT-SUM | 0.025 | 0.079 | 99.87 | 62 |
| WHT-PRO | 0.025 | 0.056 | 99.92 | 78 |

It can be observed from Table 3.5 that the fusion of 3D-MC with 3D-GC presents fewer results as compared to the case when either 3D-MC or 3D-GC palm-print features alone are used. with EER = 0.024 (%) , ROR=99.92 (%) for product rule, and EER =0.0025(%) for the rest rules.

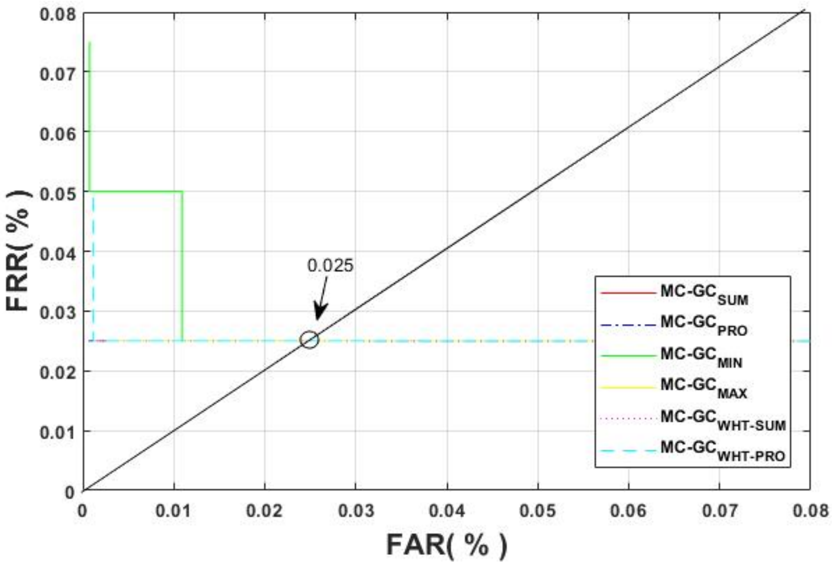


Figure 3.9: The curve of ROC (FRR in the function of FAR) of the system multi-modal 3D-MC with 3D-GC

## Fusion 2D with 3D-MC with 3D-GC

Table 3.6: Multi-modal system results (fusion between $2D$ and $3D - MC$ and $3D - GC$ )

| Rules of fusion | Open-set | | Closed-set | |
|---|---|---|---|---|
| | EER (%) | $T_0$ | ROR (%) | RPR |
| SUM | $8.72 \times 10^{-3}$ | 0.1403 | 99.975 | 2 |
| PRO | $3.75 \times 10^{-4}$ | $5.01 \times 10^{-3}$ | 100 | 1 |
| MIN | $6.26 \times 10^{-3}$ | 0.0150 | 100 | 1 |
| MAX | 0.0233 | 0.1239 | 99.425 | 4 |
| WHT-SUM | $1.0002 \times 10^{-3}$ | 0.1110 | 99.975 | 2 |
| WHT-PRO | $8.77 \times 10^{-4}$ | 0.1260 | 99.975 | 2 |

We note that the product rule gives the best result in a Multi-modal system (2D with 3D-MC with 3D-GC) which is EER= 3.75 x 10-4, ROR= 100 (%), and RPR = 1.

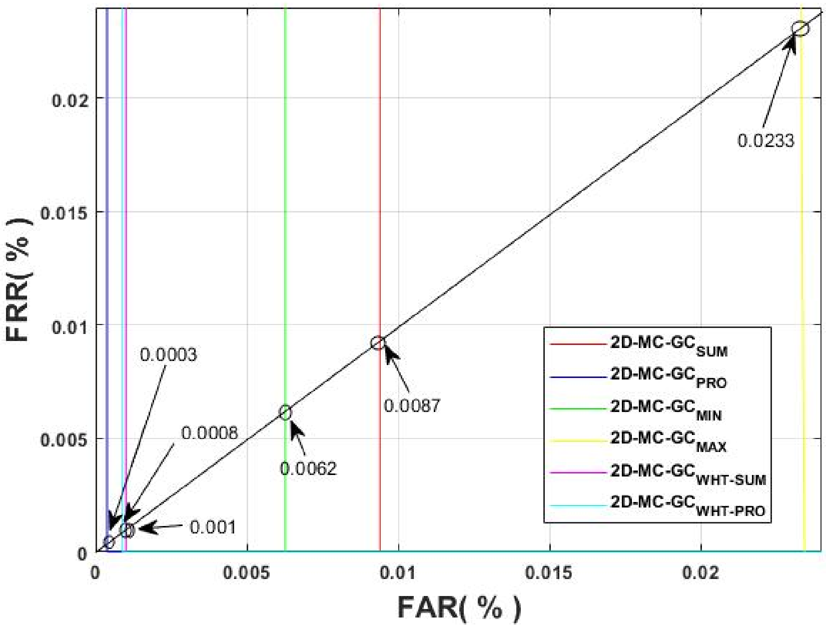


Figure 3.10: The curve of ROC (FRR in the function of FAR) of the system multi-modal 2D with 3D-MC with 3D-GC

According to the combination which has been illustrated in Tables 3.3, 3.4, 3.5, and 3.6. We notice that the results have been improved and the smallest identification errors multi-modal were obtained using the 2D combination with 3D-GC (ERR = 0 (%)) with the Min Rule fusion (MIN). This performance is significantly higher as compared to the case when either 2D or 3D-GC palm-print features alone are used.

## 3.9 Conclusion

In this chapter, we made a study of the uni-modal and multi-modal systems for the recognition of people using 2D and 3D palm prints based on deep learning. According to studies carried out on an alexNet pre-trained network, the results obtained, that for the uni-modal system, the best results obtained for the open set were EER = 0.0005 ( %) and for the closed set ROR = 99.92 ( %) for the 2D spectral band. For the multi-modal system, the best results were found for the combination 2D with 3D-GC with the rule MIN and EER = 0.00 (%) in the open set and ROR = 100 (%) in the closed set.

# General conclusion

Palm-print has been widely studied for its high accuracy and low cost. Two-dimensional (2D) palm-print recognition has been well studied in the past decade, and recently three dimensional (3D) palm-print recognition techniques were also proposed. The 2D and 3D palm-print data can be captured simultaneously and they provide different and complementary information. Using the deep region of interest (ROI) and feature extraction models for palm-print verification, an approach is proposed where convolutional neural networks (CNNs) along with transfer learning are exploited.

In this work, we proposed to jointly use palm-print 2D and 3D for personal authentication or identification. We explored a technique for palm-print-based bio-metrics: 2D and 3D palm-print recognition. A 2D and 3D palm-print database with 8000 samples from 200 individuals (400 palms) was established, on which a series of identification experiments were performed. After the 2D and 3D palm-print image was captured, the region of interest (ROI) was extracted to roughly align the palm and remove the unnecessary cloud points. Then developed the curvature-based feature extraction algorithms to obtain the Mean Curvature Image (MCI), Gaussian Curvature Image (GCI) features. Then a module applied to extract the features of the palm, and after that, a matching module was used which is a transfer learning classifier. We utilized a pre-trained network (AlexNet) to obtain a discriminative feature vector. Moreover, to find the best technique, we utilized two bio-metric systems, the uni-modal and multi-modal bio-metric system. In a uni-modal bio-metric system, we used each type separately (2D, 3D-MC, and 3D-GC separated). At last, a multi-modal bio-metric system by score level fusion strategy of the types of features (2D with 3D-MC, 2D with 3D-GC, 3D-MC with 3D-GC, and at least all the types together 2D with 3D-MC with 3D-GC) was used to classify the palm prints with several fusion rules (Sum, Prod, Min, Max, weighted Sum, and Weighted Prod).

The experimental results show that both 2D and 3D palm-print types can achieve a good recognition rate (0.0005(%) for EER and 99.92(%) for ROR), and fusing them can get much higher performance (0(%) for EER and 100(%) for ROR), and it's the ideal results of recognition rate. In the future, more advanced and powerful feature extraction and matching techniques are to be developed for better recognition performance.

# REFERENCES

[1] ALOUL, Fadi A. The need for effective information security awareness. Journal of advances in information technology, 2012, vol. 3, no 3, p. 176-183.
[2] NGO, David Chek Ling, TEOH, Andrew Beng Jin, et HU, Jiankun (ed.). Biometric security.Cambridge Scholars Publishing, 2015.
[3] YAN, Haibin et LU, Jiwen. Facial Kinship Verification. Springer Singapore:, 2017.
[4] BHATTACHARYA, Samayita et MALI, Kalyani. Fingerprint recognition using minutiae extraction method. Proc. of International Conferance on Emerging Technologies (ICET-2011): International Journal of Electrical Engineering and Embedded Systems, 2011, p. 0975-4830.
[5] GREGORY, Peter et SIMON, Michael A. Biometrics for dummies. John Wiley Sons, 2008.
[6] ZHANG, David, GUO, Zhenhua, et GONG, Yazhuo. Multispectral biometrics: systems and applications. Springer, 2015.

LI, Stan Z. Encyclopedia of Biometrics: I-Z. Springer Science Business Media, 2009.
[8] ZHANG, David D., et al. Palmprint authentication. Springer Science Business Media, 2004.
[9] CHAARI, Anis. Nouvelle approche d'identification dans les bases de données biométriques basée sur une classification non supervisée. 2009. Thèse de doctorat. Université d'Evry-Val d'Essonne.
[10] JAIN, Anil K., ROSS, Arun, et PRABHAKAR, Salil. An introduction to biometric recognition. IEEE Transactions on circuits and systems for video technology, 2004, vol. 14, no 1, p. 4-20.
[11] JABIN, Suraiya et ZAREEN, Farhana Javed. Biometric signature verification.International Journal of Biometrics, 2015, vol. 7, no 2, p. 97-118.
[12] JAIN, Anil K., FLYNN, Patrick, et ROSS, Arun A. (ed.). Handbook of biometrics.Springer Science Business Media, 2007.
[13] VACCA, John R. Biometric technologies and verification systems. Elsevier, 2007.
[14] JAIN, Anil K., BOLLE, Ruud, et PANKANTI, Sharath (ed.). Biometrics: personal identification in networked society. Springer Science Business Media, 2006.
[15] JAIN, Rubal et KANT, Chander. Attacks on biometric systems: an overview. International Journal of Advances in Scientific Research, 2015, vol. 1, no 07, p. 283-288.
[16] LI, Stan Z. Encyclopedia of Biometrics: I-Z. Springer Science Business Media, 2009.
[17] EL-SAYED, AYMAN. Multi-biometric systems: a state of the art survey and research directions. IJACSA) International Journal of Advanced Computer Science and Applications, 2015, vol. 6.
[18] JAIN, Anil K. Biometric recognition: how do I know who you are?.In :International Conference on Image Analysis and Processing. Springer, Berlin, Heidelberg, 2005. p. 19-26.
[19] ROSS, Arun A., NANDAKUMAR, Karthik, et JAIN, Anil K. Handbook of multibiometrics. Springer Science Business Media, 2006.
[20] AMIRTHALINGAM, Gandhimathi et RADHAMANI, G. A multimodal approach for face and ear biometric system. International Journal of Computer Science Issues (IJCSI), 2013, vol. 10, no 5, p. 234.
[21] KIM, Phil. Matlab deep learning. With machine learning, neural networks and artificial intelligence, 2017, vol. 130, p. 21.
[22] MUELLER, John Paul et MASSARON, Luca. Deep Learning For Dummies. John Wiley amp; Sons, 2019.
[23] AGGRAWAL, C. C. Neural Networks and Deep Learning: A Textbook. 2018.
[24] HEATON, Jeff. Artificial Intelligence for Humans, Volume 3: Neural Networks and Deep Learning, 1.0.
[25] VASILEV, Ivan, SLATER, Daniel, SPACAGNA, Gianmario, et al. Python Deep Learning: Exploring deeplearning techniques and neural network architectures with PyTorch, Keras, and TensorFlow. Packt PublishingLtd, 2019.
[26] Tree-based Convolutional Neural Networks, Principles and Applications (Lili Mou, Zhi Jin ).
[27] Computer Vision Using Deep Learning Neural Network Architectures with Python and
[28] JIN, Kyong Hwan, MCCANN, Michael T., FROUSTEY, Emmanuel, et al. Deep convolutional neural networkfor inverse problems in imaging. IEEE Transactions on Image Processing, 2017, vol. 26, no 9, p. 4509-4522.
[29] WANI, M. Arif, BHAT, Farooq Ahmad, AFZAL, Saduf, et al. Advances in deep learning.

Springer, 2020.
[30] PAOLETTI, Mercedes E., HAUT, Juan M., PLAZA, Javier, et al. Deepamp;dense convolutional neural network for hyperspectral image classification. Remote Sensing, 2018, vol. 10, no 9, p. 1454.
[31] MICHELUCCI, Umberto. Advanced applied deep learning: convolutional neural networks and object
[32] GAD, Ahmed Fawzy, GAD, Ahmed Fawzy, et JOHN, Suresh. Practical computer vision applications using
[33] BURKOV, Andriy. The hundred-page machine learning book.Canada : Andriy Burkov, 2019.

## الملخص

في الوقت الحاضر، تعد القياسات الحيوية حلاً للعديد من مشاكل التعرف التلقائي على الأفراد. في هذه الوثيقة، قدمنا اقتراحًا لأنظمة القياسات الحيوية لتحديد الهوية أحادية الوسائط ومتعددة الوسائط وكيف أن نظام القياسات الحيوية متعدد الوسائط أفضل من نظام أحادي الوسائط. الشبكات العصبية التلافيفية (CNNs) هي بنية خاصة للتعلم العميق. أحد الاستخدامات الأكثر شيوعًا لهذه البنية هو تصنيف الصور. سنقوم بتجربة نظام قياس بيولوجي أحادي الوسائط ومتعدد الوسائط يعتمد على بصمة الكف باستخدام بنية CNN ونقل التعلم. لذلك، استخدمنا قاعدة بيانات PolyU التي تحتوي على صور ثنائية وثلاثية الأبعاد، وسيتم تحويل الأخيرة من خلال وظيفتين تعنيان الانحناء (MC) والانحناء الغاوسي (GC). النتائج التي تم الحصول عليها مثالية في الغالب للحالة متعددة الوسائط.



**الكلمات الدالة:**

القياسات الحيوية، بصمة الكف ثنائية / ثالثية ألبعاد، تحديد الهوية، التعلم العميق، نقل التعلم، CNN ،أحادي الوسائط، متعدد الوسائط.



## Abstract

Nowadays, biometrics is a solution to several problems of automatic identification of individuals. In this document, we have given a proposition of efficient Unimodal and Multimodal identification biometric systems and how a multimodal biometric system is better than a unimodal one. Convolutional Neural Networks (CNNs) is a special architecture for deep learning. One of the most common uses of this architecture is for image classification. We will be piloting a unimodal and multimodal palmprint-based biometric system using CNN architecture and transfer learning. For that, we used the PolyU database containing 2D and 3D images, the latter will be transformed by two functions mean curvature (MC) and Gaussian curvature (GC). The obtained findings are perfect predominately for the multimodal case.



**Keywords :**

Biometrics, 2D / 3D palmprint, identification, deep learning, transfer learning, CNN, unimodal, multimodal.



## Résumé

De nos jours, la biométrie est une solution à plusieurs problèmes d'identification automatique des individus. Dans ce document, nous avons donné une proposition d'un système biométrique d'identification unimodal et multimodal efficace et comment un système biométrique multimodal est meilleur qu'un unimodal. Les réseaux de neurones convolutifs (CNN) sont une architecture spéciale pour l'apprentissage en profondeur. L'une des utilisations les plus courantes de cette architecture est la classification d'images. Nous pilotons un système biométrique basé sur l'empreinte palmaire unimodale et multimodale utilisant l'architecture CNN et l'apprentissage par transfert. Pour cela, nous avons utilisé la base de données PolyU contenant des images 2D et 3D, ces dernières seront transformées par deux fonctions courbure moyenne (MC) et courbure de Gauss (GC). Les résultats obtenus sont parfaits principalement pour le cas multimodal.



**Mots clés :**

Biométrie, empreinte palmaire 2D/3D, identification, apprentissage profond, apprentissage par transfert, CNN, unimodal, multimodal.