N^o d'ordre :.../.....

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY of KASDI MERBAH OUARGLA

FACULTY OF NEW INFORMATION AND COMMUNICATION TECHNOLOGIES

THESIS

Thesis submitted in partial fulfillment of the requirements for the degree of

3rd Cycle LMD Doctorate

In : Communication and Signal Processing

By : **RACHID CHLAOUA**

Theme

Combination of Multiple Biometrics for Recognition of Persons

Publicly sustained on : 08/07/2019 before the jury composed of:

Name and Surname	***Title***	***Affiliation***	***Quality***
F. Z. LAALLAM	*Professor*	*Univ. K. M. Ouargla*	*President*
A. MERAOUMIA	*MCA*	*Univ. L. T. Tebessa*	*Thesis Director*
K. E. AIADI	*Professor*	*Univ. K. M. Ouargla*	*Thesis Co-director*
M. AMROUNE	*MCA*	*Univ. L. T. Tebessa*	*Examiner*
M. BOULESBAA	*MCA*	*Univ. K. M. Ouargla*	*Examiner*
M. L. KHERFI	*MCA*	*Univ. K. M. Ouargla*	*Examiner*
D. SAMAI	*MCA*	*Univ. K. M. Ouargla*	*Invited*


بسم الله الرحمن الرحيم


وقل رب زدني علما


صدق الله العظيم

A mes parents pour m'avoir toujours et jusqu'à ce jour soutenu et encouragé à aller le plus loin possible dans mes études. Que dieu les protège

A mes frères et mes sœurs que je remercie très fort et à qui je souhaite tout le bonheur et la réussité,

A tous mes amis,

A tous ceux que j'aime et qui m'aiment, où qu'ils soient,

A tous ceux qui m'ont aidé et encouragé

RACHID CHLAOUA

Remerciements

Je remercie tout d'abord `DIEU` *le tout puissant, de m'avoir permis d'atteindre ce modeste niveau scientifique et de m'avoir donné le courage et la patience afin de mener à bien le travail réalisé dans cette thèse.*

Mes sincères remerciements et ma profonde gratitude à M. ***Abdallah MERAOUMIA****, mon directeur de thèse, MCA à l'université des Larbi Tebessi à Tebessa, pour son appui scientifique, sa disponibilité, ses orientations judicieuses et dont les compétences intellectuelles, l'expérience, la modestie et la patience ont grandement contribué à l'aboutissement de cette thèse. Qu'il trouve, ici, l'expression de mon profond respect.*

Je tiens aussi à remercier M. ***Kamal Eddine AIADI****, mon Co-Encadreur de thèse, Professeur et directeur de laboratoire LAGE à l'université de Kasdi Merbah à Ouargla, pour son suivi permanent, sa disponibilité et ses conseils pratiques qui m'ont aidé à la thèse.*

J'adresse aussi mes sincères remerciements aux membres du juré: Mme. ***F. Z. LAALLAM****, M.* ***M. AMROUNE****, M.* ***M. BOULESBAA*** *et M.* ***M. L. KHERFI*** *d'être accepter de jurer ce travaille et M.* ***D. SAMAI*** *d'être accepter mon invitation et sans oublier tous les enseignants du département d'électronique et des communications de l'université Kasdi Merbah à Ouargla pour leurs aides et leurs conseils, sans oublier tout mes collègues et mes amis.*

Enfin, mes remerciements vont à tous ceux qui ont contribué d'une quelconque manière à l'aboutissement de ce travail.

Résumé

La sécurité des informations garantit que les utilisateurs autorisés les seuls peuvent accéder au contenu requis, ce qui implique la confidentialité des informations échangées. La reconnaissance de l'identité d'une personne est un moyen de garantir cet objectif. En effet, le très grand besoin de cette reconnaissance exige à l'homme de mettre des moyens qui sont liées à des informations qu'un individu possède ou connait. Cependant, pour contourner les limitations associées à de tels moyens, d'autres moyens de sécurité ont été développés qui permet d'utiliser une information propre à chaque personne. Cette nouvelle façon de reconnaissance est la biométrie.

La technologie biométrique a suscité une grande attention ces dernières années. Dans les systèmes de sécurité biométriques, la reconnaissance de l'identité personnelle dépend de leurs caractéristiques comportementales, biologiques ou physiques. Actuellement, certaine technologies biométriques sont développées et l'un des traits biométriques les plus populaires est FKP (Finger-Knuckle-Print) en raison de sa convivialité et de son faible coût. Cette thése présente une nouvelle approche dans laquelle l'apprentissage en profondeur est appliqué pour créer un système biométrique multimodal basé sur des images de modalités FKP qui extraient leurs caractéristiques par réseaux de PCANet (Principal Component Analysis Network) et DCTNet (Discrete Cosine Transform Network). Dans la structure proposée, PCA/DCT est utilisée pour apprendre des banques de filtres en deux étapes, suivies d'histogrammes de hachage binaires simples et de blocs pour la mise en cluster au niveau de vecteurs caractéristiques, qui est adoptée en tant qu'entrée pour la classification. Ces classificateurs SVM (Support Vector Machine) et KNN (K-Nearest Neighbor) sont utilisés pour les fonctionnalités PCANet et DCTNet, respectivement. Pour améliorer les taux de reconnaissance, le système biométrique multimodal a été généré par un schéma de fusion au niveau de score. En utilisant une base de données FKP disponible, nous avons mené une série d'expériences d'identification et les résultats obtenus montrent que la conception de notre système d'identification permet d'obtenir un excellent taux de reconnaissance et une capacité anti-fraude élevée.

Abstract

The security of information is ensuring that the only authorized users are able to access the required contents, thereby entails confidentiality of exchange information. The recognition of the person identity is one means to ensure this purpose. In fact, due to the great need for such recognition, man has developed several ways that are related to information's that a person has or knows. However, to overcome the limitations associated with such traditional means, other means of security has been developed that allow obtaining the specific information of the person. It is the biometrics-based recognition.

Biometric technology has attracted a great attention in recent years. In the biometric security systems, the personal identity recognition depends on their behavioral, biological or physical characteristics. Currently, a number of biometrics technologies are developed and one of the most popular biometric trait is Finger-Knuckle-Print (FKP) due to the user-friendly and the low cost. This thesis presents a new approach, where the simple deep learning is applied to create a multi-modal biometric system based on images of FKP modalities which extracted their features by Principal Component Analysis Network (PCANet) and Discrete Cosine Transform Network (DCTNet). In the proposed structure, PCA or DCT is employed to learn two-stage of filter banks followed by simple binary hashing and block histograms for clustering at feature vectors, which is adopt as input for classification. Thus, the Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifiers are used for the PCANet and DCTNet features, respectively. To improve the recognition rates, a multimodal biometric system based on matching score level fusion scheme was generated. Using an available FKP database, we conducted a series of identification experiments and the obtained results show that the design of our identification system achieves an excellent recognition rate and having a high anti-counterfeiting capability.

Contents

List of Figures

List of Tables

Abbreviations

AI	:	Artificial Intelligence	AMD	:	Assembled Matrix Distance
ATM	:	Automated Teller Machine	BDCT	:	Block based Discrete Cosine Transform
BLPOC	:	Band-Limited Phase-Only Correlation	BM	:	Boltzmann Machine
CMC	:	Cumulative Match Curve	CNN	:	Convolutional Neural Networks
CPU	:	Central Processing Units	DBN	:	Deep Belief Networks
DCT	:	Discrete Cosine Transform	DCTNET	:	Discrete Cosine Transform Networks
DFT	:	Discrete Fourier Transform	DIP	:	Distal Inter Phalangeal
DNA	:	Desoxyribonucleic Acid	DWT	:	Discrete Wavelet Transform
EER	:	Equal Error Rate	EM	:	Expectation-Maximization
FA	:	False Acceptation	FAR	:	False Acceptance Rate
FERET	:	Face Recognition Technology	FKP	:	Finger Knuckle Print
FR	:	False Rejection	FRR	:	False Rejection Rate
FTC	:	Failure To Capture	FTE	:	Failure To Enrol
GAR	:	Genuine Acceptance Rate	GPU	:	Graphics Processing Units
ICA	:	Independent Component Analysis	IBG	:	International Biometric Group
KNN	:	K-Nearest Neighbor	LDA	:	Linear Discriminant Analysis
LIF	:	Left Index Fingers	LMF	:	Left Middle Fingers
LP	:	Laplacian Pyramid	MACE	:	Minimum Average Correlation
MCI	:	Mean Curvature Image	MVN	:	Multivariate Normal Density
NIR	:	Near Infra-Red	NIST	:	National Institute of Standards Technology
PCA	:	Principal Component Analysis	PCANET	:	Principal Component Analysis Networks
PIN	:	Personal Identification Number	PLM	:	Palmprint
RBF	:	Radial Basis Function	RBM	:	Restricted Boltzmann Machine
RFT	:	Random Forest Transform	RIF	:	Right Index Fingers
RMF	:	Right Middle Fingers	ROC	:	Receiver Operating Curve
ROI	:	Region Of Interest	ROR	:	Rank One Recognition
RPR	:	Rank of Perfect Recognition	SIFT	:	Scale Invariant Feature Transform
ST	:	Surface Types	SVM	:	Support Vector Machine
TA	:	True Acceptance	TR	:	True Rejection

Chapter 1

INTRODUCTION

THE security of information is ensuring that the only authorized users are able to access the required contents, thereby entails confidentiality of exchange information. It is essentially provided by authentication user whereby an individual identity is verified through traditional means such as: username and password, smart-cards, keys, etc. With increasing adoption of technologies in the world, the conventional authentication methods no longer comply with the stringent authentication requirements [1].

The problem of information security is the main concern of users in all organizations and governments. For ensuring that, they focus their attention to create ways are able to protect the contents available in digital media, especially. Content owners are losing billions of dollars annually in revenue due to the illegal copying and sharing of information. In order to address this growing problem, authentication systems are being deployed to regulate the duplication and dissemination of content [2]. The critical component of this system is user authentication which determines whether a certain individual is indeed authorized to access the content available in a particular digital medium. In a generic cryptographic system, the user authentication method is based on [2]:

- **Using an acquaintance:** The first way to verify or determined the identity of an individual is based on the knowledges of the user "What user knows?" [2], these are usually a password, PIN (Personal Identification Number), etc. For instance, most passwords are so simple, that they can be easily guessed or broken by simple dictionary attacks. Simple passwords are easy to guess and, thus, compromise security; complex passwords are difficult to remember and, thus, are expensive to maintain. Some users tend to "store" complex passwords at easily accessible locations. Furthermore, most people use the same password across different applications; an impostor upon determining a single password can now access multiple applications. This type of security is used to access online services, building, computer and network, etc.

- **Using a possession:** The second way is based on the possession of a "What user has?" [2] as a smart card, badge, document, key. However, for the token based method, these physical elements can be easily used by other person or stolen, lost or falsified. So some users are able to falsify the identity of the legitimate person and to trick a system.
- **Using a biometric:** In contrast to both of these methods, biometrics appears "What user is?" to be a solution to overcome restrictions of conventional authentication methods and biometric authentication cannot be forgotten or lost [2]. Furthermore, in the authentication process, providing biometrics to the system is the proof of the claimant's presence. Unlike the password or ATM card, a biometrics is more difficult to copy or to falsify. Additionally, a biometrics can be combined with password or/and an ATM card to form two or more authentication factors. By doing so, the authentication rate can be further enhanced without having to replace these existing systems [1, 2].

Biometrics-based personal authentication systems have recently gained intensive research interest due to the unreliability and inconvenience of traditional authentication systems. Biometrics recently became a vital component of any effective person identification solutions as biometric traits cannot be forged, shared, lost, duplicated, stolen or even forgotten [2, 3, 4]. The following table (Table. 1.1) presents comparison between different traditional authentication and biometric.

Methods	Lost	Steal	Copy	Forget
ID/Badge	Yes	Yes	Yes	Yes
U/Password	Non	Non	Yes	Yes
Key	Yes	Yes	Yes	Yes
Biometric	Non	Non	Non	Non

TABLE 1.1: Comparison between different traditional authentication and biometric.

1.1 Biometric Security

Biometric security is one of the best primary functions of security system which is based on the measurement and statistical analysis of persons physical or behavioral characteristics, such as face, fingerprint, hand geometry, iris, DNA, signature, voice, etc [5, 6], that can be used for automated recognition. Biometric technologies offer several advantages over traditional authentication schemes and they have more reliable characteristics than other methods. One of the advantages of biometric methods is the requirement of the person to

be presented at the time of authentication. Furthermore, it is difficult to attack biometric systems because it requires more time, money, experience and it is unlikely for a user to repudiate having accessed the content using biometrics. Thus, a biometrics based authentication scheme is a powerful alternative into security systems because has a several reasons for uses such as [7]:

1. High security and difficult to fraud.
2. Comfort by replacing traditional methods.
3. Security and confidence of biometric authentication.

For instance, biometrics can be used in conjunction with passwords to enhance the security offered by the authentication system. All biometric technologies differ according to security level, social acceptability, reliability, cost, performance, etc. The most acceptable biometric technologies are those extracted from hand due to their higher discriminatory which based on the fact that each human hand is highly unique. In our days, the physical technology, which has been attracting much attention, is the Finger-Knuckle-Print (FKP) modality [8]. In this modality, several features can be used such as structure and shape (*e.g.* length, width and thickness, joints); characteristics of the skin surface such as creases and ridges in fingers. There are several main factors which make a person's FKP unique, among these factors, the ease of the use as well as the rich texture information seems to be the most significant, which makes the biometric system work with higher accuracy.

1.2 Objective and Contributions

Biometric is a potential technology powerful to meet the needs of privacy and security of the public information. Indeed, among the methods found in the literature that are used to the identification of people, those based on the finger. For several years, the fingerprint has been key feature of identification systems. But recently, the identification based on the Finger-Knuckle-Print (FKP) motivated a large number researchers. In fact, biometric systems based on the FKP, are accepted in access control applications. On the other hand, the finger joint is better adapted in these cases, as these systems are now considered appropriate because they do not cause users anxiety.

In pattern recognition system, feature extraction is the process by which key features are selected or improved for the sample (generally sample is in form of image). In biometrics systems, as one of pattern recognition applications, the feature extraction process is

based on a set of algorithms; the method varies depending on the type of image characteristics such as texture, line and shape. In recent years, high hopes were invested in biometrics systems, mostly thanks to modern algorithms like machine learning and artificial intelligence.

An overview of recent publications of the last decade does not show much of works that are based on deep learning feature extraction for recognition systems. Most of the systems developed are based on the feature extraction of this modality, at the external level from the finger image. However, these fingers can be combined to improve the performance of the biometric system.

The objective of this thesis is realized a multimodal biometric system that uses a biometric modalities of the finger joints. The main goal is trying to design a biometric system based on new machine learning methods which are deep architectures. The major motivation behind this objective is the hierarchical learning structures of this method which have powerful ingredients to create a more sophisticated identification system dedicated to image classification, which could serve as a use to train and to adapt with different traits biometrics. This work will be to provide answers to essential research questions, such as: To what extent does image quality affect the image classifications generated by deep learning models? Are deep model architectures more robust than others classical methods against various challenges? How should image descriptors be computed? Answers to those and similar questions are in our opinion crucial for a better understanding of deep learning-based biometric recognition.

To achieve this goal, firstly, we choose a best classifier and fusion strategy. On the other hand, we propose to use several methods of classification of the biometric characteristics (algorithms) for each finger. In our work, we propose a multimodal biometric system, where information from the different FKP modalities are fused through matching score level to improve the identification rate. In this thesis, we propose a method to increase the performance of a multimodal biometric security system. The main contribution lies in the efficient consolidation of information obtained from using the deep learning processing for features extraction. The detailed contributions of this thesis are summarized below:

- In this doctoral thesis, we develop a multimodal biometric system based on FKP biometric traits to meet the recent extensive security requirements for high performance. This system can alleviate most of the drawbacks associated with unimodal biometric systems.
- The main feature of multimodal biometric system is information fusion that is, what information needs to be consolidated and how? Thus, in this doctoral research, we

use match score level fusion which is relatively most successful approach to combine multimodal biometric information.

- In order to increase the level of confidence, we employ PCA-Network and DCT-Network architectures allow to extract deep features for biometric authentication. Further, more improvement in terms of performance level and the outcomes can be obtained through these methods.
- To demonstrate the advantages of proposed methodology over other multimodal biometric systems, we test our system by the FKP database to compare the results with various FKP recognition systems.

1.3 Organization of Thesis

The thesis has been structured as follows. Chapter 2 describes a general concepts, history and applications of biometric and it also gives a detailed idea about the different biometric technologies. After that, the description of biometric systems and the functionality principle of each biometric system are presented. Expressing the performance of a biometric system requires description of some parameters in this chapter. The end of chapter highlighted at shortcomings and imperfections in biometric recognition system based on single source of biometric information.

The Chapter 3 is devoted to the presentation of multimodal systems and the necessity of them, moreover we present the principles data fusion and different scenarios of multimodal biometric systems. In this chapter, the fusion strategies as well as the different levels of fusion are described with give researches for that. This chapter also discusses on the design issues which involved in multimodal system development process and challenges.

Chapter 4 describes the concept of biometric features, and different types of features such as: textures, lines and shapes. Also, it gives a detailed idea about the feature extraction methods. Moreover, this chapter presents the description of classical and deep learning methods as gives the principle challenges to developing of traditional algorithms to solve these limitations of classical methods requires uses of deep learning and to benefit its strengths. The last of chapter, we justify our choice to the simple deep learning methods and their hierarchy structures, where can use for artificial intelligence applications.

FKP is the most common biometric identifier and it is used by most of the biometric researchers for identity authentication. In Chapter 5, the proposed biometric methodology is illustrated in system based on FKP technology. All different processes and proposed fusion

strategy for our multimodal system are also described in more detail in this chapter.

The efficiency of the proposed biometric identification system were tested in Chapter 6. For that, Chapter 6 shows the outcomes of the experiments performed on FKP database frameworks. Also, this chapter include an experimental setup for selection of parameters and adapt our algorithms. After that, the results of uni-modal and multimodal system are presented. Through our experimental results and literature researches, we made comparative studies between classical methods and deep learning methods, which can be demonstrated the feasibility and effectiveness of our proposed biometric systems.

Finally, Chapter 7 summarizes the thesis and the contribution and presents some concluding remarks. Possible future directions of this research are also discussed in this chapter.

Chapter 2

BIOMETRICS TECHNOLOGIES OVERVIEW

2.1 Biometrics Overview, History and Applications

BIOMETRIC is derived from the Greek words, composed by two parts: "Bios" meaning life and "metros" meaning measures [9], biometric system is defined as "a system which automatically distinguishes and recognizes a person as individual and unique through a combination of hardware and pattern recognition algorithms based on certain physiological or behavioral characteristics that are inherent to that person" [10].

Using parts of the human body as a mean to identity authentication goes back to very old times. It is reported two thousand years ago that in ancient Babylon, merchants sealed deals with fingerprints on clay tablets to record their trading transactions [11]. The Chinese in the 3rd century B.C. used thumbprints and fingerprints on clay tablets as signatures to seal the official documents. While in the 14th century A.D., various official document papers dated in Persia bore fingerprint impressions [12, 13].

A systematic and scientific basis for human identification started in the 19th century when a French police officer, Alphonse Bertillon [14] invented a number of anthropomorphic measurements, called Bertillonage, for identifying criminals. His system was built on the assumption that the body of people do not change in basic characteristics. Bertillon's system involved measuring five primary measurements of body parts such as head length; head breadth; length of the middle finger and the length from elbow to end of middle finger (see Fig. 2.1). Afterward, every major heading was additionally classified into three categories of: small, medium and large. The length of the little finger and the eye color were also recorded.

FIGURE 2.1: Bertillonage or anthopometric measurements (from [14]).

Biometric system is essentially a pattern recognition system which makes a personal identification decision by measuring the specific physiological or behavioral characteristics. These are usually presented by the user when comparing biometric features with the stored feature of user. Biometric is a constantly growing technology which has been widely used in many applications. It can help to make operations, transactions and everyday life both safer and more convenient like what we see today. Biometrics have been extensively applied in various fields not just to identify a criminal. According to International Biometric Group (IBG), the biometrics worldwide market was expected to expand to a very high values by next years. The usage of biometrics not only has been driven by the public sector, the private sector has also increasingly shown its interest in such applications. These applications of biometrics are illustrated in the Fig. 2.2 and can be divided into four main groups:

1. **Government applications:** The key application for biometric technology is helping automatic control process such as national identity card, driver's permit, passport, border control, airports, etc.

2. **Justice/Law applications:** Biometric technology and law enforcement have many important identity management such as body identification, criminal investigation, terrorist identification, etc.

3. **Logical/Physical Access applications:** Major area of biometric technology application. Whether, it's securing the apps on your smartphone, computer and network access, or home, car, hotel, etc.

4. **Commercial applications:** Such as credit card, bank account, cell phone, medical registry management, distance learning, etc.

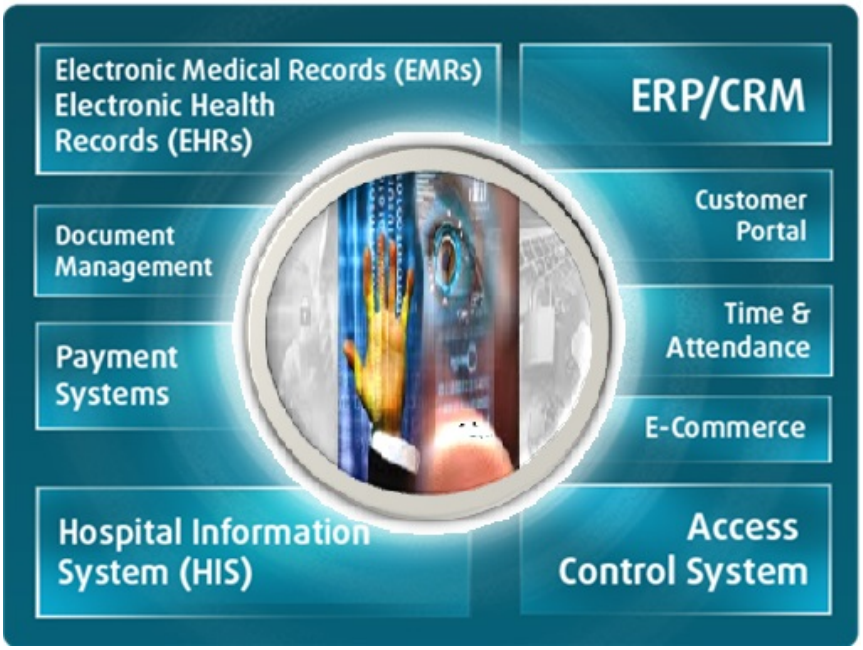


FIGURE 2.2: The applications of biometrics.

2.2 Classification of Biometric Modalities

The goal of this section is to introduce the variety of different biometric traits, by discussing the principles of acquiring biometric information from humans, and enabling the reader to identify these concepts for each of the traits introduced. Biometric modalities used different scientific references for classified, and can be evaluated into two groups:

- Intrusive techniques: These techniques require physical contact with the individual to recognition, such as fingerprints, palm prints or the hand geometry, etc.

Biometrics	Universality	Uniqueness	Permanence	Collectability	Performance	Acceptability	Circumvention
Face	High	Low	Med	High	Low	High	Low
Finger print	Med	High	High	Med	High	Med	High
Hand geometry	Med	Med	Med	High	Med	Med	Med
Keystrokes	Low	Low	Low	Med	Low	Med	Med
Hand veins	Med	Med	Med	Med	Med	Med	High
Iris	High	High	High	Med	High	Low	High
Retinal scan	High	High	Med	Low	High	Low	High
Signature	Low	Low	Low	High	Low	High	Low
Voice	Med	Low	Low	Med	Low	High	Low
Facial thermograph	High	High	Low	High	Med	High	High
Odor	High	High	High	Low	Low	Med	Low
DNA	High	High	High	Low	High	Low	Low
Gait	Med	Low	Low	High	Low	High	Med
Ear Canal	Med	Med	Low	Med	Med	High	Med

TABLE 2.1: Comparison of biometric technologies [3].

- Non intrusive techniques: These techniques do not require the cooperation of the individual. Their application can be done remotely using sensors that do not require direct contact with the user.

Further, we are pointed out that one of the basic challenges in research on biometrics is finding adequate modalities, fulfilling the main aspects of ascertainability. There are seven factors defined by Jain, Bolle, and Pankanti [15] that determine the suitability of a physical or a behavioral trait to be used in a biometric application. From [3], the Table. 2.1 presents a brief comparison of the physiological and behavioral biometric techniques based on these seven factors described:

* **Universality:** each person accessing the application should possess the trait.

* **Uniqueness:** the given trait should be sufficiently different across individuals comprising the population.

* **Permanence:** the characteristic should be sufficiently invariant with respect to the matching criterion over a period of time.

* **Collectability:** the characteristic should be measured quantitatively.

* **Performance:** the recognition accuracy and the resources required to achieve that accuracy should meet the constraints imposed by the application.

* **Acceptability:** individuals in the target population that will use the application should be willing to present their biometric trait to the system.

* **Circumvention:** this reflects how easily the system can be fooled using fraudulent methods.

The choice of different biometric characteristics depends on that particular application scenario. There exists three principles for obtaining information about personal traits for measurement of biometrics.

2.2.1 Physiological Biometric

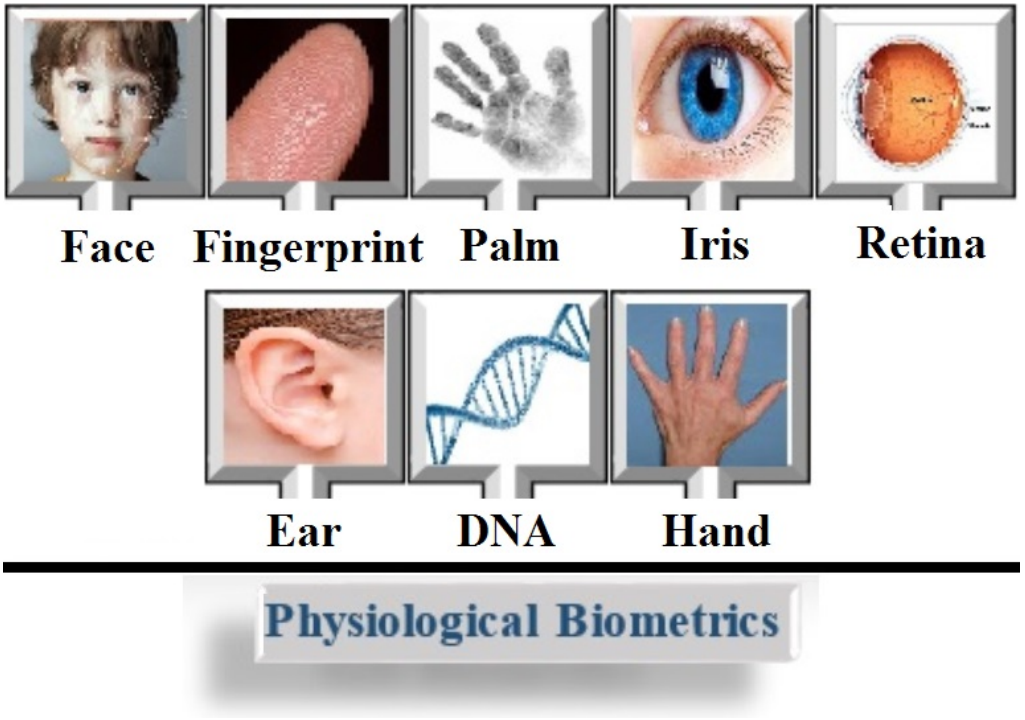


FIGURE 2.3: Various physiological biometric characteristics.

Physiological biometrics refer to physical measurements of the human body, including face, fingerprint, hand-geometry, ear, iris, retina, DNA, palmprint, hand and finger vein etc. The recognition system based on physiological characteristics has a relatively high accuracy. In addition, the use of physiological biometrics introduce reliable identification since the body characteristics are irreplaceable. Fig. 2.3 shows physiological biometric characteristics which can be used in biometric systems for person authentication. Next we list some example of those biometrics:

a. Fingerprint Recognition:

Fingerprint recognition is the most widely used method of biometric authentication [16]. The technology uses unique features from the fingerprint to develop the template. These features are known as minutiae, which are a combination of ridge bifurcations and ridge endings. The template only uses the information gathered describing the minutiae of the fingerprint and not the entire image of the fingerprint [16]. This is important to note because it is not possible to reconstruct an image of the fingerprint from the information stored in the database.

There are advantages and disadvantages to a fingerprint biometric authentication system. One advantage of fingerprint recognition is that it has a long history of use. In relative terms, the use of fingerprints as an automated authentication tool is new compared to the centuries of manual fingerprinting of individuals for identification. Other advantages include factors such as the ability to use multiple fingers to scan for a template, the fingerprint is permanent and it does not change patterns with age, it is easy to use, and the sensors are inexpensive [11]. The disadvantages of fingerprint recognition include issues with public perceptions about its use such as touching the sensor will spread germs and the scanned image of the fingerprint could be reproduced or used for criminal investigations [11]. Research has also been performed on print quality in elderly individuals, which shows that as people grow older, there is a higher rate of reject rates in sensor recognition.

b. Face Recognition:

Humans have been using facial recognition to identify each other as a part of daily life for centuries. There are two categories of facial recognition: facial appearance and facial geometry [16]. The method of facial appearance is also called the eigenface method because it collects a number of face images that form a two dimensional gray-scale image which in turn produces a biometric template. Facial geometry gathers measurements of the face that do not change over time such as the distance between the eyes, the length and width of the face. In contrast to fingerprint biometrics, there is no contact made in facial recognition biometrics. The disadvantage of this type of biometric is the condition of the environment while obtaining the sample can affect the quality of the image poor lighting, camera quality, and obstructions on the face by the individual requesting access can make a significant difference in the initial enrollment.

c. Hand and Finger Geometry:

Hand geometry recognition systems are based on a number of measurements taken from the human hand, including its shape, size of palm, lengths and widths of the fingers. The geometry of a hand and fingers is not very distinctive, and cannot be used for systems requiring identification of an individual from a large population [15].

d. DNA Recognition:

Among the various known types of biometric traits, deoxyribonucleic acid (DNA) is the most reliable personal identification biometric trait. DNA is the genetic material found in most organisms, including human beings, and remains unchanged during a person's life or even after the death. DNA based identification is the most accurate biometric technology that never fails. DNA can be easily found in the blood, urine or any other liquid that comes out from a human body. The results of a DNA test are very fast and can be obtained within one to two hours. DNA is currently used mostly in the forensic applications for person recognition.

e. Retina Recognition:

In the retinal scan technology retina of an individual is used for his/her identification. Retina is the surface on the back of the eye that processes light entering through the user. The basis of this technology is blood vessel pattern in the retina of the eye, which forms a unique pattern. This blood vessel pattern in the retina of an individual can be used as tamper proof personal identifier. The pattern of the blood vessels is unique and stays the same for a lifetime. However, it requires about 12-15 seconds of careful concentration to take a good scan. It is claimed to be the most secure biometric since it is not easy to change or replicate the retinal vasculature [16].

f. Iris Recognition:

Iris recognition uses the pattern of the iris as a unique identifier. Although the coloration of the iris is found to be genetic, the pattern of the iris results from the development process of the eye during the prenatal stage of growth [11, 16]. A high resolution digital camera is used as the sensor for acquiring the image of the iris. An individual must line his or her eye up within a field of view in order to minimize the amount of noise (i.e., eyelashes, eyelids) in the image. Just as with the facial recognition biometric, there is no physical contact with a sensor. Noise such as eyelids, eyelashes, and contact lenses can decrease the accuracy of the biometric. There is also a negative public misperception that the eye is scanned with a light source, and that it would damage the eye [11]. Although the automated technology is new and consumer education is needed to reduce fears.

g. Ear Recognition:

Another biometric authentication technique is conducted which based on the recognition of the unique shape and appearance of human being ear. Naturally, a person is born with a visual shape of his/her ears. However, human ear is not subject to change while a person's growth and even aging [15]. It is based on matching the distance of salient points on the pinna from a landmark location on the ear. The evidence from study [15] supports the hypothesis that the ear contains unique physiological features.

It has a dependable stability which increases its level of security as a proposed method for the security identification/verification of individuals.

h. Palmprint Recognition:

The palms of human hands contain patterns of ridges and valleys much like the fingerprints. The area of the palm is much larger than the area of a finger and as a result, palmprints are expected to be even more distinctive than the fingerprints [17].

i. Hand and Finger Vein:

The pattern of heat radiated by the human body is a characteristic of an individual and can be captured by an infrared camera [18].

2.2.2 Behavioral Biometric

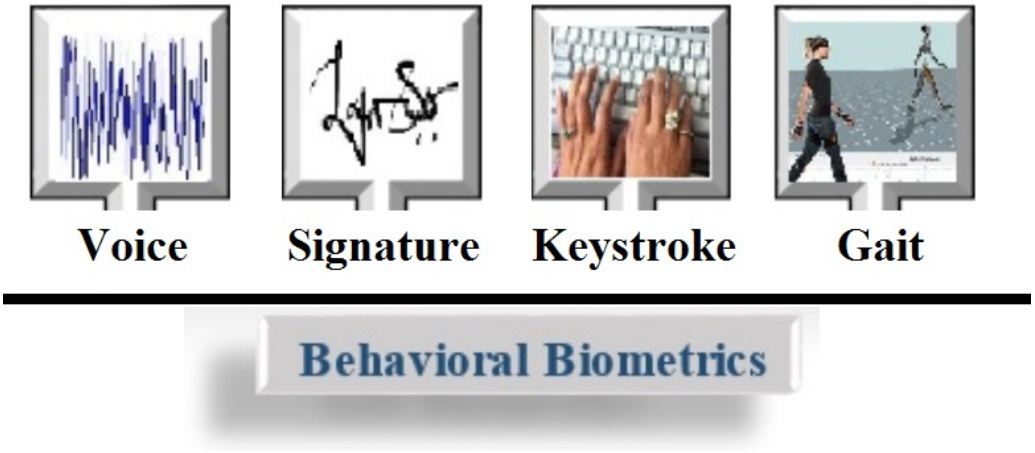


FIGURE 2.4: Various behavioral biometric characteristics.

Behavioral characteristics are based on an action taken by a person. On the other hand, behavioral biometrics are based on measurements and data derived from an action and indirectly measure characteristics of the human body. Fig. 2.4 shows behavioral biometric characteristics which can be used in biometric systems for person authentication. The following are the examples of biometric techniques based on behavioral characteristics:

a. Gait Recognition:

Is the way one walks and is a complex spatio-temporal biometric. Gait is not supposed to be very distinctive, but is sufficiently discriminatory to allow verification in some low-security applications [15].

b. Signature Recognition:

The way a person signs his/her name is known to be a characteristic of that individual. Signatures change over a period of time and are influenced by physical and emotional conditions of the signatories [19].

c. Voice Recognition:

Voice recognition systems use the characteristics of the voice in order to recognize a person. The behavioral part of the speech of a person changes over time due to age, medical conditions, emotional state, etc. Therefore, voice is not very distinctive and may not be appropriate for large-scale identification [20].

d. Keystroke Dynamics:

It is hypothesized that each person types on a keyboard in a characteristic way. It is not unique to each individual but it offers sufficient discriminatory information to permit identity verification [21].

2.2.3 Soft Biometric

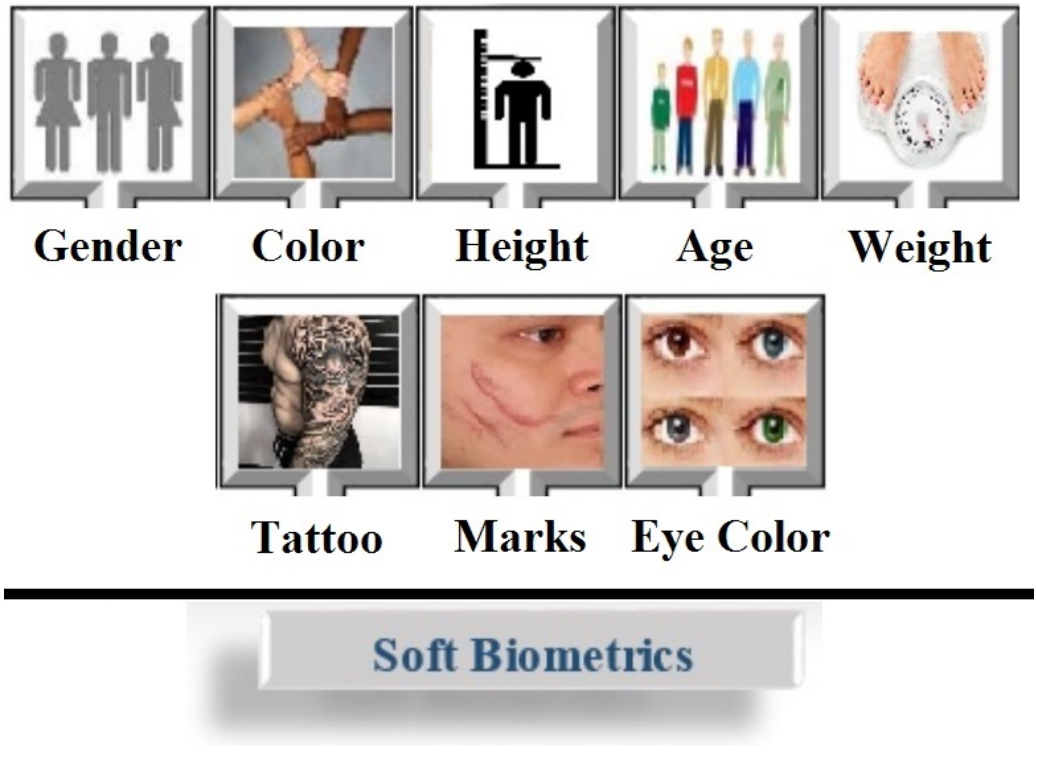


FIGURE 2.5: Various soft biometric characteristics.

Soft biometric characteristics such as: gender, weight, height, color, ethnicity, age, scar, eye color, marks and tattoo, etc; cannot provide reliable user recognition because they are not distinctive and permanent. Recently, soft biometric traits are started to be used in person

recognition along with some physiological or behavioral characteristics and can complement the identity information provided by the primary biometric traits [2]. Fig. 2.5 shows soft biometric characteristics which can be used in biometric systems for person authentication.

2.3 Biometric Systems and Functionalities

A biometric system is basically a pattern recognition system that can recognize a person based on specific features by acquiring biometric data from an individual, extracting a feature set from the acquired data, and comparing this feature set against the template feature set in the database [3]. Thus, biometric system components can be divided into five main modules according to their functionalities.

1. **The sensor module:** represents the interface between the user and the system acquires the biometric data from an individual, through a variety of instruments based on the type of biometrics such as: camera, fingerprint sensor, speaker, etc.

2. **The feature extraction module:** extracts features from the acquired biometric trait, which ideally should be unique for each person. Also it includes quality assurance to determine if the quality of the biometric is good enough to be used in the process. The feature set obtained during enrollment is stored in the system database as a template.

3. **Database unit:** is important component for any biometric system where all the enrolled biometric templates are being stored and where the templates are being retrieved from in the authentication process.

4. **The matching module:** compares the newly acquired biometric template with the template stored in the database and determine the degree of similarity/dissimilarity between the two feature sets.

5. **The authentication decision:** is taken at the decision module based on this degree of similarity/dissimilarity and on decision rules determines either if the presented biometric is a genuine/impostor.

A biometric systems function through the enrolment and recognition phases. However, the recognition can be used for identification and verification modes. Fig. 2.6 illustrates biometric enrolment, biometric verification and identification processes in a biometric system.

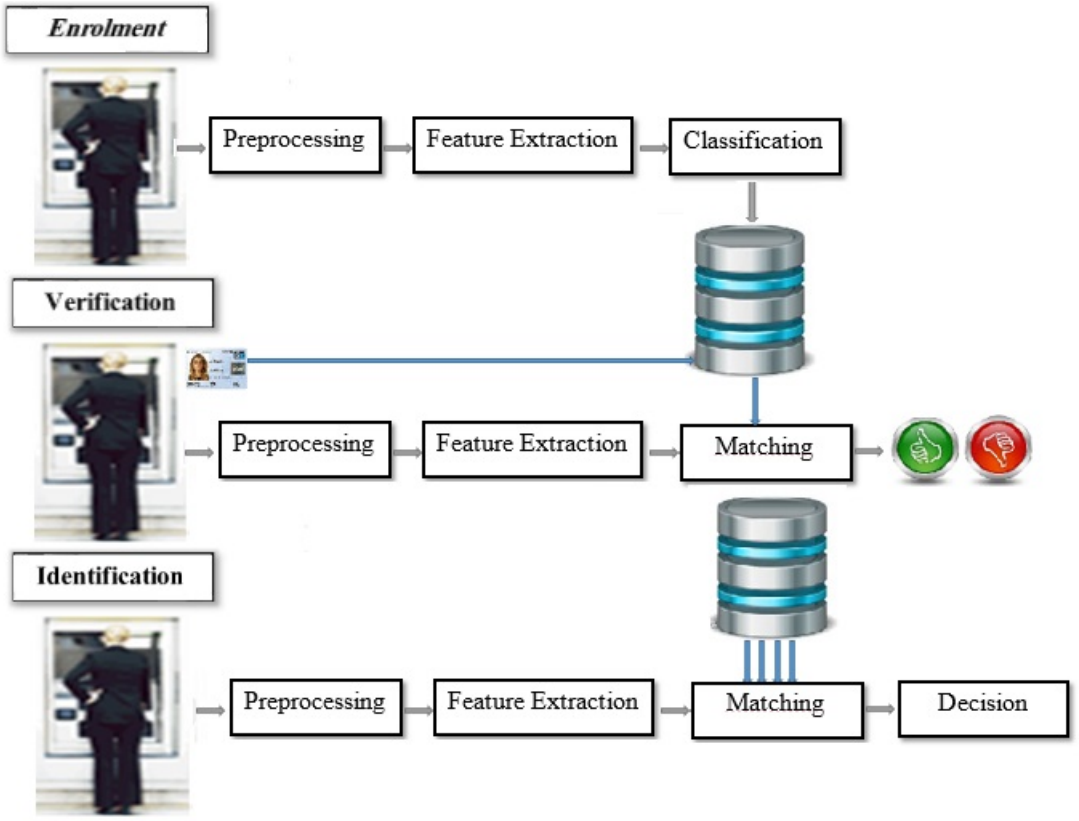


FIGURE 2.6: Biometrics Systems enrolment, verification and identification.

2.3.1 Enrollment Phase

Is an apprenticeship phase, which is used to create the reference database from biometric characteristics of individuals. During this phase, the biometric characteristics are captured by sensor then represented in signature forms, and finally stored in the database.

2.3.2 Recognition Phase

– **Verification Mode:**

Person verification answers the question, "Am I who I claim to be?" and is the process of establishing the validity of a claimed identity by comparing a verification template to an enrollment template [4]. Verification requires that an identity be claimed, after which the individual's enrollment template is located and compared with the verification template. Thus the comparison needed for verification is termed as 1×1 comparison [22]. During verification, usually some knowledge about the identity (such as ID) is given to the system along with the biometric identifier. This additional factor uniquely presents an enrolled identity and extracted biometric features to the system database and hence an associated biometric machine representation [23].

– **Identification Mode:**

Biometric identification establishes a person's identity by answering the question "Whose biometric data is this?" by searching the entire templates in database. The system conducts a $1 \times N$ comparison to establish an individual's identity [3]. In [22], the authors mentioned two types of identification systems: positive identification and negative identification.

They define a positive identification systems as those systems which are designed to find a match in a biometric authentication system to answer the question "Who am I?" [22]. An example of a positive identification system would be an access control system in an office setup to confirm that the employee is on the designated access list of the office.

Negative identification [22] systems ensure that a person is not present in the database. This can be used in benefits programs to prevent users from enrolling under multiple identities.

2.4 Performance Evaluation of Biometric Systems

The performance of biometric systems is an important issue in high security applications. Where, the matching between the stored template and the template constructed generates a confidence score to verify whether they are an impostor or a genuine user.

2.4.1 Error Rates

For each type of decision, there are two possible outcomes, true or false. Therefore, there are a total of four possible outcomes: a genuine is accepted (True Acceptance (TA)) or a False Rejection (FR) occurred, and an impostor is rejected (True Rejection (TR)) or a False Acceptation (FA) occurred [4]. Moreover, there is always overlap region between the score distributions of the genuine user and impostor for a practical biometric system as shown in Fig. 2.7. It causes the difficulty in classifying the claimant into the correct categories. In evaluating the performance for any biometric based recognition system, there are mainly two types of factors: False Acceptance Rate (FAR) and False Rejection Rate (FRR). A verification threshold, T_0 is needed in the overlap region as a reference to do the classification.

According to the distribution shown in Fig. 2.7, T_0 is used to establish the security level of a biometric systems. It can be seen that for those who obtain a similarity matching score less than T_0 will be classified as an impostor. If one is verified with the similarity matching score higher or equals to the threshold, his (her) claimed identity will be accepted as a genuine. A higher T_0 represents a High-security level. Undoubtedly, less impostors will

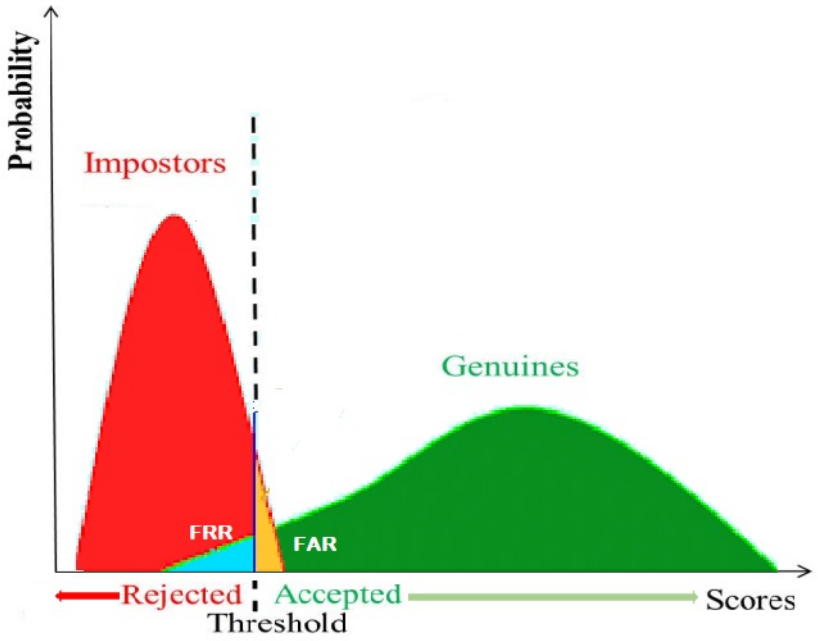


FIGURE 2.7: Distribution of curves impostor and genuine users.

get through verification but a genuine user with score less than T_0 will also be rejected at the same time. Conversely, by adjusting the threshold to a lower level will reduce the number of the genuine users being falsely rejected. However, this will also cause an increase of falsely accepted impostors. In brief, there is a trade-off between these two types of errors.

1. **False Accept Rate (FAR):** is defined as the probability of an impostor being accepted as a genuine individual [4]. That is, in a biometric authentication system, the FAR is computed as the rate of number of people is falsely accepted FA over the total number of the impostor (NI) for a predefined threshold T_0. This is denoted

$$FAR = \frac{FA(T_0)}{NI} \times 100\%.$$

2. **False Rejection Rate (FRR):** is defined as the probability of a genuine individual being rejected as an impostor [4]. That is, in a biometric authentication system, the FRR is computed as the rate of number of people is falsely rejected FR over the total number of total genuine user (NG) for a predefined threshold T_0. The formula for the FRR is denoted

$$FRR = \frac{FR(T_0)}{NG} \times 100\%.$$

3. **Genuine Accept Rate (GAR):** is used to measure the accuracy of a biometric system [4]. It is measured as the rate of number of people is genuinely accepted over the total number of enrolled people for a predefined threshold. In other words, GAR can be

obtained by subtracting the number of falsely rejected people from the total number of genuine people. The GAR is denoted

$$GAR = 1 - FRR(\%).$$

4. **Equal Error Rate (EER):** is a point defines the trade-off between the false rejects and the false acceptances, based on FAR and FRR. Thus, EER is a common way of evaluating the performance of a biometric system where low value of EER is considered to represent a biometric system with highly accurate performance. In general, the EER is the value on $FFR = FAR$.

Other errors that may arise in a biometric system are Failure To Capture (FTC) and Failure To Enrol (FTE). These two errors are crucial for live applications. The FTC error takes place when the data acquisition unit is not capable to capture a satisfactory quality of the biometric trait. Whilst, the error of FTE usually occurs when the user tries to enrol in the recognition system are unsuccessful. All these factors are dependent on the decision threshold T, and by varying decision threshold we can obtain a multiple operating points of the system.

2.4.2 Performance Curves

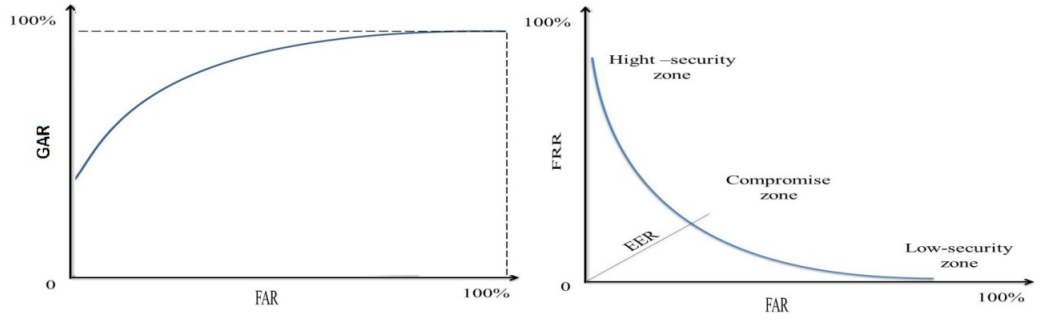


FIGURE 2.8: ROC Curves.

The values of the performance metrics are usually plotted in different graphs or curves to represent the recognition accuracy of the biometric system. The most commonly used plotting curve is the Receiver Operating Characteristics (ROC) curve [24]. It is as shown in Fig. 2.8, the ROC curve plots the GAR against FAR in a semi-logarithmic scale in biometrics research field. Also can be represented the variation of the FRR as a function of FAR; this graph graphically represents the performance of a verification or identification system. The equality error rate (EER) squares at the intersection of the ROC curve with the first bisector.

It is frequently used to give an overview of the performance of a system. It is observed that the curve illustrates in the Fig. 2.8.

Another commonly used curve is Cumulative Match Characteristics (CMC) curve [25] which is mainly used for closed set identification. The Fig. 2.9 illustrates an example for CMC curve. This curve gives the percentage of people recognized according to a variable called rank. This curve is associated by two criteria Rank of Perfect Rate (RPR) and Rank-One Recognition (ROR); ROR represents the most commonly used measure but it is not always sufficient. RPR which corresponds to ROR = 100% [26]. CMC curves show the chance of a good system will start with a high identification rate for low ranks identities.

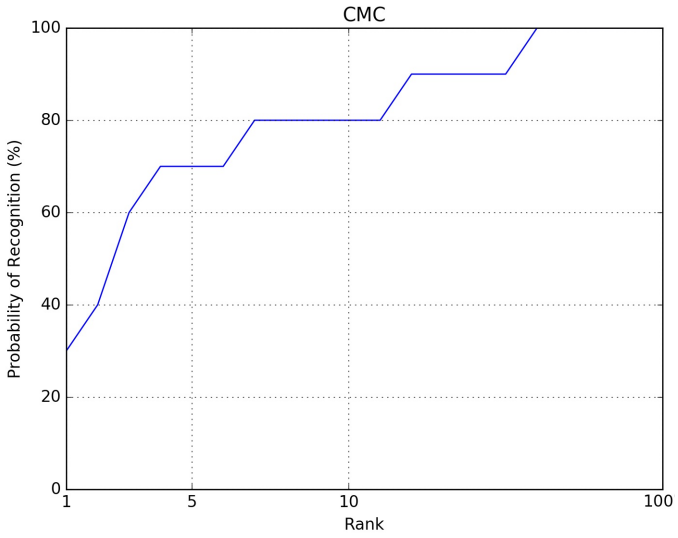


FIGURE 2.9: CMC Curve.

2.5 Single Biometric Systems Limitations

2.5.1 Overview

The majority of biometric systems use a single biometric trait such systems, are called unibiometric systems. Regardless of significant advances in the latest years, there are still several limitations derived from utilising one biometric trait. Nevertheless, single biometrics performance in term of enrolment rate is not sufficient for larger population coverage. For example, 2% of the population as reported in [27] failed to enroll to the fingerprint system due to damaged fingerprints. Some other reasons of enrolment failures can be due to the fact that they are born with less discriminative biometrics or because of physical body changes. Fig. 2.10 show example of unimodal biometric system limitations.

FIGURE 2.10: Example of unimodal biometric system limitations.

2.5.2 Limitations

The limitations of unimodal biometric systems are summarized in five elements as follows:

▷ Noisy Data Acquired by Sensor

The recognition rate of any biometric system is very sensitive to biometric sample quality and noisy data can seriously reduce the overall accuracy of the system [28]. During enrollment, the template quality deteriorates if the biometrics is not properly provided. This is as a result of imperfect conditions or significant variation in the biometric itself. For example, a poorly illuminated face image may cause to reject the sample, appearance of wrinkles due to aging or presence of facial hair, presence of scars in a fingerprint, etc. The variations may be due to improper interaction of the user with the sensor causing the sensor to less effectively capture the biometrics. Also, to the use of different sensors during enrollment and verification or change in the ambient environmental conditions. For instance, the speech/voice biometrics uses the voice print or sound wave to extract the biometrics features. Ambient noise might be integrated into the acoustic signal and cause significant negative impact on the authentication results.

▷ Distinctiveness

The biometric characteristics extracted from different persons may be not distinctive, and have large inter-class similarities used to represent these traits. While it is expected to vary significantly across individuals. For case, identifying identical twins are failed in face recognition system. Inter-user similarity refers to the overlap of the biometric samples from the different individuals. In [3], the distinctiveness of the fingerprint is classified as high whereas the hand geometry is classified as low. Usually, this limitation of distinctiveness increases the errors of a biometric system.

▷ **Non-universality**

The principle of universality is an essential condition in any efficient biometric recognition. However, all biometric modalities are not really universal. The National Institute of Standards and Technology (NIST) has reported that it is not possible to obtain a good quality fingerprint from about 2% of the population (people with disabilities related to the hand, people with oily or dry fingertips, etc.) [27]. Consequently, such people cannot be sign up in a fingerprint verification system. Similarly, persons having those suffering from eye abnormalities or diseases cannot provide good iris images for automatic recognition [29]. Hence, it is possible that some users do not possess that particular biometric characteristics.

▷ **Sensitivity to Attacks**

Many studies [30, 31] demonstrated that it is possible to spoof a number of fingerprint authentication systems using simple techniques with molds made from range of materials such as plastic, clay, silicon or gelatin. There are more spoofing examples given in [32] where biometric systems using facial and iris recognition, are spoofed by using high resolution digital images of face or iris. Moreover, behavioral biometric modalities are more susceptible to this kind of attack.

▷ **Intra-Class Variations**

The biometric sample obtained from a user throughout the identification or verification phase is not identical to the sample which was collected to generate the reference database from the same user during the enrolment phase. This is known as the intra-class variations.

The above mentioned single biometrics restrictions result in errors such as the rejection of genuine user or the acceptance of impostor. These limitations of single biometrics can be overcome by using multiple biometric modalities [33, 34, 35], this integration of more evidence is a feasible way to enhance the biometrics performance.

2.6 Chapter Summary

Biometric is the automated method of recognizing a person which based on a physiological or behavioral characteristic. Biometrics became a significant part of efficient person recognition as biometric traits cannot be stolen, shared or even forgotten. Further, it has pointed out that one of the basic challenges in biometrics research, is finding adequate modalities

and fulfilling the seven main aspects of certainty. In this chapter, we focused on IT security by discussing in more details about the three categories of biometric technologies physiology, behavior and soft. Measurements of these different categories are called biometric modalities, and formed the basis for techniques in various biometric domains. Here, we covered architecture of a biometric system to achieve the objective of recognition. Biometric system consists of both enrolment and recognition, generally includes sensors to be able to read data out and software portion makes use of algorithms to enhance and recognize this data to generate a template unique to the individual. Also, the performance metrics of biometric systems is discussed in this chapter.

The majority of currently used biometric systems usually use a single biometric feature, such systems are called unimodal systems. But biometric system based solely on a single biometric suffers from several limitations such as noisy sensor data, intra-class and spoof attacks, etc. It is generally believed that by integrating various biometrics into one system, the limitations of unimodal systems can be alleviated, given that the several biometric sources usually compensate for the weaknesses of a single biometric. Multimodal biometrics systems are described in the next chapter.

Chapter 3

MULTIMODAL BIOMETRICS FUSION

3.1 Introduction

The unimodal biometric system take a single source of information for authentication. For that, it has various challenges such as lack of secrecy, non-universality of samples, spoofing attacks on stored data, etc. Some of these challenges can be addressed and the security of the resources and information can be further enhanced by employing a multimodal biometric system. As the name depicts, multimodal biometric systems work on accepting information from two or more biometric inputs take from the users for authentication. The multimodal biometrics refers to the process which seeks to manage or coordinate the usage of various biometric modalities in a manner that improves the process of data fusion and perception, synergistically.

3.2 Necessity of Multimodal Biometric Systems

Multimodal biometric systems, provide an improved performance over unimodal systems in their ability to authenticate a user. The advantages of multimodal biometric systems stem from the fact that there are multiple sources of information. The most prominent implications of this are increased and reliable recognition performance, fewer enrolment problems and enhanced security [36, 81]. From these, it can be summarized that multimodal biometrics is a preferred approach than the other scenarios to tackle single biometrics limitations by the following reasons:

3.2.1 Enhanced Security

Multimodal biometrics is more difficult to spoof because biometric traits have to be presented at the same time. The advantage of multimodal systems is that the impostor would have to be able to spoof more than one biometric trait simultaneously, which would be significantly more challenging. Further, some multimodal biometric systems employ challenge-response [4] mechanism to fight against spoof attacks by asking the user to present a random subset of traits at the point of acquisition. Multimodal biometric systems can also serve as a fault tolerant system [4]. If any single trait is unavailable in a multimodal biometric system, the system can still work with other available traits.

3.2.2 Fewer Enrolment Problems

Multimodal biometric systems address the problem of non-universality or the insufficient coverage, it provides alternative biometric options for a claimant who is unable to provide a specific biometrics. By this, the Failure to Enrol (FTE) rate can be significantly reduced significantly [37]. Depending on the system design, many multimodal biometric systems can perform matching even in the absence of one of the biometric samples. For example, in a fingerprint and face based multimodal system, a person cannot enrol his fingerprint information to the system, the system can still perform authentication using the facial characteristics of that person.

3.2.3 Increased and Reliable Recognition Performance

As multimodal biometric systems use more biometric traits, hence each of those traits can offer additional evidence about the authenticity of any identity. Therefore, higher information gain can be achieved because have not any correlation among the sources biometrics. In addition, the usability of the multimodal biometrics system is also better than the single biometrics. For example, in a face and voice based multimodal biometric system, due to ambient noise, if the voice signals cannot be accurately measured, the facial characteristics may be used for authentication. Increased and reliable recognition performance of multimodal biometric systems is ability to effectively handle the interclass similarities, noisy or poor data, etc.

3.3 Multimodal Biometric Fusion Scenarios

Multimodal biometric system can be based on one or more sources of the biometric data obtainable from individual's traits. The scenarios of multimodal system differ from system to

system depending on the application requirements. Generally, the term multimodal biometric system refers specifically to those biometric systems where multiple biometric modalities are used [4]. But in reality, the term multibiometric is more generic and includes multimodal systems and some other configurations using only one biometric modality with different samples instances or algorithms [38]. According to Fig. 3.1, and based on the sources of information, the author in [4] propose the following five possibilities to create multibiometric systems:

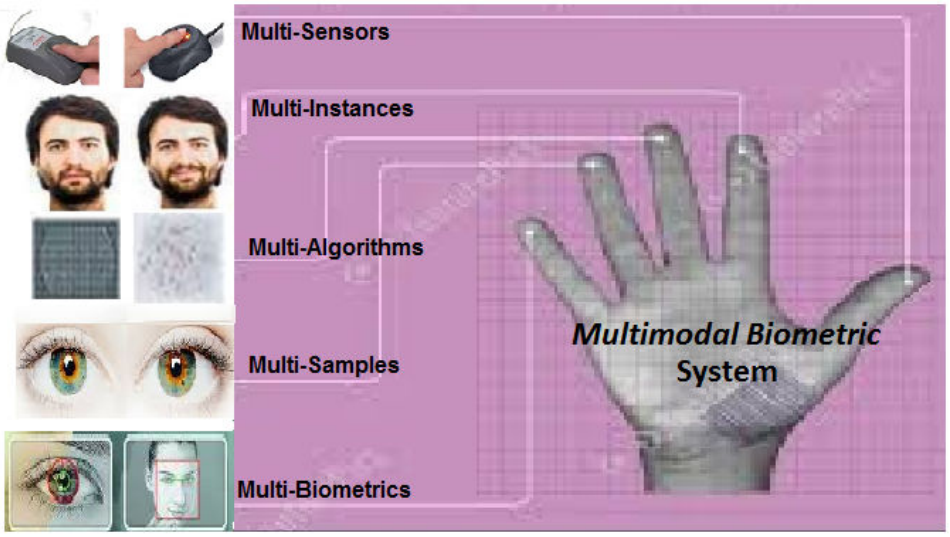


FIGURE 3.1: The information sources of multimodal biometric system.

3.3.1 Multi-Sensors System

In these systems, one biometric trait are used for capturing by different sensors to extract different representations from registered images. For example in [39], the face images of an individual obtained using a thermal infrared camera and a visible light camera. It uses an infrared camera because it's robust against ambient lighting and other variations such as facial hair, wrinkles and expression. It overcomes some limitations of the conventional visual camera used for facial recognition.

3.3.2 Multi-Instances System

In these systems, multiple instances of the one biometric trait are used for the variations that can occur within this modality with the same sensor several times. These systems are cost efficient, as the same sensors or the same feature extraction and matching algorithm can

be used. For example, a facial recognition system can capture different images of the face with changes the positions of these images such as frontal, left and right profiles.

3.3.3 Multi-Algorithms System

These systems process different feature extraction and matching algorithms from a single biometric trait captured through a single sensor. Then, the individual results from each matcher are combined to obtain the final decision. For example, a texture-based algorithm and minutiae based algorithm can operate on the same fingerprint image [3]. Furtherer, these systems is cost effective but suffer with the poor quality of input data.

3.3.4 Multi-Samples System

These systems use only a single sensor but multiple samples of the same biometric trait. An example of multiple samples is using left and right iris images for identity recognition. However, its implementation costs will not be as high as a multibiometric systems.

3.3.5 Multi-Biometric System

These systems use more than one biometric modalities and combine the evidence presented by different body traits for identification systems. The cost of these systems is high since must be using multiple sensors to extract the biometric traits. Usually the identification accuracy of these systems are proportional to the number of traits. Therefore, these scenarios are more frequently used to improve unimodal biometrics performance. Nevertheless, these multimodal biometric systems can be implemented to further enhance the usability and performance. For example, a biometric recognition system based on combining face and ear attributes would be considered a multimodal system.

3.3.6 Hybrid Systems

Hybrid systems concern other types of systems, they are composed of several scenarios from those presented above. Therefore, hybrid systems combine information and have advantages than previous systems. For example, a biometric system may use two iris matching algorithms and three face matching algorithms in one face and iris based multimodal biometric system.

3.4 Multimodal Biometric Architecture

The next step after determining which biometric sources are to be integrated, is to build the system architecture. As Fig. 3.2, any multimodal system can operate in one of three different operational modes: serial, parallel or hierarchical mode [122, 123].

a. Parallel Mode

In this mode of operation, the information from multiple modalities is processed concurrently, independently and all at once. Then, the results are combined to make the final classification decision [123] such as an authentication system based on fingerprint and face recognition. So, if it would be operated in a parallel mode, the user had to present the two traits in the same time for validation.

b. Serial Mode

This mode called cascade mode, each modality is examined before the next modality is investigated. Therefore, multiple biometric traits do not have to be captured at the same time. Furthermore, a decision could be obtained before acquiring the rest of traits. As a result, the overall recognition duration can be decreased. For example, in authentication system based on voice, fingerprint and iris traits. Initially the user uses the voice validation unit, and if this fails fingerprint validation is applied. If the last validation is failed the iris unit is required. The reward of such systems is that many users will enrol to the system using single trait [123].

c. Hierarchical Mode

In this operational mode, individual classifiers are combined in a treelike structure. This mode is preferred when a large number of classifiers are expected. Most of the current multimodal biometric systems operate either in the serial mode or in the parallel mode. The serial mode is computationally efficient, whereas the parallel mode is more accurate [123].

3.5 Multimodal Biometric Fusion Levels

According to [40], “Information fusion can be defined as an information process that associates, correlates and combines data and information from single or multiple sensors or sources to achieve refined estimates of parameters, characteristics, events and behaviors”.

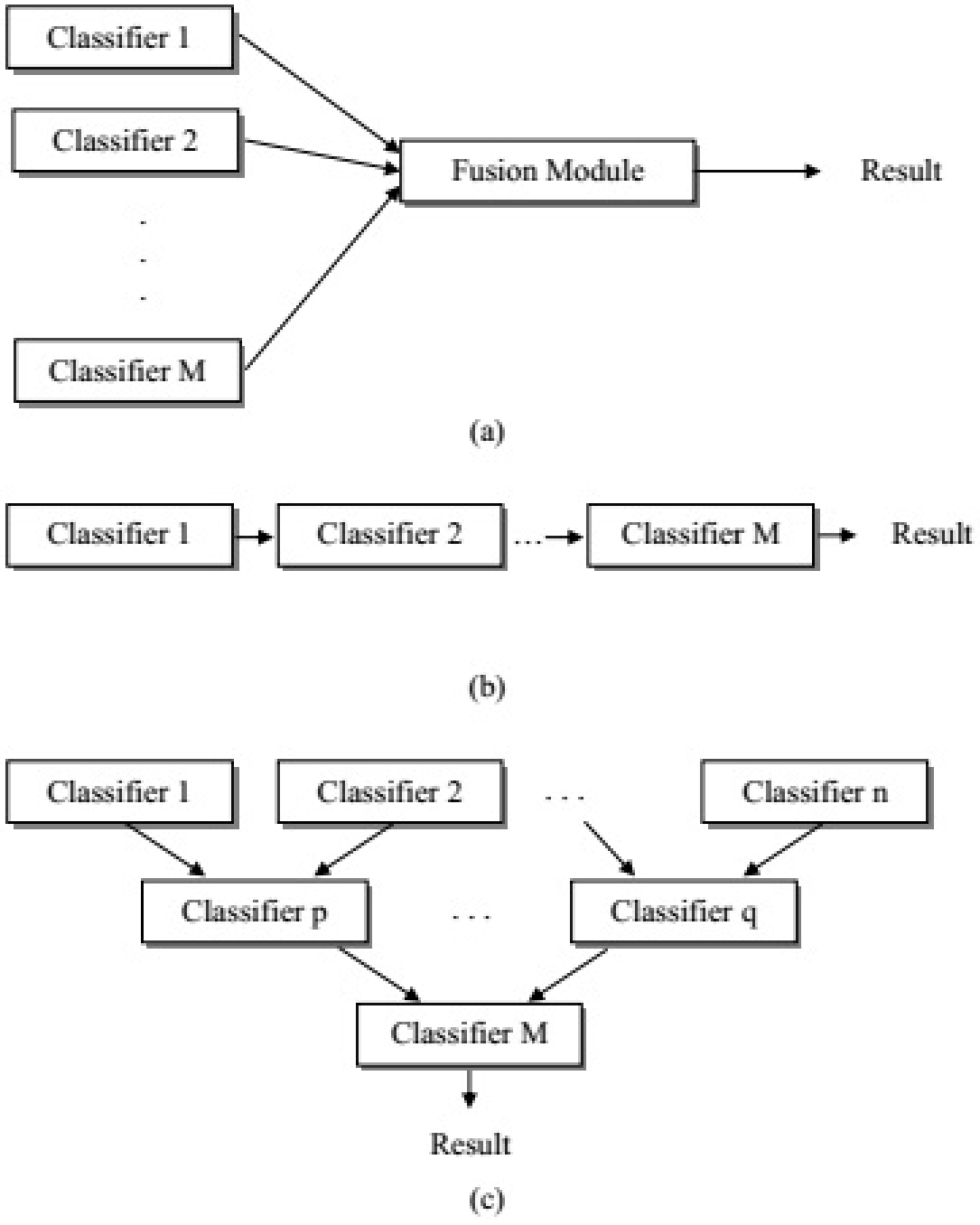


FIGURE 3.2: Architecture for several classifier combinations, from [123], (a) parallel, (b) serial, (c) hierarchical.

A good information fusion method allows minimizing the influence of unreliable sources compared to reliable ones [41]. Since, multimodal biometric systems rely on the evidence presented by multiple sources of biometric information, information fusion is essential for processing of such information. In their research, Sanderson and Paliwal [42] categorized the fusion methods into two broad categories: fusion before matching and fusion after matching. Fusion before matching category contains sensor level fusion and feature level fusion, while the fusion after matching contains match score level fusion and decision level fusion [82]. Fusion classification levels are illustrated in Fig. 3.3.

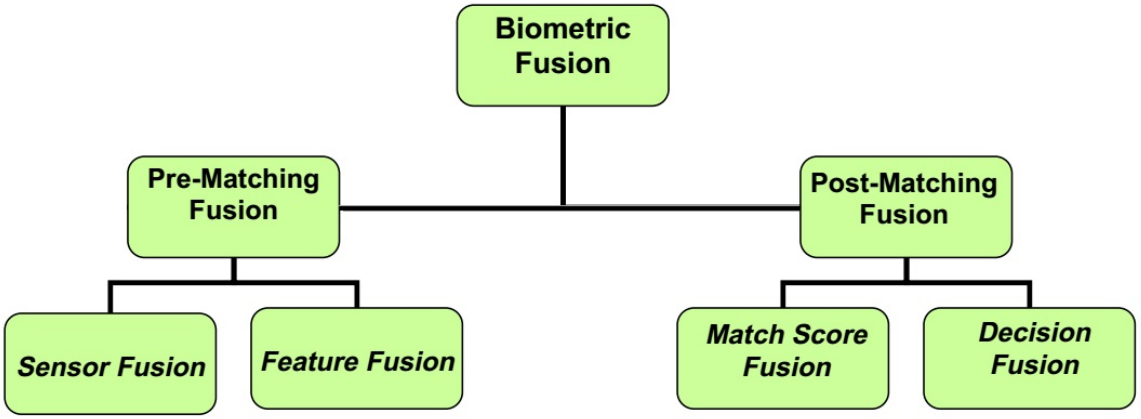


FIGURE 3.3: The block diagram of biometric fusion classification.

3.5.1 Fusion Pre-Matching

– Sensor Level Fusion

In this early stage of fusion, the raw data, derived from the biometric characteristic with two or more sensors, is combined. Sensor level fusion is defined as "the consolidation of evidence presented by multiple sources of raw data before they are subjected to feature extraction" [4]. The fusion at the sensor is relatively little used because it requires homogeneity between the biometric characteristics. Indeed, the sensor level fusion can be done on images of multiple instances of the same biometric modality (images fingerprints obtained from several cameras, or several instances of the same trait biometric obtained from a single sensor) [43]. The resulting information from this initial level represent the richest source of information, whilst the other levels contain a smaller amount of information. Among, the tricks used is Discrete Wavelet Transform (DWT) algorithm, which is a good example of sensor level fusion. For example, in [44], authors combined multiple instances of faces captured by

using a single camera and by mosaicking method to obtain better recognition performance. The Fig. 3.4 show an example of sensor level.

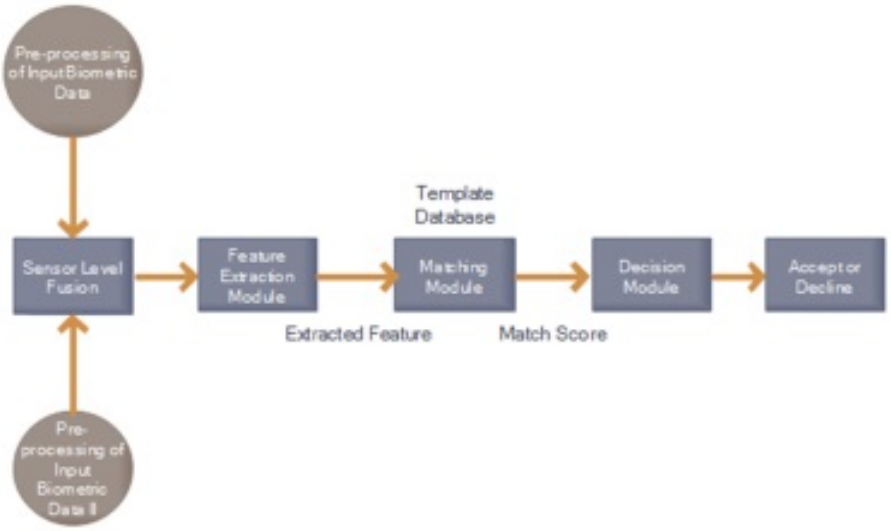


FIGURE 3.4: Process of fusion at the Sensor Level [45].

– Feature Level Fusion

Feature level fusion consists in combining different vectors of characteristics, that are obtained from multiple data sources to create a new feature set to represent the individual. The fusion of features extraction into one feature vector usually involves applying appropriate feature normalization, selection and reduction techniques [4]. Certainly, the feature level is much richer and exploits more useful information, fusion at feature level may be helpful for integrating features of the same modality with multiple sensors. However, such fusion type is not always feasible because finding relationship between the feature sets is difficult [4]. There are some difficulties if the feature sets originate from multiple biometric traits, when feature vectors are heterogeneous. Moreover, the relationship between the feature spaces of the joint biometrics may not be known exactly. In addition, concatenating two feature sets or more may leads to the curse of dimensionality problem. For example, in many approaches the given features might not be compatible due to differences in the nature of modalities, we can concatenate them to form a single feature vector (See Fig. 3.5).

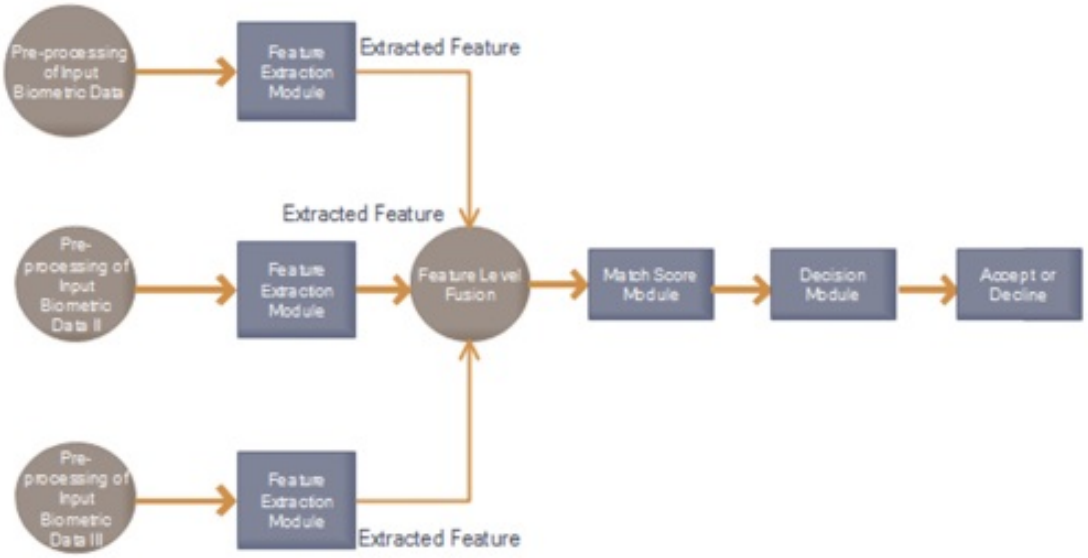


FIGURE 3.5: Process of fusion at the Feature Level [45].

3.5.2 Fusion Post-Matching

– Match Score Level Fusion

This level is also known as measurement level. Fusion at this level is much more effective than fusion at the decision level. Match score level fusion method consolidates matching scores generated from different classifiers and can be applied to most of the multibiometric scenarios. Matching score is a measure of similarity between features derived from a presented sample and a stored template. Each unimodal biometric system measures and calculates its own matching scores and these matching scores are fused to reach a final match/non match decision(See Fig. 3.6).

As different matching scores from different algorithms may not share the same underlying properties or the score range, score normalization is necessary in match score level fusion methods. Min-max, decimal scaling, z-score, median, median absolute deviation, double sigmoid, tanh-estimator are some examples of score normalization techniques. Normalization process is costly in terms of time and choosing inappropriate normalization can lead to very poor recognition accuracy.

For obtaining a single matching score, this fusion method applies arithmetic operations, such as sum, subtraction, maximum, minimum, and median on to different matching scores. As an example, the match scores generated by three different matchers for the face, fingerprint and hand modalities of a user may be combined via the simple sum rule in order to

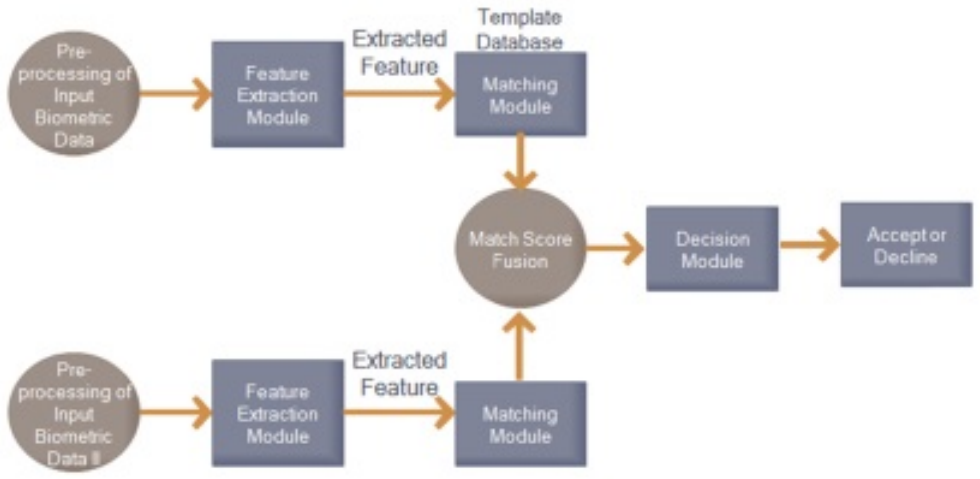


FIGURE 3.6: Process of fusion at the Match Score Level [45].

obtain a new match score which is then used to make the final decision [46]. In [47], the advantages of fusion at matching score level are analyzed in several aspects:

- Matching score fusion does not affect the existing biometric systems, these unimodal biometric systems can be easily combined into a multimodal biometric system given.
- The data from prior evaluations of single modal biometric systems can be reused. This avoids live testing or re-running individual biometric algorithms.
- The matching scores contain the richest information with, it is much easier to access and to combine the scores generated by the different matchers.

– Decision Level Fusion

Since the biometrics verification decision is only accept or reject, very limited information is available for fusion at this level. Therefore, its performance is normally not comparable to the feature and score level fusion. In this level, a separate authentication decision is made for each biometric trait. These decisions are then combined into a final vote, as shown in Fig. 3.7. Thus, fusion at such a level is the least powerful [45]. Methods proposed in the literature for decision level fusion include AND and OR rules [48], majority voting [49], weighted majority voting [50], Bayesian decision fusion [51], the Dempster-Shafer theory of evidence [51] and behavior knowledge space [52].

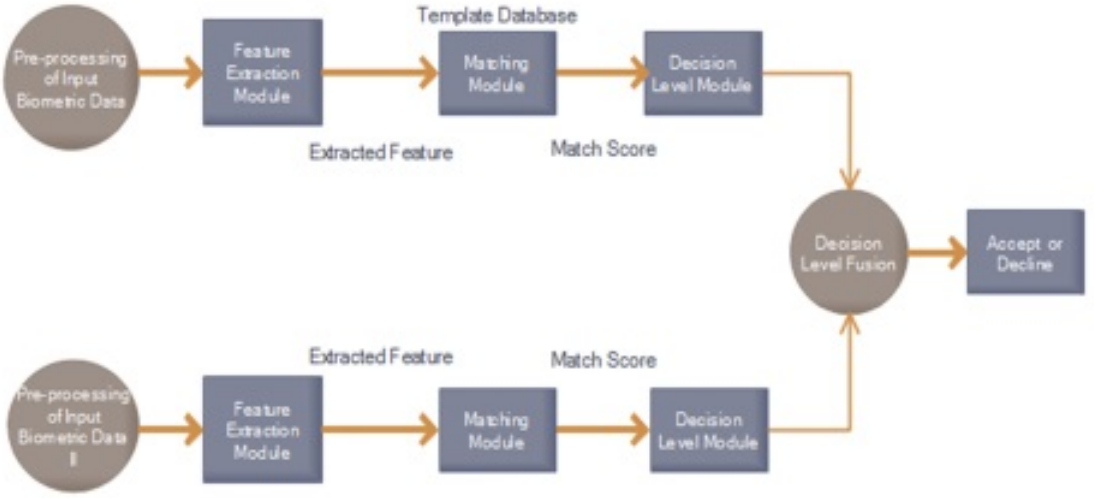


FIGURE 3.7: Process of fusion at the Decision Level [45].

3.6 Multimodal Biometric Fusion Research

Due to some problems associated with the unimodal biometric data, the use of multimodal biometrics is a first choice solution [36]. The main objective of a multimodal biometric system is to improve the recognition performance of the system and to make the system robust over the limitations associated with unimodal biometric systems. Over the years, several approaches have been proposed and developed for multimodal biometric authentication system with different biometric traits and with different fusion mechanisms. The following sub-sections discussed some of the research using different fusion methods for multimodal biometric systems:

a. Research on Sensor Fusion

A multisensory multimodal biometric system fuses information presented by multiple sources of raw data (image, video, sound, text, symbols etc.) at sensor level [4] and is expected to produce more accurate results than the system that integrate information at later stages due to the availability of more information.

In 2003, Liu and Chen [44] propose a face mosaicking technique. This is a method for combining two or more images of the same face. The authors used a 3D ellipsoidal model to approximate human head images. Later, using geometric mapping, authors projected 2D face images onto the ellipsoidal model and used CMU PIE database [124] and a patch based probabilistic model for classification.

Another key contribution in this area is the research reported in [125]. The authors proposed an approach to combine information obtained from face and palmprint image using particle swarm optimization (PSO). The Kernel Direct Discriminant Analysis (KDDA) and the nearest neighbor method are used for feature extraction and classification. Using FRGC face database [126] and polyU palmprint database [127], the authors tested the recognition performance with match score level fusion and with genetic algorithm applied on the same set of databases.

b. Research on Feature Fusion

Feature level fusion consolidates information from multiple biometric feature sets of the same individual. As the most features regarding the identity of a person is available at this level, so feature level fusion is expected to perform better than other level fusion methods [128]. However, there are some inherent drawbacks associated with this fusion method. The feature spaces of different biometric traits may not be compatible and the feature level fusion may lead to the "curse of dimensionality" problem by concatenating several features as one [4]. Due to these drawbacks, the study on feature level fusion is seldom reported.

In 2004, Feng et al. [129] developed a system for face and palmprint using feature level fusion technique. The authors used ORL face database and polyU palmprint database [127] and employed concatenation method for feature fusion. Two feature extraction approaches PCA and ICA to see which results in a better recognition performance were also investigated. As noted by authors, ICA performed better than PCA in both monomodal and multimodal validation framework.

In another attempt to develop a multimodal biometric system, in [130], Rattani et al. proposed a multimodal biometric system combining face and fingerprint information at the feature level. In their research, the authors extracted feature sets from face and fingerprint images and then concatenated (after necessary normalization) them to obtain combined feature set for their system. The authors also employed dimensionality reduction method to handle the problem of "curse of dimensionality" and implemented several feature reduction techniques for the proposed system. In [131], the authors conducted experiments on BANCA face database and a local fingerprint database to evaluate the recognition accuracy with match score level fusion for the same set of database.

c. Research on Match Score Fusion

Matching score fusion consolidates matching scores generated from different classifiers and can be applied to most of the multibiometric scenarios because of its content of adequate information to make genuine and impostor case distinguishable and because of the easy availability of the scores [132]. But to use a different matching scores from different classifiers, normalization of these scores is required which can be a bottleneck of this system for the time requirements. Also choosing inappropriate normalization technique can produce very low recognition accuracy.

In 2003, a bimodal approach was proposed identification system based face, fingerprint and hand geometry with a fusion method at the match score level (integrating the matching scores of different classifiers and making a decision based on decision tree and linear discriminant function) [46]. The MSU fingerprint database and a public domain face database results showed that the system achieved higher recognition accuracy using match score level fusion than when using any single biometric trait [133].

In 2005, Jain et al. proposed a multimodal approach for face, fingerprint and hand geometry, with fusion at the score level [134]. The authors examined simple sum-rule, max-rule and min-rule method of match score fusion with seven normalization techniques. The final outcomes showed that all fusion approaches (except for the median MAD normalization technique) exhibit better recognition performance than monomodal approaches [134].

d. Research on Decision Fusion

Decision level fusion method integrates the final decisions of single biometric matchers to form a consolidated decision. The consolidated decision can be obtained by employing various techniques including “AND/OR”, majority voting, weighted majority voting, decision table, Bayesian decision and Dempster-Shafer theory of evidence. Decision level fusion is too rigid and comparatively less sophisticated than other fusion methods as it operates only on binary information [4].

In 2000, Frischholz and Dieckmann developed a commercial multimodal approach, BioID, for a model-based face classifier, a vector quantization (VQ)-based voice classifier and an optical-flow-based lip movement classifier for verifying persons [135]. Weighted sum rule and majority voting approaches of decision level fusion method were used for fusion. Their experiments on 150 persons for three months demonstrated that the system can reduce the FAR significantly.

In another attempt in 2009, the research of Yu et al. [136] presented a multibiometric approach which combines palmprint, fingerprint and finger geometry collected by a digital camera at decision fusion level. Three decision fusion rules, including "AND" rule, "OR" rule and majority voting, are employed to perform the fusion. Experimental results conducted on a database of 86 hands (10 impressions per hand) showed that the proposed decision fusion methods are effective. Among the three decision fusion methods, majority voting was more accurate than the other two decision methods.

3.7 Challenges for Multimodal Biometric Systems

Although the development in the field of multimodal biometrics has received considerable attention, there are several challenges that are associated with the multimodal biometric system design process including source of information, choice of biometric traits, fusion strategy and level of robustness and so on. Multibiometric systems are more expensive and require more resources for computation and storage. They generally require additional time for user enrollment, causing some inconvenience to the user. This difficulty is related to many issues such as:

3.7.1 Multimodal Datasets

The development of multimodal systems is still limited because of the lack of consistent multibiometrics databases. The main reason behind this lack of multimodal databases is that contained a certain degree of difficulty and challenges in the data acquisition phase. Moreover, most of the publicly available multimodal databases comprised of matching scores obtained by a number of biometric approaches operating on particular modalities [3, 45]. Consequently, this does not allow additional research to be held on other types of fusion levels other than the matching scores level. There are currently a few multimodal person authentication databases that are reported in the literature, some examples are listed in Table. 3.1.

Due to the difficulties in constructing multimodal databases, some researchers have assumed that different biometric traits of the same person are statistically independent in order to simplify the fusion algorithm design [46]. Experiments in multimodal biometrics have been conducted on combining biometric trait of a user from a database with different biometric trait of another user from another different database to generate virtual or so-called chimeric databases [45].

Database	Modalities
BANCA	Face and speech
XM2VTS	Face and speech
VidTIMIT	Face and speech
BIOMET	Face, speech, fingerprint, hand and signature
NIST	Face and fingerprint
MYCT	Fingerprint and signature
UND	Face, ear profile and hand
FRGC	Face modality captured using camera at different angles and range sensors in different controlled/uncontrolled settings
IDIAP	Score of XM2VTS database
MyIDea	Face, speech, fingerprints, signature, handwriting, palmprint and hand Geometry
BioSec	Fingerprint, face, iris and voice

TABLE 3.1: List of available multimodal biometric databases, from [37].

3.7.2 Incompatibility of Information Resources

As one of the most important challenges of the multimodal biometric system is the incompatibility of the information resources, where the integration of biometric information in early stages is thought to be more valuable, since the amount of information available to the fusion process decreases as it moves from one level to the next level of fusion [2, 4, 53, 54]. Nevertheless, fusion at early stages such as sensor and feature levels is not always possible due to the incompatibility of the gathered information. For example, in a multimodal biometric system based on fusing face and voice print, it is not possible to fuse the raw images of face with the voice signal.

3.7.3 Privacy and Acceptance

There is a number of serious privacy concerns raised concerning the implementation of biometrics, due to the fact that biometric technologies have the potential to provide governments and organizations with increased power over individuals [55]. Privacy concerns are related to data collection, unauthorized use of recorded information and improper access to biometric records. As such, a trade-off between security concerns and privacy issues may be necessary by enforcing data protection laws and standards through common legislation [55]. Nevertheless, Biometrics from the positive point of view provides valuable tools to implement liable logs of system transactions [55].

3.7.4 Optimum Design

The improvements in a multimodal biometric-based approach address key design questions. The main question is about the integration of modalities. This strongly depends upon the application and the required level of security concerns [55]. This will also decides the complexity in designing the authentication system. From [55], other design questions ought to be asked such as, What are the best combinations of modalities? How do we choose a best set of samples for a particular biometric? What is the smallest size sample set? Which level fusion is appropriate? Which is the best fusion scenario and processing architecture? What is expected performance? What is the cost involved in developing and deploying a real-time system? A part from the above mentioned factors there are still open questions to be addressed before deploying multimodal biometric system in a real time environment.

3.8 Chapter Summary

The development of biometric system is governed by the limitations of performance effectively and usability. The proposed approaches done to overcome these limitations which are addressing many issues as are: information sources and fusion of information to design multimodal biometric system. Multimodal biometric system is emerging as a current trend which is capable of using, more than one source biometric traits (physiological or behavioral) and allows to fuse the informations extracted from these biometric traits. In this chapter, we have discussed an advantages and necessity of multimodal systems compared to unimodal biometric systems, as well as the different scenarios involved in multimodal biometric system development. Amongst the different scenarios, multiple biometric traits is preferred due to the fact that biometric characteristics are inherently different, so more information gain can be obtained compared to other scenarios. It further provides alternative biometrics entry option. As a result, the system is more robust to adapt with a different operating ambiance (e.g. dim light or noisy conditions) and to cover larger population. At the same time, multimodal biometrics enhances the system usability and makes the spoof attack more difficult.

In this context, biometric sources are being integrated to build the system architecture, where multimodal systems can operate through many architectures modes: serial, parallel or hierarchical mode. Effective fusion of the biometrics information plays a key role for biometrics system improvement. The biometrics information fusion attempts are performed at four potential levels: sensor, feature, matching and decision levels. The sensor and the feature levels are referred to as a fusion before matching while the matching score and the decision levels are referred to as a fusion after matching. Vast majority of works focus on

score level fusion. This is because of the balance between the complexity and richness of information. Furthermore, the challenges of multimodal biometric systems are also discussed in this chapter where can be summarised as more expensive, choice of compatible biometric traits, difficulty of fusion and robust of system design, etc.

Chapter 4

FEATURE EXTRACTION: FROM CLASSICAL TO DEEP LEARNING METHODS

4.1 Introduction

GENERALLY , in a biometric recognition system and after having acquired the biometric traits and preprocessing them, the feature extraction is performed. Feature extraction is important factor for the success of the recognition and classification process, and should be able to extract more information even under difficult conditions, such as: bad lighting, noise and redundant data.

A feature is defined as an "interesting" part of an image, and is used as a starting point in main primitives for algorithms. The features given by the extraction process can be used to gain more statistical information by using parametric deep learning algorithms. This techniques are a new framework for feature extraction which aim to represent deep local features of biometric modalities. In this chapter, the effectiveness of the deep learning methods are extensively examined by using several types of analysis and comparing them with those of existing classical methods. The end result of feature extraction is a set of features, commonly called a feature vector, which constitutes a representation of image.

4.2 Biometric Feature Types

Biometric feature extraction is the process by which key features of the sample are selected or enhanced. The feature is defined as a function of one or more measurements, each of which specifies some quantifiable property of an object, and is computed such that it quantifies some significant characteristics of the object. Typically, the process of feature extraction relies on a set of algorithms; the method varies depending on the type of biometric features used. According to the abstraction level, they can be divided into: Pixel-level, Local and Global features. The most important types of general features which can be considered are:

4.2.1 Texture Features

Texture is one of the important characteristics used in identifying objects or regions of interest in an image. Texture can be defined as superficial phenomenon of natural objects. Texture does not occur over a point but it rather occurs over a region. Texture can be analyzed by quantitative and qualitative analysis. According to quantitative analysis one of the first descriptions given by the Tamura [146] proposed six textural properties and gave descriptions common over all texture patterns. These are six different texture features given by Tamura: Coarseness, Contrast, Directionality, Line-Likeness, Regularity and Roughness.

Texture is an important property of image and is a powerful regional descriptor that helps in the retrieval process [147]. Texture, on its own does not have the capability of finding similar images, but it can be used to classify textured images from non-textured ones and then be combined with another visual attribute like color to make the retrieval more effective [148]. Textural features are:

- Statistical measures: Entropy, Homogeneity and Contrast
- Wavelets
- Fractals

4.2.2 Line Features

Line features are usually correspond to the object contours or boundaries, also are one of the basic elements of inerratic objects. These line features and principal lines identify the length, position, depth and size of the various objects, which are an important clue for vision perception and essential basis for image interpretation. In the image based geometric

measurement, it is important for us to describe the acquired or interpret images by abstracting and positioning line features in a high precision. Line features may not be sufficiently distinctive to be a reliable identifier in themselves, but wrinkle features are highly distinctive and not easily duplicated. There are varieties of features extraction methods, such as: Hough Transform, Boundary Tracing and Curve Fitting, etc [149].

4.2.3 Shape Features

Visual features of objects are called the shape characteristics or visual features. Shape is an important visual feature and one of the primitive feature for image content description. Shape representation is mainly based on the shape features which are based on the boundary plus region content, moment, etc. These representations can be used for matching shapes, object recognition, or for making measurements of shapes. Shape content description cannot be defined exactly because the measuring the similarity between shapes is difficult [150]. Therefore, two steps are essential in shape based image retrieval, they are: feature extraction and similarity measurement between the extracted features. Shape descriptors can be divided into two main categories:

1. Contour based methods which use the whole area of an object for shape description.
2. Region based methods which use local features as boundary segments.

The fourth type, which is also considered important in features classifications, can be added. The color feature is one of the most widely used visual features in image classification. Images characterized by color features, have many advantages such as: robustness, effectiveness and computational simplicity.

4.3 Feature Extraction Methods

The all features can be coarsely classified into low-level features and high-level features. Low-level features can be extracted directly from the original images, whereas high-level feature extraction depends on low level features. A various methods of features extraction can be used for biometric recognition, although these techniques, a recognition systems are need to be robust to the different challenges in representation images area.

4.3.1 Classical Methods

In image processing literatures, the researchers have invested many classical methods for features extraction process. Often, traditional methods depend on the quality of the images

in their representation and are highly affected to noise, which reduces the efficiency of these methods. In this section, we will mention only to two examples of traditional algorithms to extract the feature, that our help to use in subsequent sections.

a. Principal Component Analysis (PCA)

The principal components analysis is unsupervised learning algorithm that provides a means of compressing data. PCA learns a representation that has lower dimensionality than the original input [140]. It also learns a representation whose elements have nonlinear correlation with each other. This is a first step toward the criterion of learning representations whose elements are statistically independent. To achieve full independence, a representation learning algorithm must also remove the nonlinear relationships between variables [140]. Thus, it can use PCA as a simple and effective dimensionality reduction method that preserves as much of the information in the data as possible.

The PCA transform applied to a set of images, can be used to find the subspace that is occupied by all of the images from the analyzed set. The methodology for calculating principal component is given by the following method [140]:

- Let the training set of vectors of original data (each vector with dimension N)

$$X = [x_1, x_2, x_3, ..., x_N]\,. \tag{4.1}$$

- Compute the mean of original data of the set

$$\check{X} = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{4.2}$$

- Subtract the mean from each original data to generate the mean removed data

$$\varphi_i = x_i - \check{X}. \tag{4.3}$$

- Form the matrix using mean removed data of $N \times N$ dimension

$$D = [\varphi_1, \varphi_2, \varphi_3, ..., \varphi_N]\,. \tag{4.4}$$

- Compute the sample covariance matrix C, where

$$C = \frac{1}{N}\sum_{n=1}^{N} \varphi_n \varphi_n^T = DD^T. \tag{4.5}$$

- Compute the **eigen values** of the covariance matrix and the eigen vectors of the **eigen values**.
- Finally, keep only the eigen vectors corresponding to L largest **eigen values**, these **eigen values** are called as principal components.

The process of obtaining a single subspace consists of finding the covariance matrix C of the training set of ROI sub-images, and computing its eigen vectors. Each original ROI sub-image can be projected into this subspace. The eigen vectors spanning the print-space can be represented as images with the same dimensionality as the ROI sub-images used to obtain these eigen vectors [151].

b. Discrete Cosine Transform (DCT)

The feature extraction principle is a transformation of image from spatial representation to frequency representation by provide a value that is defined at each point. The recognition is then carried with this new representation mode. Recently, the transformation by blocks has emerged as a representation of image and has been widely applied in images processing and pattern recognition, particularly. In the transformation by blocks, the image is subdivided into many blocks (sub-images) to reduce data size, memory space and the computation time. The most popular method for transformation by blocks is the 2D-Block based on DCT (2D-BDCT) [151]. The DCT allows us to suppress some non-important data by to removing the high frequencies of the image while keeping the important data represented by the low frequencies.

The application of the DCT by block consists of subdividing the dimensions of image $H \times W$, into size blocks $(N \times N)$, and transforming each block to obtain a block of coefficients at same size. Thus, the image f is split into B_{ij} blocks, with: $i \in \{0, 1, ..., \eta_h - 1\}, j \in \{0, 1, ..., \eta_w - 1\}$.

Where η_h represents the number of horizontal blocks, η_w the number of vertical blocks, and η number of total blocks [151]:

$$\eta \, blocks = \eta_h \times \eta_w = \left[\frac{H}{N} \times \frac{W}{N} \right]. \tag{4.6}$$

The 2D-DCT is defined by [151]

$$F_{ij}(u, v) = \alpha(u)\,\alpha(v) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f_{ij}(n, m)\,\psi(n, m, u, v), \tag{4.7}$$

where

$$\psi(n,m,u,v) = cos\left[\frac{(2n+1)\,u\pi}{2N}\right] cos\left[\frac{(2m+1)\,v\pi}{2N}\right], \tag{4.8}$$

with: $u,v = 0,1,...,N-1$, $F_{ij}(u,v)$ are the DCT coefficients of blocks B_{ij}, f_{ij} is the intensity value of pixel of block B_{ij}, and $\alpha(*)$ is defined by:

$$\alpha(\gamma) = \begin{cases} \sqrt{\frac{1}{N}} & si \;\; \gamma = 0, \\ \sqrt{\frac{2}{N}} & si \;\; \gamma \neq 0. \end{cases} \tag{4.9}$$

The DCT makes possible to decompose the blocks into a matrix of coefficients that represents the influence of each frequency constituting the block. The first value is the equivalent of average value of block. For the line, the different values correspond to the horizontal frequencies contained in the block. For the column, the different values correspond to the vertical frequencies contained in the block [151]. For block B_{ij}, the DCT matrix of coefficients covers all frequency components of the block. These coefficients are organized to arrive at the vector of characteristics that represents each block. The large amplitudes coefficients are mainly located in the upper left corner of the DCT matrix. Consequently, the traversing of the DCT matrix by zigzag from the upper left corner converts the block into one vector. Fig. 4.1 shows the properties of the DCT coefficients of block 8×8, and the model called zigzag [151].

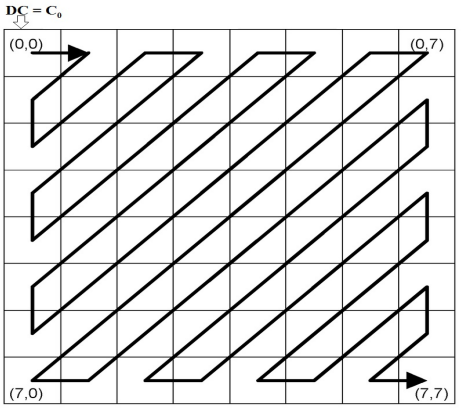


FIGURE 4.1: The Model zigzag of block 8×8.

The DC coefficient is defined

$$F_{ij}(0,0) = C_0 = \frac{1}{N}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} f_{ij}(n,m). \tag{4.10}$$

The 63 remaining coefficients (AC coefficients) denote the intensity changes in the block. Finally, the image is represented by the vector characteristics generated by the concatenation of all the vectors characteristic of the blocks. However, the vector of each block B_{ij} is

represented by [151]

$$o_{ij} = \left[F_{ij}\left(0,0\right), F_{ij}\left(0,1\right), F_{ij}\left(1,0\right) ..., F_{ij}\left(7,7\right)\right]^T = \left[C_0, C_1, C_2, ..., C_7\right]^T . \tag{4.11}$$

The total DCT vector of the characteristics of the image is

$$V_T = \left[o_{11}, o_{12}, o_{13}, ..., o_\eta\right] . \tag{4.12}$$

As a method of feature extraction, DCT transforms the high dimension images (see Fig. 4.2) into low dimension space, with the important features of the image, such as the main lines and the crests, are kept.

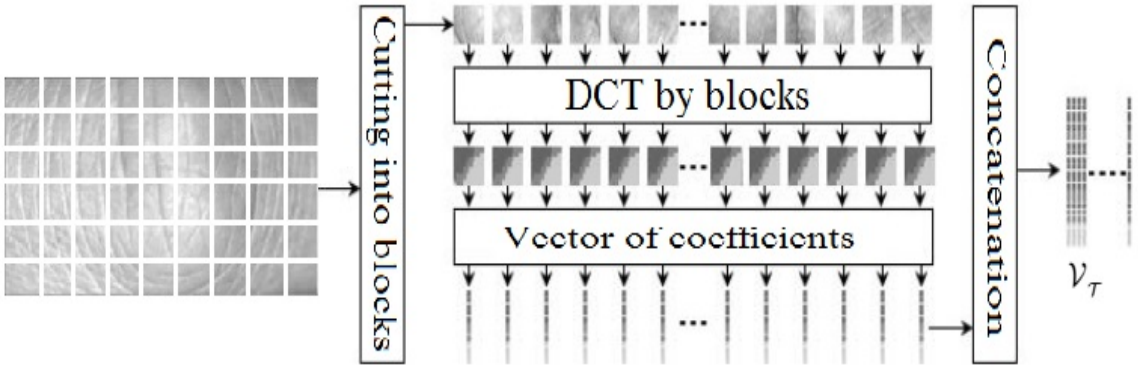


FIGURE 4.2: Generation of features vectors based on the DCT [151].

4.3.2 Deep Learning Methods

Compared with the classical feature extraction methods, one of the foremost advantages offered by the deep learning based feature extraction methods is flexibility and so its discriminative [61, 62] because it is possible to use a higher level features into multiple levels of representation to extract some discriminant information of the image trait. Hence, a lot of feature extraction methods rely on deep learning techniques have been proposed in literature [63, 85].

a. PCANet Feature Extraction

PCANet technique is one of simple deep learning techniques, this technique provides a reliable solution to extract the majority of information in image which can be used in a greater range of pattern recognition systems to discriminate the images. In contrast to other

deep learning techniques, like Convolutional Neural Networks (CNN) [64, 79, 84] and Deep Belief Networks (DBN) [65, 80], PCANet method offers some advantages such as their very suitability for texture analysis as well as their simplicity (complexity relatively simple vis-a-vis almost all deep learning techniques).

The PCANet model [66] cascades many filter bank convolutions (extracted from input image by PCA technique [67]) with an intermediate mean normalization step, followed by two other steps which are the binary hashing and the histogram composition step. PCANet algorithm can execute multiple stages of PCA filters to extract higher level feature vectors. We give hereafter an example of two stages PCANet based feature extraction method (see Fig. 4.3).

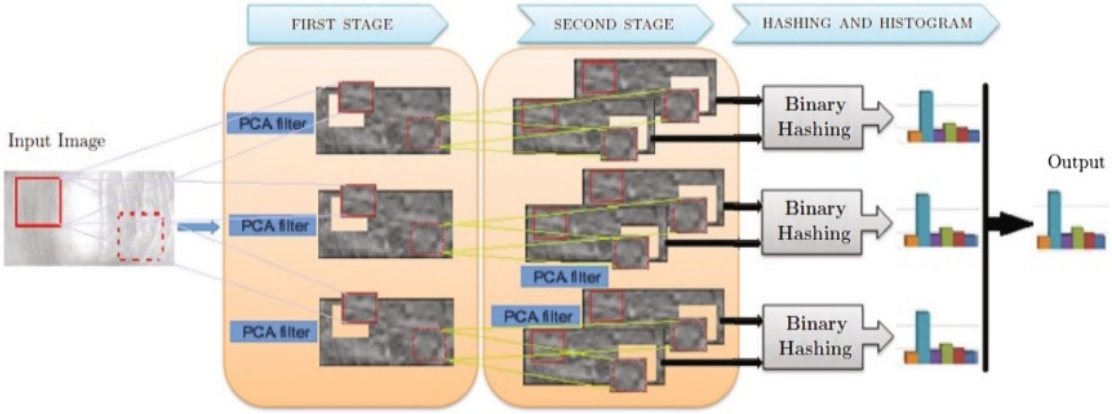


FIGURE 4.3: Example of the PCANet extracts features from an FKP image.

In the first stage, the filter banks are estimated by performing principal components (PCA) technique over a set of vectors where each vector represents the $k_1 \times k_1$ points around each pixel. Before performing this technique, the mean of each vector must be subtracted from it (normalization process).

After applying PCA over the normalized vectors, a $k_1 \cdot k_1 \times L_1$ retained, where L_1 is the primary eigen vectors. Next, each principal component $(1 \cdot \cdot L_1)$ is a filter and may be converted to $k_1 \times k_1$ kernel which is convolved with the input image. So, using L_1 vectors (L_1 convolution filter), we can convert the input image I into L_1 output filtered images

$$I_{F_i}^{s_1}(x, y) = (I * \mathcal{F}_i^{s_1})(x, y), \tag{4.13}$$

where $I_{F_i}^{s_1}$ $(i \in [1 \cdot \cdot L_1])$ denote the i^{th} filtered image using the $\mathcal{F}_i^{s_1}$ filter for the first stage and $*$ denotes the discrete convolution.

In the second stage, the same algorithm used in the first stage is iterated over each of the

L_1 output filtered images (I_F^i)

$$I_{F_{ij}}^{s_2}(x,y) = (I_{F_i}^{s_1} * \mathcal{F}_j^{s_2})(x,y), \tag{4.14}$$

where $I_{F_{ij}}^{s_2}$ ($i \in [1 \cdot \cdot L_1]$ and $j \in [1 \cdot \cdot L_2]$) denote the j^{th} filtered image using the $\mathcal{F}_j^{s_2}$ filter (with size of $k_2 \times k_2$) for the second stage.

If the number of filters in second stage equal to L_2, the output of the last convolution layer produce $L_1 \cdot L_2$ output filtered images.

Subsequently, the finally outputs ($I_{F_{ij}}^{s_2}$) are converted into binary format by using a Heaviside step function [68] which their value is one for positive entries and zero otherwise, this step called binary hashing step.

$$I_{ij}^B(x,y) = \begin{cases} 1 & \text{if} \quad I_{F_{ij}}^{s_2}(x,y) \geq 0, \\ 0 & \text{Otherwise}, \end{cases} \tag{4.15}$$

where I_{ij}^B is a binary image. After that, around each pixel, the vector of L_2 binary bits is viewed as a decimal number

$$I_i^D(x,y) = \sum_{j=1}^{L_2} 2^{j-1} I_{ij}^B(x,y), \tag{4.16}$$

where I_i^D is an image whose every pixel is an integer in the range $[0,\ 2^{L_2-1}]$.

Finally, the histograms of the obtained images are computed and then concatenated to form a feature vector which represents the input image, this step called histogram composition. Thus, the feature vector of the input image I is then defined as:

$$\upsilon_I = [\upsilon_1, \upsilon_2, \cdots \upsilon_{L_1}]. \tag{4.17}$$

where υ_i denotes the histogram of the I_i^D image.

Lastly, it is important to note that in the PCANet technique it must choose the optimal values of the PCANet parameters which are the number of stages (N), the filters sizes in each stage ($k_1, k_2, \cdot \cdot k_N$) and the number of filters in each stage ($L_1, L_2, \cdot \cdot L_N$).

b. DCTNet Feature Extraction

PCANet was worken in various image classification tasks. In this thesis, we also used a data-independence network, dubbed DCTNet for biometric recognition system in which we adopt DCT as filter banks in place of PCA by the fact that 2D DCT basis is indeed a good approximation for high ranked eigen vectors of PCA [152]. DCTNet adopts a similar structure to PCANet except there is an extra layer at the histogram output for histogram

normalization as shown in Fig. 4.4. The detail of each component is described below.

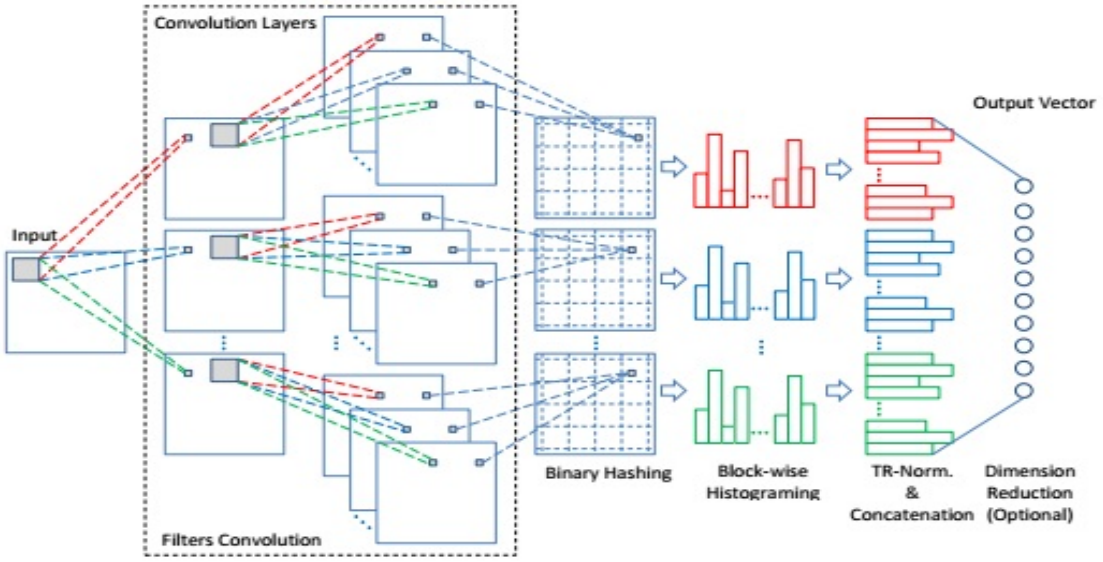


FIGURE 4.4: The block diagram of the DCTNet from [152].

∗ Convolution Layer

Assume that filter size of all stages have the same size $k \times k$. Given an input image I_d of size $m \times n$ with D channels (multiple channel image or input from previous layer), boundary of each channel d is zero padded with pad size $(k-1)/2$ before convolution to keep the size of output O_d^p same as I_d. With a set of 2D-DCT basis selected as denoted [152]

$$W_p^l \in \mathbb{R}^{k \times k}, \quad p = 1, 2, 3, ..., P_l, \tag{4.18}$$

where P_l is the number of filters at layer l, convolving each with I_d yields

$$O_p^d = \left\{ I_d * W_p^l \right\}_{p=1}^{P_l}. \tag{4.19}$$

The number of output of each layer is $d \cdot P_l$. Cascading this layer can form a deeper network. Since, there is no nonlinear operation between the previous convolution layer and the next layer, DCT basis of each layer can be combined to form a flat single layer network. The number of bases formed is:

$$N_{bases} = \prod_{i=1}^{L} P_i \ , \tag{4.20}$$

where L represents the number of convolution layers for the sake of convenience without storing large number of combined filters and to ease the binarization process.

∗ Binarization and Block-wise Histograming

The last convolution layer of DCTNet forms D sets of real valued outputs. Each set has a total of P_L outputs where the outputs are the response of DCT filters. Binarization is performed on each set separately by first binarizing the responses with threshold at zero (value one for positive response, zero otherwise) which denoted by $BIN(*)$. Followed by binarization, each binary string is encoded as a single integer number z:

$$z = \sum_{p}^{P_L} 2^{p-1} BIN\left(O_d^p\right), \tag{4.21}$$

and forming an "image" for each set of d^{th} output where each pixel has an integer range of $\left[0, 2^{P_L-1}\right]$. Then, each of these D binarized "image" is partitioned into B non-overlapping blocks. Histogram of each block denotes by: $H_b^d, b = 1, 2, ..., B; d = 1, 2, ..., D$ with bin $\left[0, 2^{P_L-1}\right]$ is obtained as the input of histogram normalization layer.

It is also worth to mention that block-wise histogram not only encodes spatial information [152], it also provides local translation invariance in the extracted features within each blocks. The combination of binarization and block-wise histograming are expected to be able to extract discriminative features.

∗ Histogram Tied Rank Normalization

The first stage of Tied Rank normalization uses tied rank principle that computes rank of a given vector X which produces a vector $\bar{X}$ that has a range from 1 to the length of X where each element $\bar{x_i}$ corresponds to the ascending order rank of x_i. In case of ties, their average rank is assigned to all ties which may produce a non-integer values. Given H as the extracted block-wise histogram of a given face data, where

$$H = \left\{ H_b^d \right\}_{p=1,d=1}^{B,D}, \tag{4.22}$$

each H_b^d is ranked with tied ranking without considering the bin with zero occurrence denoted by $\bar{H_b^d}$. This is because bin with zero occurrences is not a sample in histogram, it should be ignored in the ranking process. In order to make $\bar{H_b^d}$ to be more evenly distributed, first apply square root on $\bar{H_b^d}$ forming

$$v_b^d = \sqrt{\bar{H_b^d}}. \tag{4.23}$$

Follow by $L2$ norm normalization which follows the idea of intra-normalization uses to obtain $\hat{v}_b^d$ [152]. The final Tied Rank normalized histogram feature vector is constructed by

Algorithm :
Input : Extracted block-wise histogram of an image: H.
Output : Tied Rank normalized histogram feature vector: V.
1. For each H_b^d compute tied rank without bin with zero occurrence yields $\bar{H}_b^d$
2. $v_b^d = \sqrt{\bar{H}_b^d}$.
3. Normalize v_b^d with $L2$ norm to obtain $\hat{v}_b^d$
4. Repeat step 1 to step 3 for $b = 1, 2, ..., B; d = 1, 2, ..., D$
5. Concatenate all $\hat{v}_b^d$ to obtain the final output V

TABLE 4.1: Algorithm of Histogram Tied Rank Normalization

concatenating all $\hat{v}_b^d$:

$$V = \left[\hat{v}_1^1, \hat{v}_2^1, \hat{v}_3^1, ..., \hat{v}_B^1, \hat{v}_1^2, \hat{v}_2^2, ..., \hat{v}_B^D\right] . \tag{4.24}$$

Finally, the dimension of the resulting Tied Rank normalized block-wise histogram vector is optionally compressed to obtain the final feature vector where the projection matrix is learned. The pseudo code of histogram Tied Rank normalization is shown in Table. 4.1 [152].

4.4 Classical Methods Limitations

Feature extraction is the key factor that affects the performance of image classification. Traditional feature extraction algorithms often use single feature and directly use the low-level features, thus, these methods bring limitations. The existing classical methods mainly have the following defects:

- The difficulty of feature extraction can increase due to noisy features. Noise in a dataset is defined as "the error in the variance of a measured variable" which can result from errors in measurements [153]. These algorithms tend to be affected by noisy data, for that noise should be reduced as much as possible in order to avoid unnecessary complexity in the models and improve the efficiency of the algorithms.

- In classical learning as the dimensionality of the data rises, the amount of data required to provide a reliable analysis, and processing a large number of data involves high computational cost [154]. When the dimensionality of a dataset grows significantly, there is an increasing difficulty in the classification due to overfitting. An overfitted model can mistake the important variance of data which can lead to classification errors.

For example, PCA does not select features, it just try to find linear combination of features that are almost as informative as the initial features, but in smaller number. Obviously, it is a dimensionality reduction method, because it's used to find the principal components of features. By contrast, it eliminates candidate features that are irrelevant, thereby using classical PCA is often useless. For DCT, the performance is affected due to certain limitations such as: The correlation between pixels of the neighboring blocks is ignored. It considers only spatial correlation of the pixels inside the single 2-D block, additional time and effort must be put to correct the scaling factor.

4.5 Deep Learning Varieties

4.5.1 Deep Learning Strengths

Deep learning is an architecture that can be easily adapted with a new problems of pattern recognition. The advantages of deep learning are:

▷ No Need for Feature Engineering

Feature engineering is a process of the creation of feature extractors from raw data to reduce the complexity of the data and make patterns more visible to learning algorithms to work. In machine learning, most of the features need to be identified by a features extraction methods, the performance depends on how accurately the features are selected and extracted. But, one of deep learning's main advantages is its capacity to execute feature engineering on its own. The deep learning algorithms scan the data to search for features that correlate and combine them to enable faster learning, thereby reduces the need for feature engineering, one of the most time-consuming parts of machine learning practice. Besides, the neural networks that a deep learning algorithm is made, can uncover new, more complex features that traditional methods can miss. Deep learning algorithms do to learn high-level features from data. This is a very distinctive part in deep learning and a major step ahead in the high-level representation.

▷ The Efficient Performance

The most important strengths of deep learning is its efficient performance with the big data. This is because deep learning algorithms need a large amount of data to understand it perfectly. Fig. 4.5 summarizes this fact. On the other hand, deep learning algorithms can be trained using different big data formats correctly that the quality of deep learning work never diminishes despite of the raw data increases.

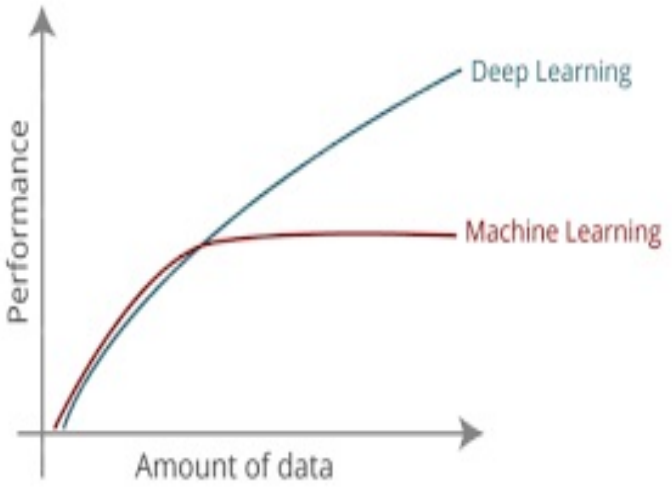


FIGURE 4.5: The efficient performance of deep learning.

▷ **Advanced Analysis and Problem Solving**

While machine learning works only with labeled data, deep learning supports unsupervised learning techniques that allow the system to become smarter in its own. The capacity to determine the most important features allows deep learning to efficiently provide data with concise and reliable analysis results.

When solving a problem using traditional learning algorithm, it is generally recommended to break the problem down into different parts, solve them individually and combine them to get the result. In contrast, deep learning allows processes cycle reduction and efficient utilization of resources that can lead to the automation of tasks to solve the problem end-to-end.

4.5.2 Deep Learning Weakness

One fact that cannot be ignored is that the techniques and applications of deep learning play a more significant role in the success of biometrics recognition by focusing on more powerful model architectures and better learning techniques. Based on comprehensive experimentation, we identify the weaknesses of the deep learning models, and present key points:

▷ **Cost and Hardware**

Deep learning algorithms heavily depend on high-end machines, contrary to traditional machine learning algorithms, which can work on low-end machines. This is because the requirements of deep learning algorithm is a simple but broad task which most of the computation in the training process is large amount of matrix multiplication operations. Graphics Processing Units (GPUs) use a parallel architecture, is very good at handling many sets of simple instructions. Over time, GPUs started to show massive improvements over CPUs for training models, these units are getting more expensive. Also, advances in data storage are some of the major challenges of deep learning which means more and more memory hardware.

▷ **Execution Time**

Fortunately, Deep learning algorithms are becoming increasingly accurate, and are becoming more widely used. Deep learning is trainable neural networks that needs a lot of training data. Usually, a deep learning algorithm takes a long time to train. This is because there are so many parameters in a deep learning algorithm that training them, takes longer than usual. For example, state of the art deep learning algorithm can takes about two weeks to train completely. At test time, the time increases on increasing the size of testing data. Although, this is not applicable on all deep learning algorithms, as some of them can have a small testing times. In conclusion, the processing is difficult and expensive in terms of time and expertise.

▷ **Interpretability**

In general, anything that requires interpreting or applying the scientific method, this is out of reach for deep learning models. This factor is the main reason of think long before its use. Classical algorithms give us crisp rules as to why it chose? or what it chose? So, it is particularly easy to interpret the reasoning behind it. In contrast, the performance of deep learning is quite excellent, but we don't know: what there neurons were supposed to model? What these layers of neurons were doing? And why it has given these parameters? As a sequence, we fail to interpret the results.

4.6 Simple Deep Learning Methods

Image classification plays a very important significance in our life which is a process including image processing, image segmentation, key feature extraction and matching identification. During the rise of deep learning, feature extraction and classifier have been integrated

to a learning framework which overcomes the difficulties of feature selection in traditional methods. The idea of deep learning is to discover multiple levels of representation, with the hope that a high-level features represent more abstract semantics of the data. But these algorithms require a large amount of data with take more time and expensive machines as noted in last sub-section (Section. 4.5.2).

According of these previous facts, we justify the use of simple deep learning algorithms in Section. 4.3.2 as a compromise solution has the ability to learn deep feature representations of biometric modalities. Including the compatibly with our machine, because a training more convolutional neural networks, when the iteration times increase, is difficult in order to produce a good model by limited processing units and that requires more than more time. Lastly, Annex D provides a detailed course in the general basis and important functions of deep learning that will be applied in the image processing fields.

4.7 Chapter Summary

In this chapter, we started with a definition of what a biometric feature is, and presented feature extraction tasks. The biometric feature types can be divided into three categories: Texture features, Line and Shape Features; we described these categories and gave some examples of each category. Then the classical feature extraction methods are described from applied statistics algorithms with increased emphasis on the PCA and DCT. Comparing with the classical methods, we discussed on the deep learning architectures and how the hierarchical brain structure can help to improve features engineering? For biometric systems, the deep learning methods select and analyze deep features to easily finding patterns. Furthermore, we described a some factors of limitation in traditional methods to extract the features and proving lack competency required. These limitations have motivated the development of deep learning algorithms that overcome these obstacles.

Deep learning is an aspect of machine learning that is concerned with emulating the data mining approach to gain certain types of pattern. Moreover, it applies a transformation on its input data and uses to create a statistical model. Iterations continue until the model has reached an acceptable level of accuracy. This chapter also presented a basic of deep learning as strengths and weakness of deep learning algorithms, to estimate the use of deep learning functions as make them popular approaches in the recent researches.

Chapter 5

PROPOSED FKP RECOGNITION SYSTEM

5.1 Introduction

THIS chapter describes the design and implementation of FKP identification system using deep learning in multimodal biometric based system. The deep learning technique is subfield of machine learning research [56, 57] which applied learning algorithms to scan multiple levels of representation to modeling complex relation within the data. Thus, it identified high level features and concepts based on the lowest of them. Recently, deep learning methods are known a great interesting in the field of classification and pattern recognition. In general, a deep neural network algorithms consist on multiple trainable stages stacked on top of each other [58], each stage generally comprises of convolutional filter bank layer, a nonlinear processing is adopted for feature extraction as well as texture patterns analysis and classified [83]. However, learning a network useful for classification which critically depends on expertise of parameter tuning and some setting tricks [59]. Also, this chapter discusses on the investment of the power of these algorithms in feature extraction and fusion for multimodal biometric system based on FKP recognition.

5.2 Finger Knuckle Anatomy

Each finger has three joints. There are three bones in each finger called the proximal phalanx, the middle phalanx and the distal phalanx. The first joint is where the finger joins the hand, called the proximal phalanx [91]. The second joint is the proximal inter phalangeal joint, or PIP joint. The last joint of the finger is called the distal inter phalangeal joint, or

DIP as shown in Fig. 5.1 (Source: Finger Joint Anatomy JointReplacement.com).

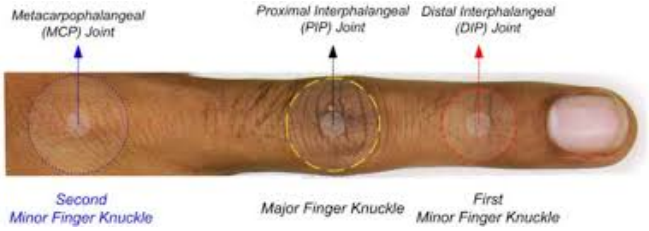
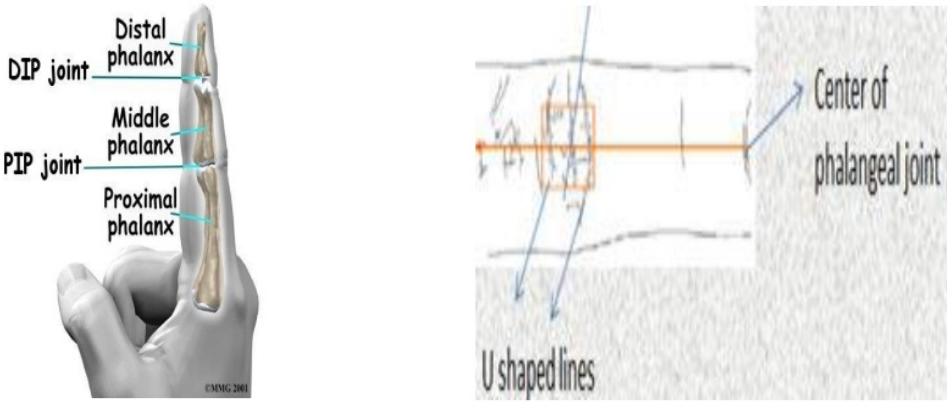


FIGURE 5.1: Illustration of finger knuckle.

Finger knuckle is the back surface of finger, it is also known as dorsum of the hand. The inherent skin patterns of the outer surface around the phalange joint of one's finger, has high capability to discriminate completely different people. Such image pattern of finger knuckle is unique for authentication. In FKP, features are center of phalangeal joint, U shaped line around the middle phalanx, number of lines, length and spacing between lines. Knuckle crease patterns and stray marks are means of photographic identification. Such features are unique and can be used for identification. Extraction of options of knuckle for identification is completely depends upon the user. Some of the researchers of science extracted the options for authentication as shown in Fig. 5.2.



(*a*) (*b*)

FIGURE 5.2: Illustrations of finger knuckle anatomy, *(a)* Distal inter phalangeal joint and *(b)* Centre of phalangeal joint.

FKP uses feature detection and matching techniques in its hard core design. It works similar at almost authentication system. Fig. 5.3 shows the finger knuckle features. Biometrics authentication must provide the security level, unattended system, spoofing and reliability [91].

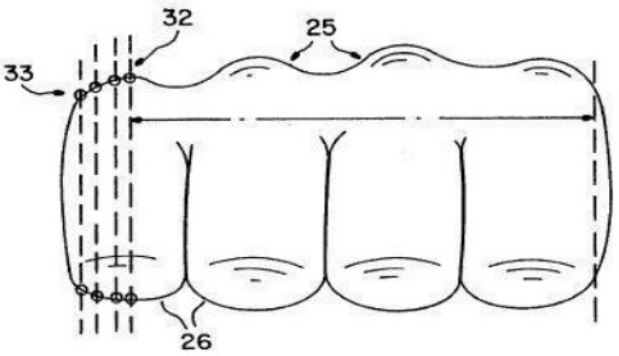


FIGURE 5.3: Illustration of finger knuckle features.

5.3 Finger-Knuckle-Print (FKP) Recognition System

5.3.1 FKP Researches

Many authentication systems have been used widely in business-related and law enforcement applications. The use of different biometric traits such as fingerprint, face, iris, ear, palmprint, hand geometry and voice have been well studied [92]. Based on existing research reports the skin pattern on the finger-knuckle is significantly rich in texture due to skin folds and creases, and hence, can be considered as a biometric identifier [93]. The advantages of using FKP include rich in texture features [94], easily accessible, contact-less image acquisition, invariant to emotions and other behavioral aspects such as tiredness, stable features [95] and acceptability in the society [96]. There are a lot of characteristics and advantages of using FKP as biometric identifier [100], but limited work has been reported in the literature [97].

Systems reported in literature have used global features, local features and their combinations [86] to represent FKP images. Many efforts have been made to build a FKP system based on global features. FKP features are extracted using Principle Component Analysis (PCA), Independent Component Analysis (ICA) and Linear Discriminant Analysis (LDA) in [98]. These subspace analysis methods may be effective in face recognition but they are not found to be effective to represent FKP [99]. FKP is transformed using Fourier transform and the Band-Limited Phase Only Correlation (BLPOC) is employed to match the FKP images [88].

Global feature gives the general appearance (holistic characteristics) of FKP which is suitable for coarse level representation, while local feature provides more detailed information from specific local region and is appropriate for finer representation [86]. There exist

systems where local features of FKP are extracted by using Gabor filter based competitive code, (CompCode) [88] combined orientation and magnitude information (ImCompCode and MagCode) [87]. Further, in [90], orientation of random knuckle lines and crease points (KnuckleCodes) of FKP which are determined using random transform are used as features. In Damon L. Woodard and Patrick J. Flynn [93], FKP is represented by curvature based shape index.

Morales et. al. have proposed a FKP based authentication system (OE-SIFT) using Scale Invariant Feature Transform (SIFT) from orientation enhanced FKP. SIFT features of FKP is matched using similarity threshold [101]. In [99], the authors proposed features which are extracted using Fast Feature transform (MonogenicCode). Further, [86] has proposed a verification system which is designed by fusing the global information extracted by BLPOC [102] and the local information obtained by Compcode [88]. However, there does not exist any system which is robust to scale and rotate.

5.3.2 FKP System Description

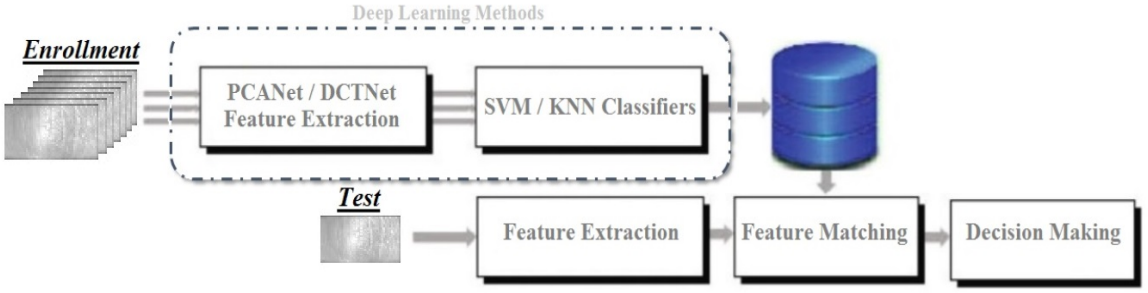


FIGURE 5.4: Block-diagram of the proposed unimodal biometric system.

In FKP the texture pattern is produced by the finger knuckle bending. It is highly unique and makes the surface a distinctive biometric identifier. Like any other biometric identifiers, FKPs are believed to have the critical properties of universality, uniqueness, and permanence for personal recognition. Comparing to other existing biometric traits, finger knuckle is one of the most popular traits in current research. In the proposed multimodal system, we combine the FKP modalities to provide better performance and to improve the security level. In Fig. 5.4, we show the block-diagram of the proposed biometric system based on the FKP images. For enrollment phase, the feature vectors are extracted from the all Region of Interest (*ROI*) sub-images [60] by deep learning network. After that, the PCANet or DCTNet feature vectors will be as a training data used to create models based on the SVM and KNN classifiers, respectively. Training data is a matrix where each column

corresponds to a feature vector. In the second phase, devoted to the identification, the same method is applied, on finger test image, for extract the feature vector, then it uses as an input to matching process in order to find the person which own the same test finger.

5.4 Classification Stage

Regarding the classification algorithm, the most important thing is its capacity to discriminate, based on the available information. SVM has been chosen since it proven advantages in handling large scale classification tasks with good generalization performance. Additionally, it has demonstrated superior results in various classification and pattern recognition problems [139]. Furthermore, KNN has also been chosen as second classifier to provide a large generalization about classification techniques especially with a big number of training samples and a high number of input variables. In this section, we will give brief background knowledge on SVM and KNN.

5.4.1 Overview of SVM

SVM has been recently proposed as a popular tool for solving many classification tasks based on the statistical learning theory which invented by Vapnik [107]. For this purpose, we use SVM to validate our approach, it is a good classification accuracy reported for many pattern recognition problems. To achieve better generalization performance of the SVM, original input space is mapped into a high-dimensional dot product space called the feature space, and in the feature space the optimal hyperplane is determined. The optimal hyperplane is found by exploiting the optimization theory, and respecting insights provided by the statistical learning theory.

▷ Linearly Separable

Given training vectors $x_i, i = 1, ..., N$ of length n and a vector y defined as follows

$$y_i = \begin{cases} 1 \ if \ \ x_i \ \ in \ \ Class1, \\ -1 \ if \ \ x_i \ \ in \ \ Class2. \end{cases} \tag{5.1}$$

The central idea of SVM is to define a separating hyperplane, so, the classification margin between the two classes is as large as possible, where measured along a line perpendicular to the hyperplane.

The SVM training model finds the separating hyperplane which gives the maximum margin or distance between the parallel hyperplanes that are as far apart as possible while still separating the data. These hyperplanes should satisfy the following constraints since the wider margin can acquire the better generalization ability.

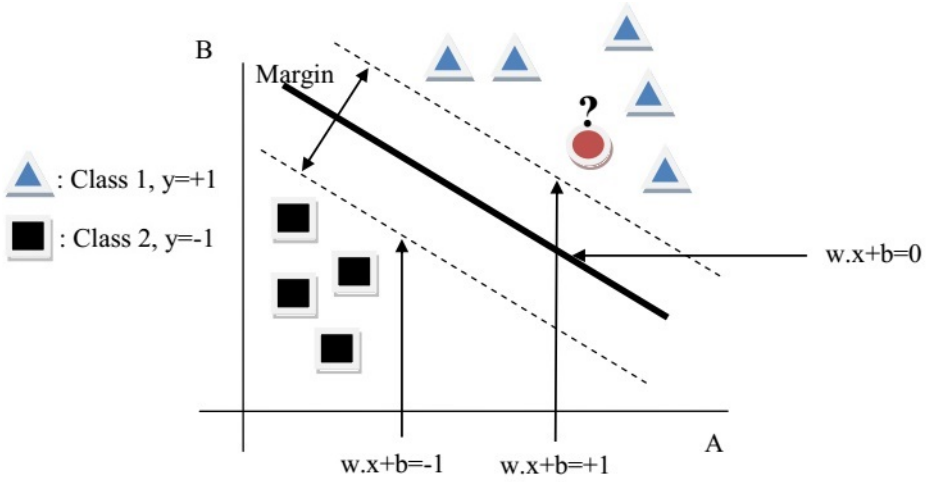


FIGURE 5.5: Linearly separable data.

In Fig. 5.5, two class's stances could be separated by bold solid line. The test sample (the circle) can be classified by based on the hyperplane. In this figure, the hyperplane that is calculated from these training examples is given by the bold line, separated from the closest training vectors by the distance d. The classification of an unknown sample is done by determining where the new instance falls at the hyperplane side. In this example, the prediction for the unknown sample would be triangle.

So, we can define a canonical hyperplanes as follows [107]:

$$\begin{cases} H_1 = w^T x + b = \ +1, \\ H_2 = w^T x + b = \ -1. \end{cases} \tag{5.2}$$

In addition, all training samples x_i satisfy

$$\begin{cases} w^T x + b \geqslant +1 \ \ for \ \ y_i = +1, \\ w^T x + b \leqslant -1 \ \ for \ \ y_i = -1. \end{cases} \tag{5.3}$$

For linearly separable data, any hyperplane $g(x) = 0$ can be written as

$$g\left(x_i\right) = w^T x + b = 0, \tag{5.4}$$

where w is an n-dimensional vector , b is the offset of the hyperplane from the origin and x represents n-dimensional vector representing any point on the hyperpalne. The vector w

and the scalar b determine the position of the separating hyperplane. The distance between each of the canonical hyperplanes and the separating hyperplane is $\frac{1}{|w|}$. Now, maximizing the separating margin is equivalent to maximizing the distance between the hyperplane H_1 and H_2. Hence, we can get the maximal width between them $m = (x^+ - x^-) \cdot \frac{w}{\|w\|} = \frac{2}{\|w\|}$. Now, we can formulate the learning problem of SVM to maximize the margin of task as follows

$$minimize \;\; g(w) = \frac{1}{2}\|w\|^2 . \tag{5.5}$$

So, $w^T x_i + b \geqslant +1, \forall i$.

This enable us to use the Lagrange formalism to obtain the primal form of the objective function L_p , which is

$$minimize \;\; L_p(w, b, \alpha_i) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i \left(y_i \left(w^T x_i + b\right) - 1\right), \tag{5.6}$$

where $\alpha_i : 1, ..., n; \alpha_i \geqslant 0$ are the Lagrange multipliers.

Solving the minimization problem is equivalent to finding the values w, b, and $\alpha_i \geqslant 0$ that minimize L_p. So, we initial differentiate L_p with respect to w and b. Then, by equating the derivatives to zero, we get

$$\frac{\partial L_p}{\partial b} = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i y_i = 0, \tag{5.7}$$

$$\frac{\partial L_p}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{n} \alpha_i y_i x_i, \tag{5.8}$$

when differentiating with respect to b and w respectively. Taking these two equalities and substituting into L_p yields the dual form of the Lagrangian. We want to maximize

$$L_p = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j, \tag{5.9}$$

under the following

$$\sum_{i=1}^{n} \alpha_i y_i = 0, \;\; \alpha_i \geqslant 0. \tag{5.10}$$

This optimization's formula is expressed by using the inner product of the training samples x_i and the numbers of training samples n.

▷ Linearly Non-Separable

In the previous section, the SVM theory was introduced as an optimization problem, under the assumption, the data are linearly separable. However, in many practical problems, data is subject to noise or outliers, so it is impossible to draw linear boundaries between

classes. Hence, in order to extend the support vector theory to solve imperfect separation, positive slack variables is introduced $\xi_i : 1, ..., n; \xi_i \geqslant 0$ into the original constraints (see there in [107]) along with an additional penalty value C for the points that cross the boundaries to consider the misclassification errors. C is a regularization parameter used to decide a trade-off between the training error and the margin. If C is chosen too small, it may cause the problem of under-fitting of the training data. If C is too large, the algorithm may increase the possibility of over-fitting. So, we have

$$minimze \;\; g(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i \,, \tag{5.11}$$

then

$$y_i \left(w^T x_i + b \right) \geqslant 1 - \xi_i \,, \;\; \xi_i \geqslant 0. \tag{5.12}$$

The primal and dual forms of the Lagrangian are built as:

$$L_p = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j, \tag{5.13}$$

and

$$\sum_{i=1}^{n} \alpha_i y_i = 0, \;\; 0 \leqslant \alpha_i \leqslant C. \tag{5.14}$$

▷ **Kernel-Trick**

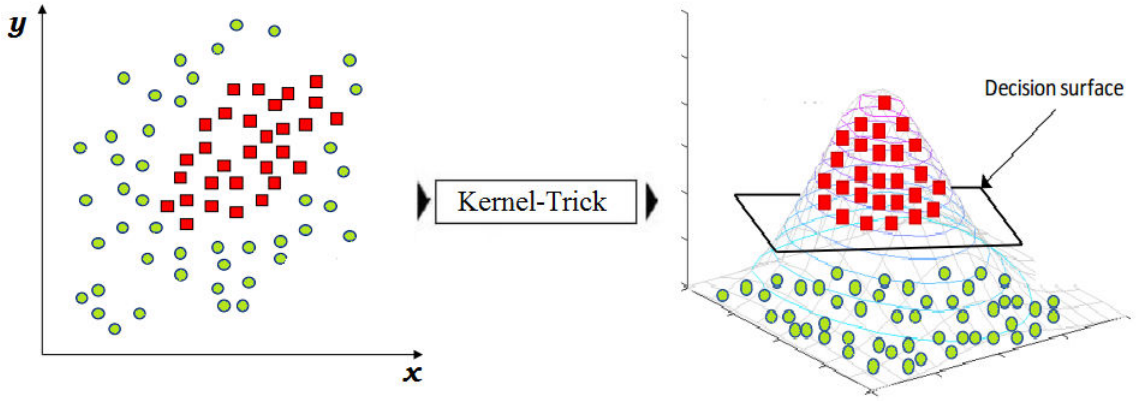


FIGURE 5.6: Kernel trick for non-linearly separable data.

The initial optimal hyperplane algorithm proposed by Vapnik [107], was a linear classifier. In [155], Boser et.al suggested a way to create a nonlinear classifiers by applying the kernel trick to extend the linear learning machine to handle nonlinear cases (see Fig. 5.6). Kernel function is essentially a weighted function designed for nonparametric function estimations.

We aimed to maximize the margin of separation between patterns to have a better classification result. The calculations can be simplified by converting the problem with Kuhn-Tucker conditions into equivalent Lagrange dual problem.

With this mapping, the discriminant function is

$$g\left(x_i\right) = w^T \Phi\left(x\right) + b. \tag{5.15}$$

And the dual form of the Lagrangian becomes

$$L_p = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \Phi\left(x_i\right)^T \Phi\left(x_j\right), \tag{5.16}$$

and

$$\sum_{i=1}^{n} \alpha_i y_i = 0, \;\; C \geqslant \alpha_i \geqslant 0. \tag{5.17}$$

Overall, any positive semi-definite functions $K\left(x_r, x_i\right)$ that satisfy mercer's condition can be a kernel functions. The function $K\left(x_r, x_i\right)$ that returns a dot product of two mapped patterns is called a kernel function. Different kernels can be selected to construct the SVM. The most commonly used kernel functions are the polynomial, linear and Gaussian radial basis kernel function (RBF).

- **Linear kernel function:**

$$K\left(x_r, x_i\right) = \left(x_i^T x_j\right). \tag{5.18}$$

- **Gaussian RBF:**

$$K\left(x_r, x_i\right) = exp\left(-\gamma \left\| x_i^T, x_j \right\|^2\right), \gamma > 0. \tag{5.19}$$

- **Polynomial kernel function:**

$$K\left(x_r, x_i\right) = \left(r + \gamma x_i^T x_j\right)^d, \gamma > 0, \tag{5.20}$$

where γ , r and d are kernel parameters.

5.4.2 Overview of K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) algorithm is one of the simplest classification algorithm and the most used. KNN is a non-parametric and lazy learning algorithm which stores all available cases and classifies new cases based on a similarity measure. KNN has been used in statistical estimation and pattern recognition [144].

▷ KNN Algorithm

KNN algorithm is a type of supervised learning, where the distance measures are very essential to find the similarity and dissimilarity between the classification data [145]. Similarity is if two objects have a same measurement and dissimilarity is two objects are different. The main aim of distance metric calculation is to find appropriate or similar distance as in [145]:

- Euclidian Distance:

$$D = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}. \tag{5.21}$$

- Chebychev Distance:

$$D = max \sum_{i=1}^{k} |x_i - y_i| . \tag{5.22}$$

- Minkowski Distance:

$$D = \left(\sum_{i=1}^{k} |x_i - y_i|^q \right)^{\frac{1}{q}} . \tag{5.23}$$

In the instance of categorical variables, the Hamming distance must be used. It also brings up the issue of standardization of the numerical variables between 0 and 1 when there is a mixture of numerical and categorical variables in the dataset:

$$D_{Hamming} = \sum_{i=1}^{k} |x_i - y_i| , \begin{cases} x = y \Rightarrow D = 0, \\ x \neq y \Rightarrow D = 1. \end{cases} \tag{5.24}$$

The rule simply retains the entire training set during learning and assigns to each query a class represented by the majority label of its k-nearest neighbors in the training set. The Nearest Neighbor rule is the simplest form of KNN when K = 1. In this method, each sample should be classified similarly to its surrounding samples. Therefore, if the classification of a sample is unknown, then it could be predicted by considering the classification of its nearest neighbor samples [144]. Fig. 5.7 shows that the KNN decision rule for K= 1 and K= 4 for a set of samples divided into 2 classes. In the right of Fig. 5.7, an unknown sample is classified by using only one known sample; in the left of Fig. 5.7, more than one known sample is used. In the last case, the parameter K is set to 4, so the closest four samples are considered for classifying the unknown one. Three of them belong to the same class, whereas only one belongs to the other class. In both cases, the unknown sample is classified as belonging to the class on the left [144].

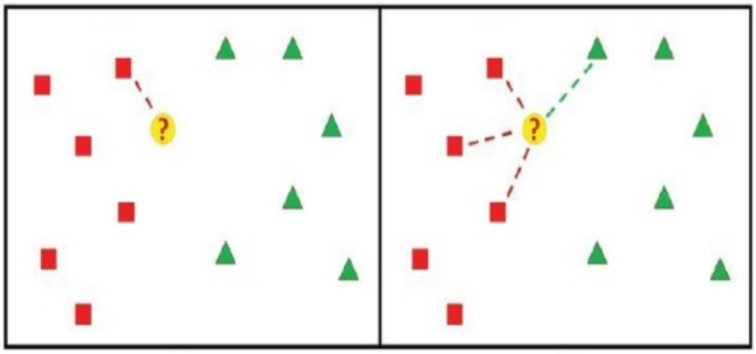


FIGURE 5.7: A example of the KNN decision rule (from [144]).

The KNN algorithm
for all the unknown samples UnSample(i)
for all the known samples Sample(j)
compute the distance between UnSamples(i) and Sample(j)
end for
find the k smallest distances locate the corresponding samples
Sample(j1),...,Sample(jk) assign UnSample(i) to the class which
appears more frequently
end for

▷ **Advantages and Disadvantages**

The KNN has several main advantages: simplicity, effectiveness, intuitiveness and competitive classification performance in many domains. It is robust to noisy training data [144].

Despite of the above advantages, The KNN has a few limitations, which can have poor run-time performance when the training set is large. It is very sensitive from the irrelevant or redundant features because all features contribute to the similarity and to the classification. Two other disadvantages of the method are the selection of the distance type and the high computation cost [144].

5.4.3 Feature Classification

In our work, which based on the PCANet or DCTNet deep learning algorithms (FKP biometric system), the SVM and KNN classifiers are used for the PCANet and DCTNet features, respectively. For PCANet case, a SVM classifier is a supervised machine learning algorithm [69], which can be used for both classification and regression challenges. However,

it is receiving increasing attention and have shown superior performance in pattern recognition [70]. For that, we used a multiclass linear SVM operate on the extracted feature vector for each image. The specific algorithm used was one-against-others support vector classification. Thus, in this algorithm, binary SVMs are combined in one-against-others, each of the binary SVMs separates a single class from all remaining classes (SVM-pairwise scheme) [71]. These classifiers are arranged in trees where each node represents a SVM.

In DCTNet case, a KNN classifier is a type of supervised learning algorithm, which can be used to measure the similarity and dissimilarity between the FKP traits for our classification process. It is important to note that in our scheme, two additional classifiers (Radial Basis Function (RBF) [72] and Random Forest Transform (RFT) [73] are tested and compared with the multiclass SVM classifier. All these classifiers SVM, RBF and RFT will be provided with more details in the Annex C.

5.5 Matching Stage and Normalization

5.5.1 Scores Normalization

Score normalization is a critical step in the design of a combination scheme of the score level fusion. To address the problem of incomparable classifier output scores in different combination classification systems, normalization methods are used to change the location and scale parameters of the matching score distributions at the outputs of the individual matchers. In such a way, various matching scores of different matchers are converted into a common domain and can be combined later [134].

It is highly desirable that the normalization of the location and scale parameters of the matching score distribution must be robust and efficient. Huber [137] defines robustness as insensitivity to the optimal estimate when the distribution of the data is known. Huber also argues even though many techniques can be used for score normalization, the challenging work is to identify a technique that can be both robust and efficient.

∗ Objectives of Normalization

Generally, we give three important issues to be considered before combining scores:

- The scores in output the individual subsystems may be not homogeneous. For example, one system can output a distance measure (dissimilarity) while another output a measure of proximity (similarity).

- The outputs of individual systems are not necessarily included in the same interval.
- The scores in output subsystems may follow different statistical distributions.

So, normalization of scores is essential for transforming subsystems scores in a same interval before fused them.

∗ Min-max Normalization

Normalization is a process that changes the different distance values in a common domain. Indeed, the normalization improves the performance of the biometric recognition system. Many methods of normalization scores can be used such as: "Decimal scaling", "Z-score normalization", "MAD normalization" and "Tanh-estimators" [138]. One of efficient technique used in normalization process is the *Min-Max* technique which is the process of transforming the different scores to value between 0 and 1 [75]. The lowest (min) value is set to 0 and the highest (max) value is set to 1. This provides an easy way to compare values that are measured using different scales of measure. *Min-Max* normalization is defined as

$$\widetilde{\mathfrak{D}} = \frac{\widetilde{\mathfrak{D}} - min(\widetilde{\mathfrak{D}})}{max(\widetilde{\mathfrak{D}}) - min(\widetilde{\mathfrak{D}})}, \tag{5.25}$$

$$\text{and} \quad \widetilde{\mathfrak{D}} = [\widetilde{d}_1, \widetilde{d}_2, \widetilde{d}_3, \cdots, \widetilde{d}_i], \tag{5.26}$$

where $\widetilde{D}$ denotes the normalized similarity/dissimilarity scores. However, these scores are compared, and the highest score is selected. Therefore, the best score is $\widetilde{d}_o$ and its equal to

$$\widetilde{d}_o = \max_i(\widetilde{\mathfrak{D}}) \qquad \text{with} \qquad i \in [1 \cdots N], \tag{5.27}$$

where N denotes the number of references vectors in system database. Finally, this score is used for making the decision in which a threshold T_o is used to regulates the system decision. The system infers that pairs of biometric samples generating scores lower than or equal to T_o, are mate pairs. Consequently, pairs of biometric samples generating scores higher than T_o are non mate pairs.

5.5.2 Feature Matching

During the identification process, the characteristic of the test FKP image, corresponding to each person, is analyzed by the deep learning algorithms. Then the similarity/dissimilarity between two given feature vectors is computed. For that, the obtained feature vector should enter a scores computing process.

In the proposed PCANet system, the matching process is based on the classification score. The SVM classifying a such observation (or feature vector) X by the signed distance from X to the decision boundary ranging from $-\infty$ to $+\infty$ [74]. For a score joined to class C, a positive score indicates that X is predicted to be in C, a negative score indicates otherwise. The predicting score for the observation X into the positive class $f(x)$, is the trained SVM classification function

$$f(x) = \sum_{j=1}^{n} \alpha_j y_j G(x_j, x) + b, \tag{5.28}$$

where $(\alpha_1, \alpha_2, ..., \alpha_n, b)$ are the estimated SVM parameters, $G(x_j, x)$ is the dot product in the predictor space between X and the support vectors and the sum includes the training set observations. The score for predicting X into the negative class is $f(x)$.
If $G(x_j, x) = x_j' x$ (the linear kernel), then the score function reduces to

$$f(x) = (x/s)'\beta + b, \tag{5.29}$$

where s is the kernel scale and β is the vector of fitted linear coefficients.

For the proposed DCTNet system, the matching process is based on Euclidean Distance. When measuring the distance in the features matching step, we use the Euclidean distance between test feature and model features in database. According to the Euclidean distance formula, the distance d is given by

$$d = \sqrt{\sum_{i=1}^{N} |\tilde{x_i} - x_i|}, \tag{5.30}$$

where $\tilde{X}$ is the vector of test, X is the vector in database and N is the size of vectors. With the distance vector $D = [d_1, d_2, ..., d_M]$, where M is the number of distances.

5.6 Biometric Modalities Combinations

Currently, fusion at the matching score level appears to be the most useful fusion level because its good performance and simplicity [76]. In this stage, normalized matching scores are fused to generate an output scores from different unimodal sub-systems, which are then used for making the final decision. During our series of tests, four different fusion schemes are experimented which are *Sum*-score, *Min*-score, *Max*-score and *Weighted*-score rules [77]. Thus, if the scalar $\widetilde{d_i}$ represents the score of the i^{th} sub-system and F_s represents the fusion score. Therefore, F_s is given by

1. **Sum-score (SUM):** Combining the scores by the sum consists to calculate F_s such that

$$F_s = \sum_{i=1}^{k} \widetilde{d}_i. \tag{5.31}$$

2. **Min-score (MIN):** In this technique, we assign to the score final (fused) the best (minimum) score calculated by the different systems. Minimum is then defined by

$$F_s = min(\widetilde{d}_1, \widetilde{d}_2, \cdots, \widetilde{d}_k). \tag{5.32}$$

3. **Max-score (Max):** In this technique, we assign to the score final (fused) the best (maximum) score calculated by the different systems. Maximum is then defined by:

$$F_s = max(\widetilde{d}_1, \widetilde{d}_2 \cdot \cdot, \widetilde{d}_k). \tag{5.33}$$

4. **Sum-weighting-score (WHT):** The weighted sum of scores consists at the extension of the sum of the scores. Indeed, the score of each system is weighted and based on the error rate associated with it, based on performance individual system or its importance in the multimodal system. The fusion of scores is calculated as follows

$$F_s = \sum_{i=1}^{k} w_i \widetilde{d}_i, \tag{5.34}$$

with k is the number of combined biometric sub-systems and the weight of i^{th} sub-system, w_i is defined as

$$w_i = \frac{1}{\sum_{i=1}^{k} \frac{1}{\mathcal{E}_i}} \times \frac{1}{\mathcal{E}_i}, \tag{5.35}$$

where $\mathcal{E}_i$ denote the Equal Error Rate (EER) of each biometric sub-systems and $\sum_{i=1}^{k} w_i = 1$.

5.7 Chapter Summary

In this chapter, the proposed FKP system based identification is having several advantages, which is rich in texture features, easily accessible, invariant to emotions and other behavioral aspects as tiredness, stable features and acceptability in the society. Also, we presented the methodology of the proposed multimodal biometric system and gave detailed explanation of the various modules of this system; beginning by the implementation of PCANet and DCTNet methods to extract the features from FKP traits. Moreover, these feature vectors obtained are classified by SVM and KNN classifiers respectively, and stored in database as templates.

Lastly in this chapter, the fusion is performed at matching scores level by using sum fusion rules. The proposed method has been successfully implemented for FKP authentication system based on deep learning descriptors and evaluated its performance by using the universal FKP database. This eligibility can be confirmed further after reviewing the results presented in the next chapter.

Chapter 6

EXPERIMENTATIONS AND RESULTS

6.1 Introduction

BIOMETRIC identification system can work into two modes, open-set identification and closed-set identification. In our study, the proposed methods were tested through the two modes. In open-set mode, the system indicates that the person presenting the acquired biometric data is an enrolled person or not in the system database. In the closed-set mode, the system select the identity of the person whose reference has the highest degree of similarity with the acquired biometric data. In this chapter, the identification tests results are divided into three parts. In the first part, a series of experiments were carried out to select the best parameters of our deep learning algorithms (PCANet and DCTNet) as a number of layers, a number and size of filters and block overlap percentage, yield the best performance. The tests results in the second part are devoted to evaluate the performance of the unimodal and multimodal biometric systems. Finally, the last part shows the comparative study between classical and deep learning methods based on test results of identification systems performance.

6.2 Database Description

The proposed biometric system investigates a personal authentication technique using finger-knuckle-print (FKP) database from the Poly University [78]. In Annex B, a specific data acquisition device is developed to capture the FKP images. The local convex direction map of the FKP image is then extracted, based on which a coordinate system is defined

to align the images and a region of interest (ROI) is cropped for feature extraction and matching [60]. FKP ROI database is established to evaluate the performance and to choose their appropriate parameters. In FKP ROI database, the person was asked to provide 12 image samples for each of Left Index Fingers *LIF*, Left Middle Fingers *LMF*, Right Index Fingers *RIF* and Right Middle Fingers *RMF*. Therefore, 48 image samples from 4 finger types were collected from each person. The Annex A provides a detailed explanation of the FKP ROI database.

6.3 Experimental Setup

In the system-design process, three images are randomly selected of twelve images of each person were used in the enrolment stage to create the system database; the remaining nine images were used to evaluate the system performance. Thus, a total of 245025 comparisons (database size equal to 165) were made. The genuine experiments were performed by comparing three test images with the corresponding class in the database in which 1485 genuine scores were made. Similarly, nine images with each class, for impostor experiments, were compared with all references in database which give 243540 impostor experiments. Thus, an identification occurs when the biometric system attempts to determine the identity of a person for that, a feature vector is extracted and compared with all the vectors in the system database. The main objective of this part is to choose the best configuration of the PCANet and DCTNet methods for the FKP biometric modalities. To choose the best configuration, we need to examine the impact of each parameter on systems performance by using *LIF* modality. To select the best parameters, it can be divided the identification tests results into two sub-sections. Firstly, we present the performance of the unimodal biometric systems based on PCANet technique in order to select the best performance. Subsequently, in the last sub-section, the identification system performance based on DCTNet technique are tested. As mentioned above, these tests are based on *LIF* modality, to adapt our deep learning methods.

6.3.1 PCANet Parameters Selection

Most of algorithms in deep learning have a set of specific parameters because most of the parameters are optional and preset to meet our specific needs. For that, we conducted a series of experiments to select the best parameters of PCANet which give a minimum errors rates. Thus, in our work, we directly choose the number of stages equal to two-stages. In the first stage, we change every time the number of PCA filters from 1 to 3 with a fixed number of PCA filters in second stage which equal to 2. The problem we are addressing is as follows: we want to choose the number of $[K_1, K_2]$ PCA filter and the block overlap percentage such

that the EER is minimized. By varying the filter number from 1 to 3 in first stage and block overlap percentage from 0% to 75%, which illustrated in Table. 6.1, the system performance as a function of the number of PCA filters selection in each stage for various block overlap percentage. The reason of presenting Table. 6.1 is to show how these parameters might have an effect on the performance of our system. From this table, we observe that the identification accuracy becomes very high at $[3, 2]$ filter number (3 filters for first stage and 2 filters for second stage), where it actually exceeds 98.788% (EER = 1.212%) and a strong decrease in identification accuracy when we go to a lower size, it is important to note that, the filter size used for this test is 5×5. Also, another series of tests is performed to select the optimum filter size with their corresponding block overlap percentage. Thus, several sizes are tested which are $(3 \times 3, 5 \times 5, 7 \times 7, \cdots, 13 \times 13)$, and the given results are plotted in Fig. 6.1. From this figure, it is very clear that a filter size equal to $[13 \times 13]$ and an overlapped block equal to 50%, are enough to achieve a good accuracy as a very lower EER equal to 0.919 %.

Overlapped blocks	Number Filters$[1, 2]$		Number Filters$[2, 2]$		Number Filters$[3, 2]$	
	T_o	EER	T_o	EER	T_o	EER
0%	0.671	8.790	0.661	8.559	0.690	1.572
25%	0.653	8.228	0.652	7.867	0.671	1.641
50%	0.665	6.646	0.650	7.272	0.676	1.481
75%	0.639	5.583	0.644	6.060	0.692	1.212

TABLE 6.1: The *PCANet* parameters test results.

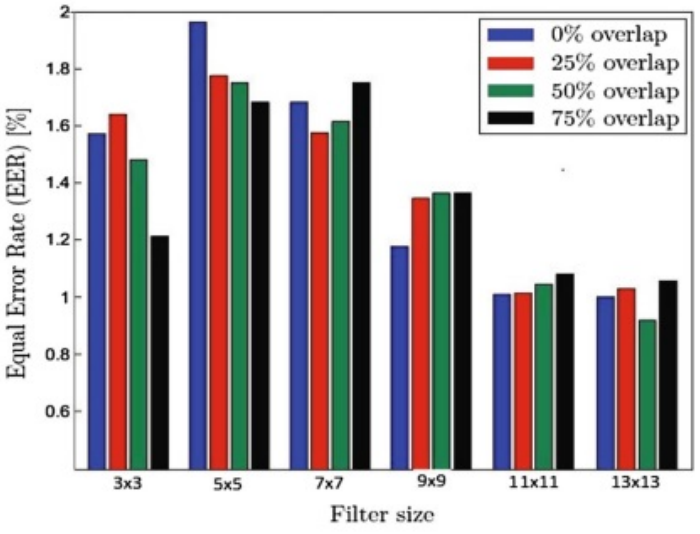


FIGURE 6.1: The *PCANet* parameters test results.

However, the results given in this part demonstrate that a filter number equal to 3 and

2 in the first and second stage, respectively, a filter size of (13×13) and a block overlap, percentage of 50% can offer better results in terms of system accuracy.

6.3.2 DCTNet Parameters Selection

As in PCANet parameters selection, the performance of DCTNet model depends on the good tuning of deep learning parameters. DCTNet algorithm also has generic parameters, such as: number of layers, number of filters and blocks size, etc. To select DCTNet algorithm parameters, we need to a new series of experiments that provide required parameters for the best performance. But these initialization procedures take more time and more expensive cost. In our case and in view of the great similarity between the two methods, we directly choose the same PCANet parameters for DCTNet method. Based on previous sub-section, DCTNet parameters can be summaried in: the number of layers equal to 2 layers, the filters number are $[3, 2]$, 3 filters for first stage and 2 filters for second stage and the blocks size equal to (13×13). Therefore, we have decided to choose these parameters in the rest of our study.

6.4 Biometric System Evaluation

6.4.1 Unimodal Systems Test Results

6.4.1.1 PCANet Based Biometric Systems

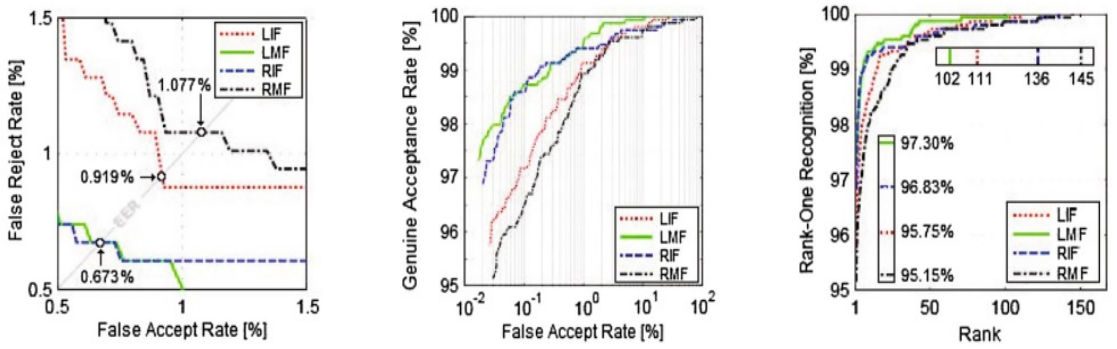
This sub-section describes the results of the proposed PCANet based identification unimodal system. When we use individually the information from four modalities *LIF*, *LMF*, *RIF* and *RMF* of each person. Thus, we can see in Table. 6.2 the test results of PCANet systems for all finger types.

Open-set: From this table, it's clear that the *LMF* and *RIF* fingers offer better results in terms of the EER. In this case, the identification system can achieve an EER of 0.673% at a threshold $T_0 = 0.715$ and $T_0 = 0.703$ for *LMF* and *RIF* modalities, respectively. Also in this table, we can observe the *LIF* modality gives EER = 0.919% at a threshold $T_0 = 0.705$ for *LIF* modality. Finally, in the case of using *RMF* modality, EER = 1.077% with $T_0 = 0.687$. So, the performance of our system is very acceptable compared with several state-of-art of FKP based biometric identifier accuracies where we can justified by the efficiency of the proposed method. The ROC curves for four fingers modalities are shown in Fig. 6.2.*(a)*, which plot the False Rejected Rate (FRR) against the False Accept Rate (FAR). The test results indicate that the *LMF* and *RIF* modalities are very efficiency at the EER point and their performances are equal. These modalities is better than the perform of the *LIF* and

RMF modalities in terms of EER. In Fig. 6.2.*(b)*, the ROC curves (Genuine Acceptance Rate (GAR) against FAR) provide a more details for the performance of proposed unimodal *open-set* identification systems.

MODALITIES	OPEN-SET IDENTIFICATION		CLOSED-SET IDENTIFICATION	
	T_o	EER	ROR	RPR
LIF	0.705	0.919	95.750	111
LMF	0.715	0.673	97.300	102
RIF	0.703	0.673	96.830	136
RMF	0.687	1.077	95.150	145

TABLE 6.2: Unimodal identification test results based on PCANet.



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

FIGURE 6.2: Unimodal *open/closed-set* identification test results for PCANet method. *(a)* The ROC curves (FRR against FAR), *(b)* The ROC curves (GAR against FAR) and *(c)* The CMC curves.

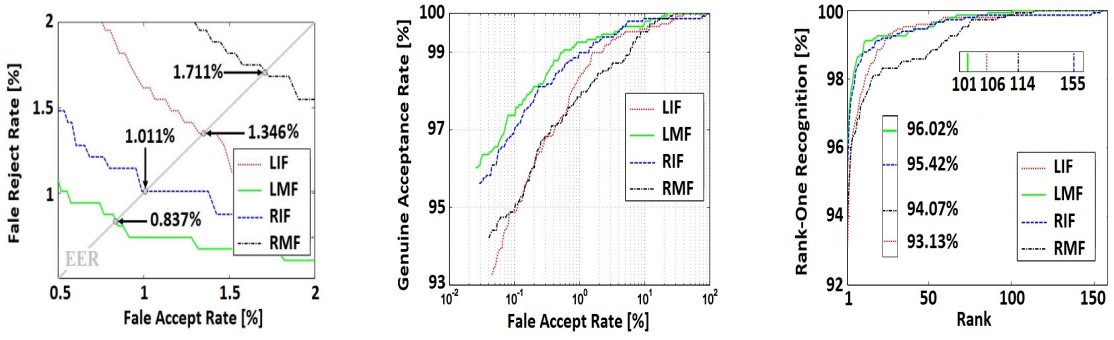
Closed-set: At *closed-set* identification systems tests, we compare the performance of the different fingers modalities to determine the best modality. The results for all fingers are also presented in Table. 6.2. From analyzing this table, we can see that the Rank-One Recognition (ROR) is between 95.150% and 97.300%. So, the system can achieve higher accuracy at the *LMF* modality compared with the other fingers modality which is produced a ROR equal to 97.300% with a Rank of Perfect Recognition (RPR) of 102. The *RIF* followed by *LIF* and *RMF* modalities can produce the ROR equal to 96.830% (RPR = 136), 95.750% (RPR = 111) and 95.150% (RPR = 145), respectively. To summarize the *closed-set* identification experiments, graphs showing the Cumulative Match Characteristics (CMC) curves using all unimodal systems were generated in Fig. 6.2.*(c)*. In conclusion, the obtained identification rates of FKP modalities for PCANet system are very efficiency.

6.4.1.2 DCTNet Based Biometric Systems

Open-set: From four modalities *LIF*, *LMF*, *RIF* and *RMF* of persons, the DCTNet based identification unimodal systems are obtain the results which described in Table. 6.3. In general, this results are very acceptable that its given a low EER rates. In the *LMF* case, the *open-set* identification system achieve better results with EER equal to 0.837% at a threshold $T_0 = 0.232$. Also according to the table, we can observe that the *RIF* modality gives EER = 1.011% at a threshold $T_0 = 0.250$ after that the use of *LIF* and *RMF* modalities can produced an EER of 1.346% at $T_0 = 0.259$ and EER of 1.711% at $T_0 = 0.307$, respectively. Finally, the performance of our DCTNet system for all fingers modalities are shown in Fig. 6.3.*(a)*, which plot the FRR against the FAR. The ROC curves in Fig. 6.3.*(b)*, the plot of GAR against FAR provide a clearly comparative between the performance of four fingers modalities in DCTNet based unimodal *open-set* identification systems.

MODALITIES	OPEN-SET IDENTIFICATION		CLOSED-SET IDENTIFICATION	
	T_o	EER	ROR	RPR
LIF	0.259	1.346	93.131	106
LMF	0.232	0.837	96.026	101
RIF	0.250	1.011	95.420	155
RMF	0.307	1.711	94.074	114

TABLE 6.3: Unimodal identification test results based on DCTNet.



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

FIGURE 6.3: Unimodal *open/closed-set* identification test results for DCTNet method. *(a)* The ROC curves (FRR against FAR), *(b)* The ROC curves (GAR against FAR) and *(c)* The CMC curves.

Closed-set: The performances of all different modalities are presented in Table. 6.3. At *closed-set* identification tests, the results can provide that the ROR is between 93% and 96%. So, the system can achieve a higher accuracy in the *LMF* modality compared with the other fingers modalities, which is produced a ROR equal to 96.026% with a RPR equal to 101. As the *RIF* modality gives the ROR = 95.420% and RPR = 155, followed by *RMF* and *LIF*

modalities which can produce a ROR equal to 94.074% with RPR = 114 and ROR = 93.131% with RPR = 106, respectively. To summarize the *closed-set* identification experiments, the CMC curves in Fig. 6.3.*(c)* present the obtained identification rates for FKP modalities.

6.4.1.3 Comparative Study

In order to show the effectiveness of the deep learning methods, a comparison study is performed based on two methods: PCANet and DCTNet algorithms. As everyone knows, the deep learning algorithms contain two parts: feature extraction part and classification part. Thus, in this comparison, PCANet and DCTNet as a feature extraction methods are used, that can capture the information from the image texture and provide a very robustness to effectively describe the image characteristics. But in same time, for PCANet method, we change the SVM classifier in classification part by using two another classifiers: Radial Basis Function (RBF) and Random Forest Tree (RFT), their obtained results are compared with DCTNet method which are used KNN classifier. In the identification tests, these test results are giving an idea about the better combination which will entirely depend at proposed process. As such, the set of experiments tends to the performance of unimodal systems. The objective of this section is to choose the best performance of unimodal system.

In *open-set* identification mode, the obtained results of the identification tests are given in terms of EER. Table. 6.4 shows the baseline obtained results by using the four combination algorithms. These results demonstrate the capability of deep learning methods to reduce identification error rate. From this table, it is observed that the use of PCANet-SVM system leads to reduce the lowest EER offered than the rest unimodal systems. However seen that, the capability is considerably improved through the *LMF* modality (EER = 0.673% and T_0 = 0.715). Regarding to the DCTNet-KNN combination of *LMF* unimodal system is also provided a good results and close to the PCANet-SVM combination (EER = 0.837% and T_0 = 0.232). On the other hand, it is observed that the use of rest systems produces limited efficiency. As the *LMF* modality gives the performance (EER = 1.305%, T_0 = 0.673) in PCANet-RBF system and (EER = 1.547%, T_0 = 0.509) in PCANet-RFT system.

In *closed-set* identification mode, the second mode of identification was tested for all methods and the results are reported in Table. 6.4 to comparison. Similarly, it is clear that the using of PCANet-SVM improves as big form the performance of our system, as compared with the PCANet-RBF and PCANet-RFT systems. In the case of PCANet-SVM method, a ROR equal to 97.30% with a lowest RPR equal to 102, which it is achieved by the ue of *LMF* modality. Using PCANet-RBF method, ROR was 94.41% with RPR = 139 and

PCANet-RFT method gives the result ROR = 92.12% and RPR = 147 for the *LMF* modality in database of 165 persons. In DCTNet-KNN case, the performance of *LMF* system is the second in terms of efficiency through a ROR equal to 96.02% and RPR = 101. Based on the results, it should be noted that absolutely, the system based on PCANet deep learning method with SVM classifier is very efficiency than these another methods DCTNet-KNN, PCANet-RBF and PCANet-RFT.

METHODS	MODALITIES	OPEN-SET IDENTIFICATION		CLOSED-SET IDENTIFICATION	
		T_o	EER	ROR	RPR
PCANet-SVM	LMF	0.715	0.673	97.30	102
DCTNet-KNN	LMF	0.232	0.837	96.026	101
PCANet-RBF	LIF	0.628	2.626	90.90	157
	LMF	0.673	1.305	94.41	139
	RIF	0.788	3.636	85.25	152
	RMF	0.733	3.271	88.28	133
PCANet-RFT	LIF	0.469	2.424	89.42	123
	LMF	0.509	1.547	91.78	133
	RIF	0.487	1.616	91.91	134
	RMF	0.434	2.223	92.12	147

TABLE 6.4: Comparison study for unimodal identification test results.

6.4.2 Multimodal Systems Test Results

Recently, the research in the field of biometrics for identification purposes, has increasingly investigated the use of multiple biometric modalities (multimodal biometrics). Multimodal biometrics refers to the use of more than one biometric modality for person identification. The multimodal systems are expected to be more reliable due to the presence of multiple templates security. A number of these systems have been proposed and differ from one to another in terms of their architecture, the number of modalities, the choice of modalities and the methods used for the information fusion.

The objective of this sub-part is to evaluate and to improve the performance of the unimodal biometric identification system by using multiple modalities information from the different finger types. The important keys to improve the accuracy of multimodal biometric system are the choice of fusion level as well as the technique deployed for data fusion. In our work, we choose only the matching score level because it's usually preferred as it's relatively easy and it can easily combine the scores presented by the different modalities. The idea behind using fusion at matching score level is the possibility to combine the scores obtained from different fingers modalities with a simple rules. The overall score is then sent to the decision module for accepting or rejecting a person.

6.4.2.1 Multi-Samples Biometric Systems

The Multi-Samples systems use multiple samples of the same biometric trait by only a single sensor. For example, our multi-samples systems are using two fingers of the left hand *LIF-LMF*, two fingers of the right hand *RIF-RMF* and four fingers *LIF-LMF-RIF-RMF* to simplify, it's noted ALL. Firstly, we present the performance of the multi-samples biometric systems based on PCANet method in order to evaluate their performance. After that, we give the performance of the multi-samples biometric systems based on DCTNet method.

a. PCANet Based Biometric Systems

COMBINATION	SUM		MIN		MAX		WHT	
	T_o	EER	T_o	EER	T_o	EER	T_o	EER
LIF-LMF	0.856	0.022	0.765	0.404	0.817	0.043	0.791	0.049
RIF-RMF	0.803	0.044	0.722	0.673	0.745	0.067	0.727	0.098
ALL	0.736	0.000	0.745	0.538	0.939	0.000	0.718	0.000

TABLE 6.5: PCANet Performance of the multimodal open-set identification system (fusion at matching score level).

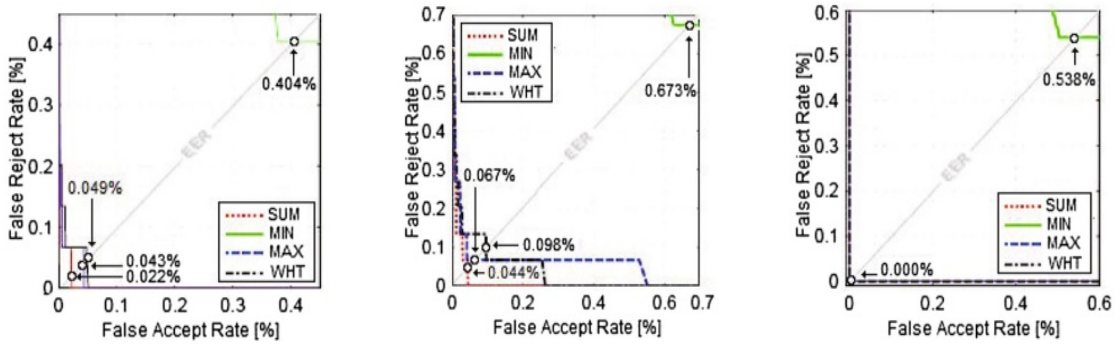


FIGURE 6.4: The ROC curves of PCANet for multimodal *open-set* identification test results (fusion at matching score level). *(a)* The *LIF-LMF* combination, *(b)* The *RIF-RMF* combination and *(c)* The ALL combination.

Open-set: Generally, the use of fusion process can improve the performance of system and provide a better result than the best unimodal identification system. Thus, to find the better performance of combinations with fusion rules, Table. 6.5 provides the results of *open-set* identification mode. From this table, we can observe that the ALL combination with almost any fusion rule (*Sum*, *Wht* and *Max*) reduces the EER to zero (100% improvement). But in the *Min* case, EER was reduced only to 0.538% ($\simeq$ 80% improvement). In *LIF-LMF* and *RIF-RMF* combinations, the EERs are very close (between 0.022% and 0.098%) in the cases

of *Sum*, *Wht* and *Max* rules and (0.404% to 0.673%) into *Min* rule. The ROC curves in Fig. 6.4.*(a)*, Fig. 6.4.*(b)* and Fig. 6.4.*(c)* present a direct comparison of the obtained performances by using four rules of fusion based on *LIF-LMF*, *RIF-RMF* and ALL combinations, respectively. Thus, the combination of higher number of finger types give a considerable improvement especially to using *Sum*, *Wht* and *Max* rules.

COMBINATION	SUM		MIN		MAX		WHT	
	ROR	RPR	ROR	RPR	ROR	RPR	ROR	RPR
LIF-LMF	99.73	10	97.84	73	99.66	7	99.79	16
RIF-RMF	99.59	12	96.90	123	99.32	9	99.46	22
ALL	100.00	1	98.11	86	100.00	1	100.00	1

TABLE 6.6: PCANet Performance of the multimodal closed-set identification system (fusion at matching score level).

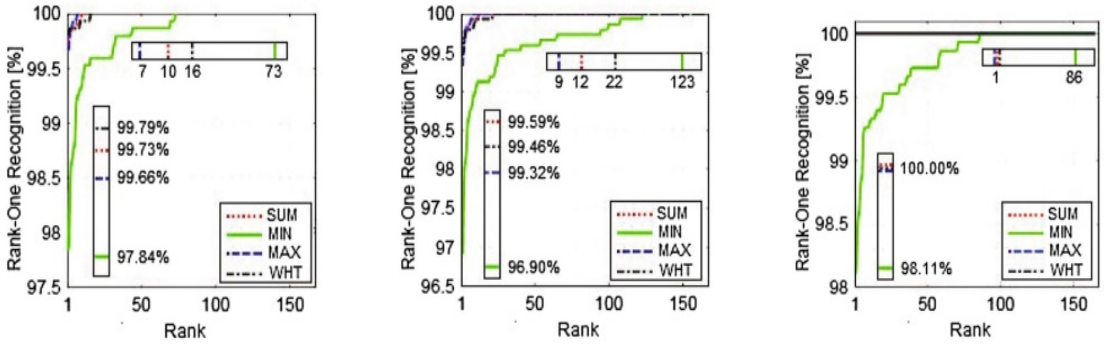


FIGURE 6.5: The CMC curves of PCANet for multimodal *closed-set* identification test results (fusion at matching score level). *(a)* The *LIF-LMF* combination, *(b)* The *RIF-RMF* combination and *(c)* The ALL combination.

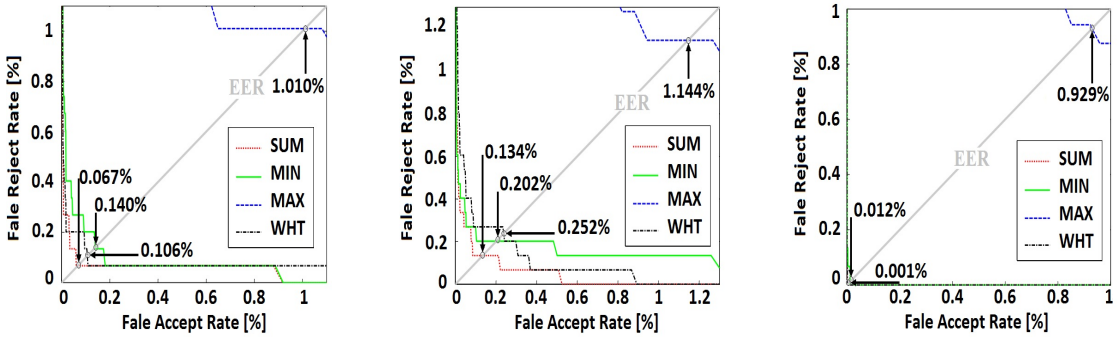
Closed-set: To validate our idea, we have run other tests for the *closed-set* identification mode. Table. 6.6 presents the results of different combinations and fusion rules to determine the best combination as well as the best fusion rule. The experiments indicate that the identification rates ROR for the ALL combination, with *Sum*,*Wht* and *Max* rules, are greater than the corresponding ones in unimodal system, so as ROR is given as 100.00% with lowest RPR of 1 in both cases. In the *Min* rule, the system is provided ROR equal to 98.110% with RPR equal to 86 but it always improve the performance. On the other hand, although the fusion rules in the *closed-set* identification case, the results in *LIF-LMF* and *RIF-RMF* combinations did not live up to higher performance such as the ALL combination. The rest of results can be clearly seen in Table. 6.6. Also, the curves in Fig. 6.5 which plot the CMCs curves for all cases, demonstrate the capability to reduce the closed-set identification error rates by combining all fingers at the matching score level. Fig. 6.5.*(a)* presents comparison of the different rules fusion based on *LIF-LMF* combination and the same presentation

Fig. 6.5.*(b)*, Fig. 6.5.*(c)* for *RIF-RMF* and ALL combinations, respectively.

b. DCTNet Based Biometric Systems

COMBINATION	SUM		MIN		MAX		WHT	
	T_o	EER	T_o	EER	T_o	EER	T_o	EER
LIF-LMF	0.170	0.067	0.182	0.140	0.263	1.010	0.199	0.106
RIF-RMF	0.212	0.134	0.258	0.202	0.263	1.144	0.209	0.252
ALL	0.159	0.001	0.158	0.012	0.258	0.929	0.193	0.001

TABLE 6.7: DCTNet Performance of the multimodal open-set identification system (fusion at matching score level).



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

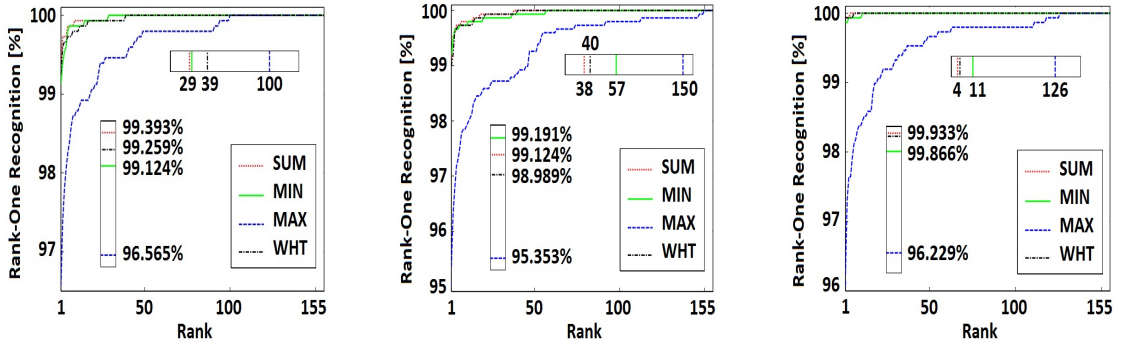
FIGURE 6.6: The ROC curves of DCTNet for multimodal *open-set* identification test results (fusion at matching score level). *(a)* The *LIF-LMF* combination, *(b)* The *RIF-RMF* combination and *(c)* The ALL combination.

Open-set: The use of DCTNet method provide genuine confirmation of efficacy of deep learning methods, through the test results of *open-set* identification which are presented in Table. 6.7. From this table, we can see that also the ALL combination with fusion rules *Sum* and *Wht* reduce the EER to zero (0.001%). But in the *Min* and *Max* cases, EER equal only to 0.012% and 0.929%, respectively. In *LIF-LMF* combination, the EERs are between 0.067% and 0.140% for the *Sum*, *Wht* and *Min* rules, and 1.010% into *Max* rule. For *RIF-RMF* combination, the rest results can be observed in Table. 6.7. The ROC curves in Fig. 6.6.*(a)*, Fig. 6.6.*(b)* and Fig. 6.6.*(c)* present a direct comparison of the obtained performances by using the all fusion rules based on *LIF-LMF*, *RIF-RMF* and ALL combinations, respectively. Thus, the combination of All finger gives a considerable improvement especially for using *Sum* and *Wht*rules.

Closed-set: For the *closed-set* identification mode, Table. 6.8 contains the results of different combinations and fusion rules in DCTNet based multimodal systems. From this table,

COMBINATION	SUM		MIN		MAX		WHT	
	ROR	RPR	ROR	RPR	ROR	RPR	ROR	RPR
LIF-LMF	99.393	29	99.124	29	96.565	100	99.259	39
RIF-RMF	99.124	38	99.191	57	95.353	150	98.989	40
ALL	99.933	4	99.866	11	96.229	126	99.933	4

TABLE 6.8: DCTNet Performance of the multimodal closed-set identification system (fusion at matching score level).



(a) (b) (c)

FIGURE 6.7: The CMC curves of DCTNet for multimodal *closed-set* identification test results (fusion at matching score level). *(a)* The *LIF-LMF* combination, *(b)* The *RIF-RMF* combination and *(c)* The ALL combination.

the ROR rate of ALL combination, with *Sum* and *Wht* rules, is greater than the rest systems, as ROR equal to 99.933% with lowest RPR of 4 in both cases. In *Min* and *Max* rules, the systems are provided ROR = 99.866% with RPR = 11 and ROR = 96.229% with RPR = 126, respectively. Furthermore, the fusion rules in *LIF-LMF* and *RIF-RMF* combinations did not live up to higher performance such as the ALL combination. The rest results of these combinations are shown in Table. 6.8. Finally, the CMCs curves plot the closed-set identification error rates for all cases and demonstrate the efficient of the matching score level. Where Fig. 6.7.*(a)* presents a comparison between the different rules fusion based on *LIF-LMF* combination and the same presentation Fig. 6.7.*(b)*, Fig. 6.7.*(c)* for *RIF-RMF* and ALL combinations, respectively.

6.4.2.2 Multi-Algorithms Biometric Systems

The Multi-Algorithms systems process biometric trait by different feature extraction algorithms. Then, the individual results from each matcher are combined to obtain the final decision. From the proposed systems, multi-algorithms study is performed based on PCANet and DCTNet algorithms through the focus of a best performance in the two methods. Factually, these PCANet and DCTNet algorithms operate on the same fusion of all fingers with

same rules.

METHODS	COMBINATION	SUM		MIN		MAX		WHT	
		T_o	EER	T_o	EER	T_o	EER	T_o	EER
PCANet-SVM	ALL	0.736	0.000	0.745	0.538	0.939	0.000	0.718	0.000
DCTNet-KNN	ALL	0.159	0.001	0.158	0.012	0.258	0.929	0.193	0.001
		ROR	RPR	ROR	RPR	ROR	RPR	ROR	RPR
PCANet-SVM	ALL	100.00	1	98.11	86	100.00	1	100.00	1
DCTNet-KNN	ALL	99.933	4	99.866	11	96.229	126	99.933	4

TABLE 6.9: Multi-Algorithms biometric systems test results.

Based on the presented results in the Table. 6.9, we can judge on the great affinity and similarity of the excellent performance in two algorithms. For instance, the results obtained in PCANet-SVM system are given a good rate EER = 0.000% that means a efficiency 100% while the DCTNet-KNN is also provided a rate of efficiency equal to 99.99% (EER = 0.001%) in *open-set* identification mode. Always in multimodal biometric systems, the results of Table. 6.9 prove again that the power of deep learning methods by high recognition rates in *closed-set* identification mode. In conclusion, the obtained identification rates of FKP modalities for the proposed systems are very efficiency especially for talking about PCANet algorithm.

6.4.2.3 Hybrid Biometric Systems

Hybrid systems are composed of several scenarios concern other types of biometric systems, therefore it has many advantages. Thus, our hybrid systems use a multiple FKP modalities with several deep learning algorithms such as PCANet and DCTNet. The idea is the same presented idea in the Section. 6.4.1.3, but the difference it concerns the type of multimodal systems. For comparison, PCANet method using SVM, RBF and RFT classifiers with DCTNet method and KNN classifier, produce the set of test results. Then it provides an idea on a better combination of *open-set* and *closed-set* multimodal biometrics system. At the matching score level fusion, it is possible to combine scores obtained from different fingers modalities by using rules: *Sum*, *Min*, *Max* and *Wht*.

For *open-set* identification mode, the all combinations with fusion rules are presented as EER in Table. 6.10. The results in this table show that the use of PCANet-SVM is better than the PCANet-RBF and PCANet-RFT systems. Moreover, it is observed that the PCANet-SVM in ALL combinations with fusion rules *Sum*, the *Max* and *Wht* successfully reduces the EER to zero for the fused biometrics, which in this case is efficiency than several previous works obtained by using FKP biometric. The use the RBF and RFT classifiers

METHODS	COMBINATIONS	SUM	MIN	MAX	WHT
PCANet-RBF	LIF-LMF	0.237	1.537	0.269	0.269
	RIF-RMF	0.740	2.750	0.673	0.673
	ALL	0.237	18.19	0.300	0.269
PCANet-RFT	LIF-LMF	0.202	0.606	0.538	0.202
	RIF-RMF	0.269	0.657	0.595	0.336
	ALL	0.002	0.353	0.213	0.002
PCANet-SVM	ALL	0.000	0.538	0.000	0.000
DCTNet-KNN	ALL	0.001	0.012	0.929	0.001

TABLE 6.10: Hybrid multimodal open-set identification test results.

provides a considerable improvement EER equal to 0.237% and 0.002% for the ALL combination at *Sum* rule, respectively. Regarding to the DCTNet-KNN method of ALL modalities is also provided a good results and very close to the PCANet-SVM system. The Table. 6.10, presents a direct comparison of the obtained results by using the combinations at matching score level fusion.

METHODS	COMBINATION	SUM		MIN		MAX		WHT	
		ROR	RPR	ROR	RPR	ROR	RPR	ROR	RPR
PCANet-RBF	LIF-LMF	98.58	50	95.42	157	97.77	32	98.51	88
	RIF-RMF	96.76	116	90.30	146	95.48	120	96.70	118
	ALL	98.58	50	91.91	157	97.77	32	98.51	88
PCANet-RFT	LIF-LMF	98.18	62	96.83	117	91.04	69	97.84	81
	RIF-RMF	98.38	49	96.83	125	91.24	78	98.31	45
	ALL	99.93	4	98.24	67	84.44	5	99.86	5
PCANet-SVM	ALL	100.00	1	98.11	86	100.00	1	100.00	1
DCTNet-KNN	ALL	99.99	4	99.86	11	96.22	126	99.99	4

TABLE 6.11: Hybrid multimodal closed-set identification test results.

For the *closed-set* identification mode, the test results are presented in Table. 6.11. As the experimental results show in PCANet-SVM case, the identification rate provides a ROR = 100% and RPR = 1 in the ALL combination for *Sum*, *Max* and *Wht* rules. Always, PCANet-SVM case is better than the another PCANet-RBF and PCANet-RFT cases. On the other hand, DCTNet-KNN method of ALL combination for *Sum* and *Wht* fusion rules gave ROR equal to 99.99%, RPR = 4. Additionally, the PCANet-RFT is provided very efficiency accuracy in this mode, the *Sum* rule for the ALL combination is provided ROR equal to 99.93% with RPR equal to 4. As in PCANet-RBF case, the ALL combination provides cogent performance such as ROR = 98.58% and RPR = 50 by using *Sum* rule. The rest results can be presented in Table. 6.11 to do a comparison of different methods based on all combinations of FKP modalities at matching score level fusion.

6.5 Classical Vs Deep Learning Methods

	Reference	Comments	EER[%]	Recognition[%]
Classical Methods	[87] L. Zhang	Competitive code	1.48	98.52
	[88] L. Zhang	Competitive code-MagCode	1.09	98.91
	[89] L. Zhang	The Fourier transform and LGIC	0.35	99.65
	[90] A. Kumar	Score-Level fusion	1.39	98.61
	[93] L. Damon	Average fusion rule	5.50	94.50
	[95] L. Zhang	RieszCompCode	1.26	98.74
	[98] A. Kumar	The LDA for (PCA-ICA) features	1.95	98.05
	[99] L. Zhang	Gabor-CompCod	1.72	98.28
	[100] A. Meraoumia	The 2D-Block based DCT	0.20	99.80
DLM	Our	**Proposed PCANet method**	0.00	100.00
	Our	**Proposed DCTNet method**	0.00	100.00

TABLE 6.12: The comparison of our results with classical literature results.

Many authentication systems have been used widely in recognition applications. Based on the research reports, the finger-knuckle is significantly rich in texture due to skin folds and creases. The advantages of using FKP are the texture features, easily, contact-less image acquisition and acceptability. These features and advantages make the using of FKP as powerful biometric identifier. The researchers have made many efforts to build a FKP system based on classical methods to features extraction and classification. This section presents a comparison study between classical methods and deep learning methods. For more credibility, we have focused on researches that using the same FKP database which has used in our proposed systems.

The Table. 6.12 shows the most obtained results from these literatures. For example, in [98], the FKP features are extracted by using classical PCA, LDA and ICA methods, however, the recognition rate have not been achieved just 98.05%. On the other hand, the our system based on deep learning method (PCANet) achieved a perfect performance, the recognition rate equal to 100%. For confirmation, [100] has proposed a identification system which is designed a fusion of the features extracted by classical 2D-Block DCT method, the recognition rate in this system is good (99.80%) but it did not rises to contest the high performance of our system based on deep learning method (DCTNet). This excellence is due to advantages of deep learning methods that which have feature engineering part. Moreover, deep learning has a capacity to extract deep features and classified into high-level features representation. Through the presented results, it is clearly shown that the proposed systems of our thesis performed much better efficiency based at matching score fusion of the FKP modalities.

6.6 Chapter Summary

The proposed FKP modalities based multimodal system has been developed and its performance has been evaluated at matching score fusion approaches. The tested database consists FKP ROI images of 165 persons from the Hong Kong Polytechnic University (PolyU). The quality of this database is good as the variation of the different finger images are very acceptable. From many experiments of the *LIF* modality, the different PCANet parameters are tested to select the best authentication performance. These parameters are as: a filters number equal to 3 and 2 in the first and second stage, respectively, a filter size of (13×13) and a block overlap percentage of 50%. For DCTNet parameters, we directly decided to based on these previous parameters because the DCTNet and PCANet methods have a great similarity.

Different obtained results from various experimentations have been illustrated by using unimodal and multimodal identification systems. Depending on the results, the recognition performance of the unimodal biometrics is slightly lower than the performance of the multimodal biometrics. Further, The performance of the identification system is significantly improved by using the fusion of all finger types and can give a GAR equal to 100% while the unimodal identification give only a GAR equal to 97.300%. Thus, it's clear that the multimodal identification systems demonstrate the efficiency and give an excellent identification rate and a higher accuracy. Further, from the FKP database, the performances of the proposed PCANet and DCTNet based on matching score fusion rules can be summarized in equal error rates (EER), which EER is the point in a ROC curve where the FAR and FRR are equal, that these performances achieved a perfect rates (EER $\simeq$ 0.00%).

Finally, this chapter presented a transparency comparison study between classical methods and competitive methods of deep learning. Depending to the results of our study, we can say that the deep learning methods have unparalleled success in recognition systems because these algorithms have strength architectures that allow the processing of detailed features superior to existing state of the art techniques and these architectures prove the ability of biometric security fields.

Chapter 7

CONCLUSIONS AND FUTURE WORKS

The conclusions chapter provides the thesis summary, presents a achieved contributions through this doctoral research and gives some possible future works in the research for multimodal biometric fusion.

7.1 Thesis Summary

The work in this thesis can be summarised as follows:

The thesis starts by introduction about information's security and biometrics' importance in our recent world. Biometric systems and challenges of these systems are also presented in chapter 1. Then, we presented the desired objectives of our thesis. A very brief methodology for the proposed multimodal biometric system is also presented in this chapter (Chapter 1).

The first step was to give the definition of biometric which refers to an automatic recognition of a person based on his/her behavioral or physiological characteristics. Many biometric traits used to build an authentication systems and are called biometric modalities. We described these modalities in more details based on classification their as physiological, behavioral and soft. Further, we have pointed out to the architecture of a biometric system and a functionalities of identification or authentication. Next, the performance metrics was summarized to evaluate the biometric systems, as the limitations of single biometric systems have been discussed with a various development issues (Chapter 2).

Multimodal biometrics alleviates many restrictions of single biometrics based on the use

of more than one source biometric traits. Due to the lower correlation between the sources, multimodal biometrics provides a maximum information gain for recognition systems. The different scenarios of multimodal biometrics provide a various options for researchers to design identification applications, through a systematical methods play an important role in improvement the performance of different combinations fusion methods and normalization techniques. Furthermore, the challenges of multimodal biometric systems are also discussed such as incompatible biometric traits, difficulty of system design and expensive cost (Chapter 3).

In Chapter 4, we have introduced the basics of biometric features and these features types. Moreover, we have presented a feature extraction methods that used in our thesis, start up from classical algorithms to deep learning networks. Also in the same chapter, we have given the motivated challenges to deep learning which can overcome the limitations of traditional feature extraction methods. As the end of chapter presented an overview of strengths and weakness of deep networks architectures.

The development procedures for the proposed FKP multimodal system have been illustrated in the Chapter 5. FKP is one of the most popular biometric modalities and has been used for the proposed recognition system because it is highly unique and the texture pattern make it a distinctly biometric identifier. In addition, this chapter presented the methodology of the proposed multimodal biometric system based on deep learning methods which it use the SVM and KNN. In this proposed system, a match score fusion is employed to improve the performance of biometric system.

Outcomes of the various experimentations have been presented and discussed in Chapter 6. The FKP dataset have been constructed from different fingers traits to evaluate the recognition performances by using a EERs, ROC and CMC curves of different unimodal and multimodal systems. Finally, based on our results, the comparative study between classical methods and deep learning methods has been illustrated in the last of Chapter 6.

7.2 Contribution to Knowledge

This section highlights on the contributions of this research to the development of a multimodal system and the critical security applications by using a simple deep learning techniques at matching score level fusion.

- In this doctoral thesis, we have fully developed and implemented a multimodal biometric system. This multimodal biometric system can overcome drawbacks associated with unimodal biometric systems.
- Several multimodal biometric systems have been developed with different biometric traits and fusion approaches. We have used finger knuckle traits from the hand region and using these traits allow the efficient and the convenient capturing of the biometric data.
- The performance of the our proposed systems give a GAR equal to 97.300% in unimodal identification. While by using the fusion of all finger are significantly improved the performance with a GAR equal to 100%. This fusion has potential to be efficiently used in different biometric application areas.
- We have developed a multimodal biometric system based on the FKP database that has a very high potential to be employed in various security critical applications.
- Our experimental results justify the use of deep learning methods which can be represent a new good direction of images processing applications, in the future years.

7.3 Future Research Works

The outcomes of this research have been published and presented through important venue Evolving Systems Journal and have benefitted both academic and enterprise applications. There are some issues and open questions left for future research.

To construct a more complicated and more sophisticated filters possibly or deeper by using more number of layers. Also, we will leave as future work to apply with a much larger database or different biometric traits. A true multimodal database is very useful for developing a reliable and efficient security application. Due to the type of collected sample data, the changes in the background and illumination are varied. True multimodal database with the identical conditions can be employed for further performance analysis.

More research can be conducted to find the optimum deep learning algorithms for unimodal biometrics to enhance the overall performance of the multimodal system. The level fusion scenarios (different fusion in different levels of the system) can be investigated to make the system faster and significantly to reduce the error rate. These represent possible future direction of research in this exciting and rich field.

Appendix A

FKP DATABASE

A.1 Overview

THE development of an identification biometrics systems involves the use of a database for the evaluation phase. During these last years, several databases, have been developed to evaluate the algorithms of biometric recognition. Thus, our experiment tests were performed using the FKP Database from the Poly University (The Hong Kong Polytechnic University) [78].

Among various kinds of biometric traits, hand based biometrics has been attracting considerable attention. Recently, it is found that the finger-knuckle-print (FKP), which refers to the inherent patterns of the outer surface around the phalangeal joint of one's finger, is highly unique and can serve as a distinctive biometric identifier. Abundant line-like textures are contained in an FKP image. The Biometric Research Centre (UGC/CRC) at The Hong Kong Polytechnic University has developed a real time FKP capture device (see Fig. A.1), and has used it to construct a large-scale FKP database. To advance research and to provide researchers working in the area of FKP recognition with a platform to compare the effectiveness of various FKP recognition algorithms, they published their FKP database, making it freely available for academic, noncommercial uses.

FIGURE A.1: FKP Capture Device.

A.2 The PolyU FKP Database Description

FKP images were collected from 165 volunteers, including 125 males and 40 females. Among them, 143 subjects were 20-30 years old and the others were 30-50 years old. We collected samples in two separate sessions. In each session, the subject was asked to provide 6 images for each of the left index finger, the left middle finger, the right index finger, and the right middle finger. Therefore, 48 images from 4 fingers were collected from each subject. In total, the database contains 7920 images from 660 different fingers. The average time interval between the first and the second sessions was about 25 days. The maximum and minimum intervals were 96 days and 14 days, respectively. Each folder is named as $nnn - finger - type$. nnn represents the identity of the person. In each folder, the first 6 images (01 to 06) were captured in the first session and the latter 6 images (07 to 12) were captured in the second session. FKP ROI.zip provide the extracted ROI images using ROI extraction algorithm described in [86].

Appendix B

FKP ROI EXTRACTION

B.1 Introduction

RECENTLY , it has been noticed that the texture in the outer finger surface has the potential to do personal authentication. This annex presents an algorithm for extraction the region of interest (ROI) based on finger-knuckle-print (FKP), which refers to the inherent skin pattern of the outer surface around the phalangeal joint of one's finger [60, 87]. A specially designed acquisition device is constructed to collect FKP images. In [86, 89], The proposed system captures the image around the finger knuckle area of a finger directly, which largely simplifies the following data preprocessing steps. Meanwhile, with such a design, the size of the imaging system can be greatly reduced, which improves their applicability. Since the finger knuckle will be slightly bend when it being imaged in the proposed system, the inherent finger knuckle print patterns can be clearly captured and hence the unique features of FKP can be better exploited.

B.2 FKP Acquisition Device

The FKP system is composed of an FKP image acquisition device and a data preprocessing module. From [60, 86, 87], the device (referring to Fig. B.1.*(a)*) is composed of a finger bracket, a ring LED light source, a lens, a CCD camera and a frame grabber. The captured FKP image is imputed to the data preprocessing module, which comprises basic step: ROI (region of interest) extraction. Refer to Fig. B.1.*(a)*, a basal block and a triangular block are used to fix the position of the finger joint [88]. The vertical view of the triangular block is illustrated in Fig. B.1.*(b)*. Fig. B.1.*(c)* shows a sample image acquired by the developed device [60].

(a) (b) (c)

FIGURE B.1: FKP image acquisition device. *(a)* Components of the acquisition device, *(b)* The position of the finger and *(c)* A sample FKP image.

B.3 ROI Extraction

It is necessary to construct a local coordinate system for each FKP image. With such a coordinate system, an ROI can be cropped from the original image for reliable feature extraction. The detailed steps for setting up such a coordinate system are as follows [60, 88, 89]:

Step 1 determine the $X-axis$ of the coordinate system. The bottom boundary of the finger can be easily extracted by a Canny edge detector. Actually, this bottom boundary is nearly consistent to all FKP images because all the fingers are put flatly on the basal block in data acquisition. By fitting this boundary as a straight line, the $X-axis$ of the local coordinate system is determined.

Step 2 crop a sub-image I_S. The left and right boundaries of I_S are two fixed values evaluated empirically. The top and bottom boundaries are estimated according to the boundary of real fingers and they can be obtained by a Canny edge detector.

Step 3 Canny edge detection. Apply the Canny edge detector to I_S to obtain the edge map I_E.

Step 4 convex direction coding for I_E, define an ideal model for FKP "curves". In this model, an FKP "curve" is either convex leftward or convex rightward. the pixels are coded on convex leftward curves as "1", pixels on convex rightward curves as "-1", and the other pixels not on any curves as "0".

Step 5 determine the $Y-axis$ of the coordinate system. For an FKP image, "curves" on the left part of phalangeal joint are mostly convex leftward and those on the right part are mostly convex rightward. Meanwhile, "curves" in a small area around the

phalangeal joint do not have obvious convex directions. Based on this observation, at a horizontal position x (x represents the column) of an FKP image, the "convexity magnitude" define as:

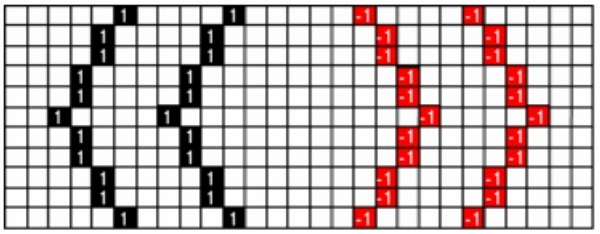


FIGURE B.2: Illustration for convex direction coding scheme.

Fig. B.2 illustrates this convex direction coding scheme and the pseudo codes are presented as follows:

Convex Direction Coding(I_E)

Output:I_{CD} (convex direction code map)
$y_{mid} =$ height of $\frac{I_E}{2}$;
for each $I_E(i,j)$:
if $I_E(i,j) = 0$
$I_{CD}(i,j) = 0$;
else if $I_E(i+1, j-1) = 1$ and $I_E(i+1, j+1) = 1$
$I_{CD}(i,j) = 0$;
else if $(I_E(i+1, j-1) = 1$ and $i \leqslant y_{mid})$ or $(I_E(i+1, j+1) = 1$ and $i > y_{mid})$
$I_{CD}(i,j) = 1$;
else if $(I_E(i+1, j+1) = 1$ and $i \leqslant y_{mid})$ or $(I_E(i+1, j-1) = 1$ and $i > y_{mid})$
$I_{CD}(i,j) = -1$;
end if
end for

$$conMag(x) = abs\left(\sum_{W} I_{CD}\right), \tag{B.1}$$

where W is a window being symmetrical about the axis $X = x$. W is of the size $d \times h$, where h is the height of I_S. The characteristic of the FKP image suggests that $conMag(x)$ will reach a minimum around the center of the phalangeal joint and this position can be used to set the $y - axis$ of the coordinate system. Let

$$x_0' = arg\,min\,(conMag(x)), \tag{B.2}$$

then $X = x'_0$ is set as the $X - axis$.

Step 6 crop the ROI image. Now that it has fixed the $X - axis$ and $Y - axis$, the local coordinate system can then be determined and the ROI sub-image I_{ROI} can be extracted with a fixed size.

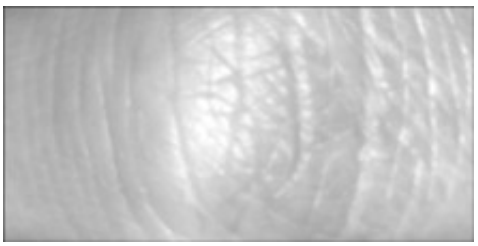


FIGURE B.3: An example of the extracted ROI image.

Fig. B.3 shows an example of the extracted ROI image.

Appendix C

MACHINE LEARNING CLASSIFIERS

C.1 Introduction

MACHINE Learning (ML) is an algorithm that is able to learn from data. Mitchell [103] provides this definition "A computer program is said to learn from experiences with respect to some class of tasks and performance measure , if its performance at tasks improves with experience". Machine learning allows us to tackle tasks that are too difficult to solve with fixed programs written and designed by humans. From a scientific and philosophical point of view, machine learning is interesting because the development of machine learning requires developing our understanding of a principles intelligence. The learning process itself is not the task, the learning is a means of attaining the ability to perform the task. For example, if we want a robot to be able to walk, then walking is the task. We could program the robot to learn to walk, or we could attempt directly to write a program that specifies how to walk manually. The most of contents of this annex is a part of the work [140].

C.2 Machine Learning Tasks

Many kinds of tasks can be solved by machine learning. Some of the most common machine learning tasks include the following:

C.2.1 Classification

In this type of task, the computer program is asked to specify the categories which some input belongs. To solve this task, the learning algorithm is usually asked to produce a function when, the model assigns an input described by a vector to a category identified by numeric class. There are other variants of the classification task, where the outputs have a probability of distribution over classes. An example of a classification task is faces recognition [67], which can be used to an image classification of people and allow computers to interact with their users.

Classification becomes more challenging if the inputs is missing. In order to solve the classification task, the learning algorithm only has to define a single function mapping from a vector input to a categorical output. The some of the inputs may be missing, rather than providing a single classification function, the learning algorithm must learn a set of functions. Each function corresponds to classifying x with a different subset of its inputs missing. This kind of situation arises frequently in medical diagnosis, because many kinds of medical tests are expensive or invasive.

C.2.2 Regression

In this type of task, the algorithm is asked to predict a numerical value given some input. For that, this type of task is similar to classification, except that the format of output is different. An example of a regression task is the prediction of the expected claim amount that an insured person will make, or the prediction of future prices of securities. These kinds of predictions are also used for algorithmic trading.

C.2.3 Clustering

In this type of task, the machine learning system is asked to cluster a subset of data which are similar. Clustering also called unsupervised learning which is the process of dividing a dataset into groups such that the members of each group are as similar (close) as possible to one another, and different groups are as dissimilar (far) as possible from one another. Clustering can uncover previously undetected relationships in a dataset. There are many applications for cluster analysis such as in business and in biology, etc.

C.2.4 Density estimation

In the density estimation problem, the machine learning algorithm is asked to learn a function model, where can be interpreted as a probability density function (continuous) or a probability mass function (discrete) on the space that the examples were drawn. To do this, the algorithm needs to learn the structure of the data. It must knows, where examples cluster tightly and where they are unlikely to occur? Most of the tasks require the learning algorithm to at least implicitly capture the structure of the probability distribution.

The types of listed tasks are only intended to provide examples of most tasks for what machine learning can do, not to define the all types of tasks. Of course, many other types of tasks are possible such as: Transcription, Machine translation, Dimension reduction, Denoising and matching, etc.

C.3 Machine Learning Categories

Machine learning algorithms can be categorized as **unsupervised** or **supervised** depending on the type of learning process. Unsupervised learning algorithms experience a dataset containing many features, then learn the useful properties of the structure of this dataset. But supervised learning algorithms experience a dataset containing features, but each example is also associated with a label or target. The term supervised learning originates from the view of the target y being provided by an instructor or teacher who shows the machine learning system what do. In unsupervised learning, there is no instructor or teacher, and the algorithm must learn to make sense of the data without this guide.

Almost, unsupervised learning involves a observing several examples of a random vector x, and attempting to implicitly or explicitly learn the probability distribution $p(x)$, or some interesting properties of that distribution, while supervised learning involves a observing several examples of a random vector x and an associated value or vector y, and learning to predict y from x, usually by estimating $p(y \mid x)$.

Unsupervised learning and supervised learning are not formally defined terms. The lines between them are often blurred. Many machine learning technologies can be used to perform both tasks. For example, the chain rule of probability states that for a vector $x \in \mathbb{R}^n$, the joint distribution can be decomposed as

$$p(x) = \prod_{i=1}^{n} p(x_i \mid x_1, \cdots, x_{i-1}). \tag{C.1}$$

This decomposition means that we can solve the ostensibly unsupervised problem of modeling $p(x)$ by splitting it into n supervised learning problems. Alternatively, we can solve the supervised learning problem of learning $p(y \mid x)$ by using traditional unsupervised learning technologies to learn the joint distribution $p(y, x)$ and inferring

$$p(y \mid x) = \frac{p(x, y)}{\sum_{y'} p(x, y')}, \tag{C.2}$$

other variants of the learning model are possible. For example, in semi-supervised learning, some examples include a supervision target but others do not. In multi-instance learning, an entire collection of examples is labeled as containing or not containing an example of a class, but the individual members of the collection are not labeled.

C.4 Overfitting and Underfitting

The central challenge in machine learning is that it must performs well on new inputs, not just those on which the model was trained. This ability to perform is called generalization. Typically, when training a machine learning model, it is can compute some error measure on the training set called the training error, and we reduce this training error. What separates machine learning from optimization is that we want the generalization error, also called the test error, to be low as well. The generalization error is defined as the expected value of the error on a new input.

Typically, when use a machine learning algorithm, we do not fix the parameters ahead of time, then sample both datasets. We sample the training set, then use it to choose the parameters to reduce training set error, then sample the test set. Under this process, the expected test error is greater than or equal to the expected value of training error. The factors determining how well a machine learning algorithm will perform are its ability to:

1. Make the training error small.
2. Make the gap between training and test error small.

These two factors correspond to the two central challenges in machine learning: Underfitting and Overfitting. Underfitting occurs when the model is not able to obtain a sufficiently low error value on the training set. Overfitting occurs when the gap between the training error and test error is too large.

We can control whether a model is more likely to overfit or underfit by altering its capacity [140]. Informally, a model's capacity is its ability to fit a wide variety of functions.

Models with low capacity may struggle to fit the training set. Models with high capacity can overfit by memorizing properties of the training set that do not serve them well on the test set.

One way to control the capacity of a learning algorithm is by choosing its hypothesis space, the set of functions that the learning algorithm is allowed to select as being the solution. For example, the linear regression algorithm has the set of all linear functions of its input as its hypothesis space. We can generalize linear regression to include polynomials, rather than just linear functions, in its hypothesis space. Doing so increases the model's capacity.

C.5 Building a Machine Learning Algorithm

The machine learning algorithms can be described as particular instances of a combine a specification of a dataset, a cost function, an optimization procedure and a model. For example, the linear regression algorithm combines a dataset consisting of X and y, the cost function is

$$J(w,b) = -\mathbb{E}_{X,Y-\hat{P}data} log \;\; p_{model}(y \mid x), \tag{C.3}$$

the model specification $p_{model}(y \mid x) = N\left(y, x^T w + b, 1\right)$, and, in most cases, the optimization algorithm defined by solving for where the gradient of the cost is zero using the normal equations.

By realizing that we can replace any of these components mostly independently from the others, we can obtain a very wide variety of algorithms. The cost function typically includes at least one term that causes the learning process to perform statistical estimation. The most common cost function is the negative log-likelihood, so that minimizing the cost function causes maximum likelihood estimation.

The cost function may also include additional terms, such as regularization terms. For example, we can add weight decay to the linear regression cost function to obtain

$$J(w,b) = \lambda \|w\|_2^2 - \mathbb{E}_{X,Y-\hat{P}data} log \;\; p_{model}(y \mid x). \tag{C.4}$$

This still allows closed-form optimization.

If we change the model to be nonlinear, then most cost functions can no longer be optimized in closed form. This requires us to choose an iterative numerical optimization procedure, such as gradient descent.

The recipe for constructing a learning algorithm by combining models, costs, and optimization algorithms supports both supervised and unsupervised learning. The linear regression example shows how to support supervised learning. Unsupervised learning can be supported by defining a dataset that contains only X and providing an appropriate unsupervised cost and model. For example, we can obtain the first PCA vector by specifying that our loss function is

$$J(w) = \mathbb{E}_{X-\hat{P}data} \|x - r(x; w)\|_2^2, \tag{C.5}$$

while our model is defined to have w with norm one and reconstruction function $r(x) = w^T xw$.

In some cases, the cost function may be a function that we cannot actually evaluate, for computational reasons. In these cases, we can still approximately minimize it using iterative numerical optimization so long as we have some way of approximating its gradients.

Most machine learning algorithms make use of this recipe, though it may not immediately be obvious. If a machine learning algorithm seems especially unique or hand-designed, it can usually be understood as using a special-case optimizer. Some models such as decision trees or k-means require special-case optimizers because their cost functions have flat regions that make them inappropriate for minimization by gradient-based optimizers. Recognizing that most machine learning algorithms can be described using this recipe helps to see the different algorithms as part of a taxonomy of methods for doing related tasks that work for similar reasons, rather than as a long list of algorithms that each have separate justifications.

C.6 Machine Learning Classifiers

Machine learning techniques employ an inference principle named induction, in which general conclusions are obtained from a particular set of examples. One of the main approaches for induction is supervised learning [105]. In supervised learning, the knowledge about the problem being modelled is presented by datasets composed of pairs in the form: input, desired output [103]. In machine learning, classification is a supervised learning approach in which the computer program learns from the data input given to it and then uses this learning to classify new observation [69]. A classifier is named model, predictor or hypothesis, will be produced in a process named training. The obtained classifier can be regarded as a function f, which receives an input x and provides an output prediction y [70, 104]. This model also provides a description of the training data. This data set may simply be bi-class or it may be multi-class too [71, 106]. There are different types of classifiers:

1. Linear Classifiers: Logistic Regression, Naive Bayes Classifier.

2. Support Vector Machines SVM.
3. Decision Trees.
4. Boosted Trees.
5. Random Forest.
6. Nearest Neighbor.

Next sections present a brief introduction to the ML classifiers techniques used in this work. Each technique employs a different approach to extract the features from raw data.

C.6.1 Support Vector Machines

Support Vector Machines (SVMs) are based on concepts from the Statistical Learning Theory [107]. The main idea of SVM is to create an optimal hyperplane to classify the data into two classes (positive and negative) and to maximize the distance between the hyperplane separating the two classes and the closet data points to the hyperlane [109]. The optimal hyperplane maximizes the separation margin between the two classes of training data, and is defined by a fraction of the input data instances (called support vectors) close to the hyperplane. SVMs have be a many advantages such as high accuracy and nice theoretical guarantees regarding to overfitting [108]. With an appropriate kernel, they can work well even in non-linear data separable of the base feature space [110]. Especially, the popular problems in classification appear in the case of a very high-dimensional spaces. The distance measurement between the data points in the high-dimensional space is defined by the kernel function [111]. Given a dataset T composed of n pairs (x_i, y_i), in which

$$x_i \in \mathbb{R}^m \quad and \quad y_i \in \{-1, +1\}, \tag{C.6}$$

for a hyperplane, we have

$$w \cdot \Phi(x) + b = 0. \tag{C.7}$$

This last able to separate the data in T with minimum error maximizing the margin of separation between the classes. In this equation, Φ represents a mapping function that maps the data in T to a space of higher dimension, such that the classes become linearly separable.

In SVM training and predictions, the mapping function appears as dot products in the form

$$\Phi(x_i) \cdot \Phi(y_j), \tag{C.8}$$

which can be efficiently computed by Kernel functions, usually simpler than the mapping function. Some of the most using of Kernel functions are the Gaussian or RBF (Radial-Basis Function) functions [72]. SVMs have a good generalization ability. Besides, SVMs also stand out for their robustness to high dimensional data. Their main deficiency concerns to the

difficulty of interpreting the generated model and their sensibility to a proper parameter tuning.

C.6.2 Radial Basis Function

▷ Radial Functions

Radial functions are a special class of functions [72]. Their characteristic feature is that their response decreases (or increases) monotonically with distance from a central point. The centre, the distance scale, and the precise shape of the radial function are parameters of the model, all fixed if it is linear [113].

A typical radial function is the Gaussian which, in the case of a scalar input, is

$$h\left(x\right) = exp\left(-\frac{\left(x-c\right)^2}{r^2}\right). \tag{C.9}$$

Its parameters are its centre c and its radius r. Fig. C.1 illustrates a Gaussian RBF with centre $c = 0$ and radius $r = 1$.

A Gaussian RBF monotonically decreases with distance from the centre. In contrast, a multi-quadric RBF which, in the case of scalar input, is

$$h\left(x\right) = \frac{\sqrt{r^2 + \left(x-c\right)^2}}{r}, \tag{C.10}$$

monotonically increases with distance from the centre (see Fig. C.1) Gaussian-like RBFs are local (give a significant response only in a neighbourhood near the centre) and are more commonly used than multi-quadric type RBFs which have a global response. They are also more biologically plausible because their response is finite.

▷ Radial Basis Function Networks

Radial functions are simply a class of functions. In principle, they could be employed in any sort of model (linear or nonlinear) and any sort of network (single-layer or multi-layer). However, radial basis function networks (RBF networks) have traditionally been associated with radial functions in a single-layer network [112] such as shown in Fig. C.2. An RBF network is nonlinear if the basis functions can move or change size or if there is more than one hidden layer. Some focus on single-layer networks with functions which are fixed in position and size. They use a nonlinear optimisation but only for the regularisation

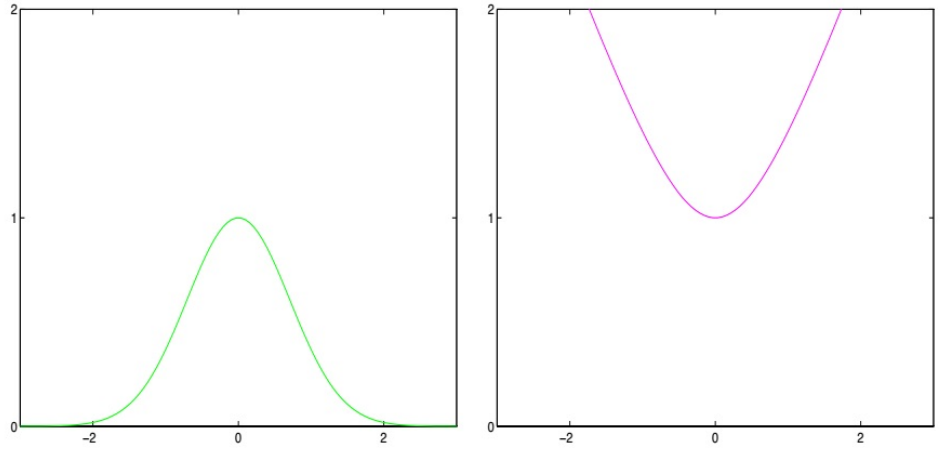


FIGURE C.1: Gaussian (left) and multi-quadric RBF.

parameters and the optimal subset of basis functions [72]. That are employed in explicitly nonlinear networks.

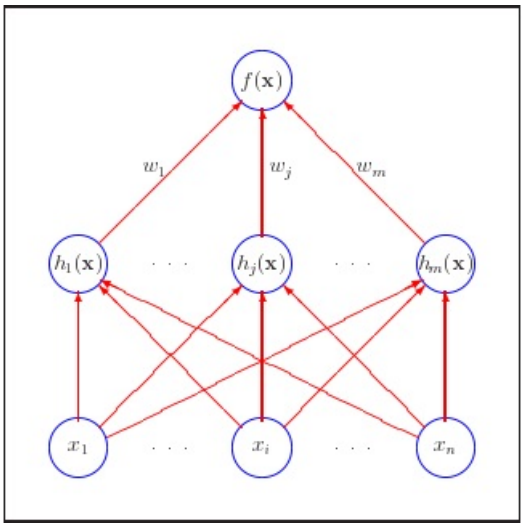


FIGURE C.2: The traditional radial basis function network.

C.6.3 Random Forests Trees

▷ RFT Description

Random Forests Trees (RFT) are combinations of tree predictors [73, 114], which is a flexible, easy machine learning algorithm that produces, even without hyper-parameter tuning. It is also one of the most used algorithms, because its simplicity and the fact that it can be used for both classification and regression tasks. RFT is a supervised learning algorithm

and from its name, it creates a random forest. This forest is an set of Decision Trees trained with the "bagging" method [115]. The general idea of the bagging method is a combination of learning models increases the overall result. In reality, RFT is a collection of Decision Trees but there are some differences. Decision tree is formulate some set of rules, which will be used to make the predictions by a training dataset input. In comparison, the RFT algorithm randomly selects observations and features to build several decision trees and averages the results. Another difference is that decision trees might suffer from overfitting. Random Forest prevents overfitting [108, 116] by creating random subsets of the features and building smaller trees using these subsets, afterwards, it combines the sub-trees.

▷ Advantages and Limitations

An advantages of random forest are that it can be used for both regression and classification tasks [118] and that it's easy to view the relative importance of the input features. RFT is also considered as a very easy to use algorithm, because its default hyper-parameters produce a good prediction. One of the big problems in machine learning is overfitting, but a random forest classifier won't overfit the model because there are enough trees in the forest [119, 120].

The main limitation of RFT is that a large number of trees can make the algorithm to slow and ineffective for real-time predictions. In general, these algorithms are fast to train, but quite slow to create predictions once they are trained. A more accurate prediction requires more trees, which results a slower model. In most real-world applications, the random forest algorithm is fast enough, but there can certainly be a situations where run-time performance is important and other approaches would be preferred. And of course, RFT is a predictive modeling tool and not a descriptive tool [121]. That means, if you are looking for a description of the relationships in your data, other approaches would be preferred.

▷ RFT Algorithm

RFT classifier is one of the big accurate learning algorithms as it produces a best classifiers for many data sets and it runs efficiently on big databases also it gives estimates of what variables are important in the classification [117]. *RFT* is composed of some number of decision trees. Each tree is built as follows:

1. Let the number of training objects be n_{tree}, and the number of features in vector be F.
2. Training set for each tree is built by choosing n times with replacement from all available training objects.
3. Number $f << F$ is an amount of features on which are randomly chosen for each node to do the decision.

4. Each tree is built to the largest extent possible.

5. Each tree gives a classification, which is called voting for that class. The forest chooses the class having the most votes.

The random forests tree algorithm (for both classification and regression) is as follows:

Algorithm : The random forests tree.

---

1. For $b = 1$ to B :

(a) Draw a bootstrap sample Z^* of size N from training data.

(a) Grow a RFT T_b to the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size n_{min} is reached.

i. Select m variable at random from the P variables.

ii. Pich the best variable/split-point among the m.

iii. Split the node into two daughter nodes.

2. Output the ensemble of trees $\{T_b\}_1^B$.

To make a prediction at a new point x:

Regression: $f_{rf}^B(x) = \dfrac{1}{B}\sum_{b=1}^{B} T_b(x)$.

Classification: Let $C_b(x)$ be the class prediction of the b^{th} RFT, then:

$C_b(x) = vote \;\; majority\{C_b(x)\}_1^B$.

Appendix D

DEEP LEARNING AND MODERN PRACTICE

D.1 Introduction

DEEP learning is a specific kind of machine learning, that a person must has a solid overview on the basis of machine learning to understand the deep learning. For that, this chapter provides a brief course in the most important principles that will be applied throughout the deep learning applications.

D.2 Deep Learning Networks

D.2.1 Definitions

Deep learning has various closely related definitions or high-level descriptions [142]:

"Deep Learning is a new area of Machine Learning research, which has been introduced as a sub-field within machine learning that is based on algorithms for learning multiple levels of representation in order to model complex relationships among data. Higher-level features and concepts are thus defined in terms of lower-level ones, and the same lower-level concepts can help to define many higher-level concepts. Such a hierarchy of features is called a deep architecture. Most of these models are based on unsupervised learning of representations."

We can provide another definition of deep learning which is:

"Deep learning is a set of algorithms in machine learning that attempt to learn in multiple levels, corresponding to different levels of abstraction. It typically uses artificial neural networks. The levels in these learned statistical models correspond to distinct levels of concepts, where higher-level concepts are defined from lower-level ones, and the same lower-level concepts can help to define many higher-level concepts. Deep learning is part of a broader family of machine learning methods based on learning representations."

D.2.2 Classes of Deep Learning

Deep learning refers to a rather wide class of machine learning techniques and architectures, with the characteristic to use many layers of non-linear information processing that are hierarchical in nature. Depending on how the architectures and techniques are intended for use, it can broadly categorize this deep learning networks area into three major classes [142]:

- **Deep networks for unsupervised or generative learning**, which are intended to capture high-order correlation of the observed or visible data for pattern analysis or synthesis purposes when no information about target class labels is available. Unsupervised feature or representation learning in the literature refers to this category of the deep networks. In the generative mode, may also be intended to characterize joint statistical distributions of the visible data and their associated classes when available and being treated as part of the visible data. In the latter case, the use of Bayes rule can turn this type of generative networks into a discriminative one for learning.

- **Deep networks for supervised learning**, which are intended to directly provide discriminative power for pattern classification purposes, often by characterizing the posterior distributions of classes conditioned on the visible data. Target label data are always available in direct or indirect forms for such supervised learning. They are also called discriminative deep networks.

- **Hybrid deep networks**, where the goal is discrimination which is assisted, often in a significant way, with the outcomes of generative or unsupervised deep networks. This can be accomplished by better optimization and regularization of the deep networks. The goal can also be accomplished when discriminative criteria for supervised learning are used to estimate the parameters in any of the deep generative or unsupervised deep networks.

D.2.3 Basic Deep Learning Terminologies.

The family of deep learning methods have been growing increasingly richer, encompassing those of neural networks, hierarchical probabilistic models, and a variety of unsupervised and supervised feature learning algorithms. Below, we review representative work in each of the above three categories, where several basic definitions are summarized [142]:

- **Deep Belief Network (DBN):** probabilistic generative models composed of multiple layers of stochastic, hidden variables. The top two layers have undirected, symmetric connections between them. The lower layers receive top-down, directed connections from the layer above.
- **Boltzmann Machine (BM):** a network of symmetrically connected, neuron-like units that make stochastic decisions about whether to be on or off.
- **Restricted Boltzmann Machine (RBM):** a special type of BM consisting a layer of visible units and a layer of hidden units with no visible-visible or hidden-hidden connections.
- **Deep Neural Network (DNN):** a multilayer perceptron with many hidden layers, whose weights are fully connected and are often initialized by using either an unsupervised or a supervised training technique.
- **Deep Autoencoder:** a discriminative DNN whose output targets are the data input itself rather than class labels; hence an unsupervised learning model. When trained with a denoising criterion, a deep autoencoder is also a generative model.
- **Distributed Representation:** an internal representation of the observed data in such a way that they are modeled as being explained by the interactions of many hidden factors. A particular factor learned from configurations of other factors can often generalize well to new configurations. Distributed representations naturally occur in a connection neural network, where a concept is represented by a pattern of activity across a number of units and where at the same time a unit typically contributes too many concepts. One key advantage of such many-to-many correspondence is that they provide robustness in representing the internal structure of the data in terms of graceful degradation and damage resistance. Another key advantage is that they facilitate generalizations of concepts and relations.

D.3 Challenges Motivating Deep Learning [140]

The machine learning algorithms described in previous annex work very well on a wide variety of important tasks. However, they have not succeeded to solve the central problems in artificial intelligence (AI). The development of deep learning was motivated by the failure of traditional algorithms to generalize well on such AI tasks.

This section is the challenges of traditional machine learning to generalizing at new examples more difficult when working with high-dimensional data, and how the mechanisms used are insufficient to learn complicated functions in high-dimensional spaces, also often impose high computational costs. Deep learning was designed to overcome these and other obstacles.

D.3.1 The Curse of Dimensionality

Many machine learning problems become exceedingly difficult when the number of dimensions in the data is high. This phenomenon is known as "the curse of dimensionality". Of particular concern is that the number of possible distinct configurations of a set of variables increases exponentially as the number of variables increases.

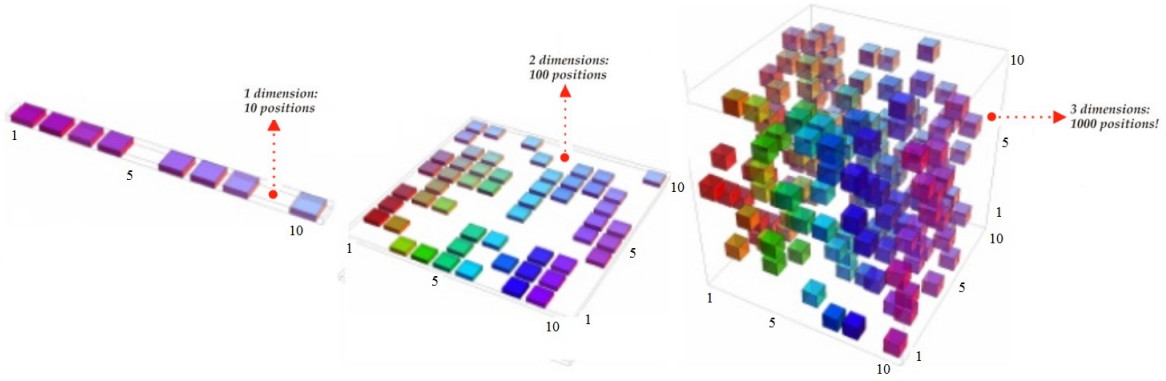


FIGURE D.1: The number of relevant dimensions of the data increases (from left to right).

Fig. D.1: As the number of relevant dimensions of the data increases (from left to right), the number of configurations of interest may grow exponentially. In this one-dimensional example (Left), we have one variable for which we only care to distinguish 10 regions of interest. With enough examples falling within each of these regions (each region corresponds to a cell in the illustration), learning algorithms can easily generalize correctly. A straightforward way to generalize is to estimate the value of the target function within each region

(and possibly interpolate between neighboring regions). With 2 dimensions (Center), it is more difficult to distinguish 10 different values of each variable. We need to keep track of up to $10 \times 10 = 100$ regions, and we need at least that many examples to cover all those regions. With 3 dimensions (Right), this grows to $10^3 = 1000$ regions and at least that many examples. For d dimensions and v values to be distinguished along each axis, we seem to need $O\left(v^d\right)$ regions and examples.

One challenge posed by the curse of dimensionality is a statistical challenge. As illustrated in Fig. D.1, a statistical challenge arises because the number of possible configurations of x is much larger than the number of training examples. To understand the issue, we consider that the input space is organized into a grid, like in the figure. We can describe low-dimensional space with a low number of grid cells that are occupied by the data.

To generalizing a new data point, we can usually tell what to do simply by inspecting the training examples that lie in the same cell as the new input. For example, if estimating the probability density at some point x, we can just return the number of training examples in the same unit volume cell as x, divided by the total number of training examples. If we wish to classify an example, we can return the most common class of training examples in the same cell. If we are doing regression we can average the target values observed over the examples in that cell. In high-dimensional spaces the number of configurations is huge, much larger than our number of examples, a typical grid cell has no training example associated with it. Many traditional machine learning algorithms simply assume that the output at a new point should be approximately the same as the output at the nearest training point.

D.3.2 Manifold Learning

An important concept underlying many ideas in machine learning is a manifold. A manifold is a connected region. Mathematically, it is a set of points associated with a neighborhood around each point. From any given point, the manifold locally appears to be a Euclidean space. In everyday life, we experience the surface of the world as a 2-D plane, but in fact, it is a spherical manifold in 3-D space.

The definition of a neighborhood surrounding each point implies the existence of transformations that can be applied to move on the manifold from one position to a neighboring one. Although there is a formal mathematical meaning to the term "manifold", in machine learning it tends to be used more loosely to designate a connected set of points that can be approximated well by considering only a small number of degrees of freedom, or dimensions. Each dimension corresponds to a local direction of variation. See Fig. D.2 for an example of

training data lying near a one-dimensional manifold embedded in two dimensional space. In the context of machine learning, we allow the dimensionality of the manifold to vary from one point to another. This often happens when a manifold intersects itself.

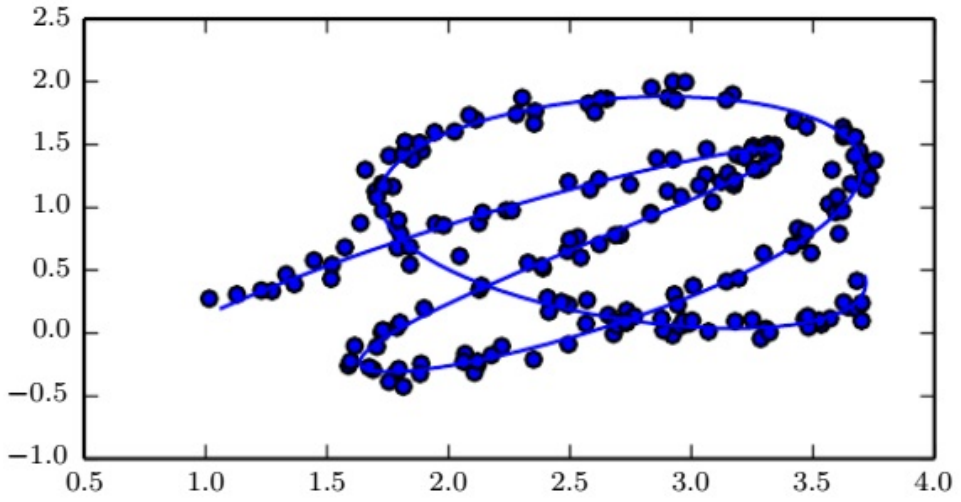


FIGURE D.2: Data sampled from a distribution in a two-dimensional space that is actually concentrated near a one-dimensional manifold, like a twisted string. The solid line indicates the underlying manifold that the learner should infer [140].

Many machine learning problems seem hopeless if we expect the machine learning algorithm to learn functions with interesting variations across all of $\mathbb{R}^n$. Manifold learning algorithms surmount this obstacle by assuming that most of $\mathbb{R}^n$ consists of invalid inputs, and that interesting inputs occur only along a collection of manifolds containing a small subset of points, with interesting variations in the output of the learned function occurring only along directions that lie on the manifold, or with interesting variations happening only when we move from one manifold to another.

Manifold learning was introduced in the case of continuous-valued data and the unsupervised learning setting, although this probability concentration idea can be generalized to both discrete data and the supervised learning setting: the key assumption remains that probability mass is highly concentrated. The assumption that the data lies along a low-dimensional manifold may not always be correct or useful. In the context of AI tasks, such as those that involve processing images, sounds, or text, the manifold assumption is at least approximately correct. The evidence in favor of this assumption consists of two categories of observations:

1. The first observation in favor of the manifold hypothesis is that the probability distribution over images, text strings, and sounds that occur in real life is highly concentrated.

Uniform noise essentially never resembles structured inputs. Concentrated probability distributions are not sufficient to show that the data lies on a reasonably small number of manifolds. We must also establish that the examples we encounter are connected to each other by other examples, with each example surrounded by other highly similar examples that may be reached by applying transformations to traverse the manifold.

2. The second observation in favor of the manifold hypothesis is that we can imagine such neighborhoods and transformations. In the case of images, we can certainly think of many possible transformations that allow us to trace out a manifold in image space: we can gradually dim or brighten the lights, gradually move or rotate objects in the image, gradually alter the colors on the surfaces of objects, etc.

When the data lies on a low-dimensional manifold, it can be most natural for machine learning algorithms to represent the data in terms of coordinates on the manifold, rather than in terms of coordinates in $\mathbb{R}^n$. In everyday life, we can think of roads as 1-D manifolds embedded in 3-D space. We give directions to specific addresses in terms of address numbers along these 1-D roads, not in terms of coordinates in 3-D space. Extracting these manifold coordinates is challenging, but holds the promise to improve many machine learning algorithms. This general principle is applied in many contexts. Fig. D.3 shows the manifold structure of a dataset consisting of faces.

FIGURE D.3: Training examples from the Multiview Face Dataset for which the subjects were asked to move in such a way as to cover the two-dimensional manifold corresponding to two angles of rotation [140].

D.3.3 Local Constancy and Smoothness Regularization

In order to generalize, machine learning algorithms need to be guided by prior beliefs about what kind of function they should learn. Among the most widely used of these implicit "priors" is the smoothness or local constancy prior. This prior states that the function we learn should not change very much within a small region.

Many simpler algorithms rely exclusively on this prior to generalize, and as a result they fail to scale to the statistical challenges involved in solving AI tasks. Throughout this, deep learning introduces additional (explicit and implicit) priors in order to reduce the generalization error on sophisticated tasks. There are many different ways to implicitly or explicitly express a prior belief that the learned function should be smooth or locally constant. All of these different methods are designed to encourage the learning process to learn a function f^* that satisfies the condition

$$f^*(x) \approx f^*(x + \epsilon), \tag{D.1}$$

for most configurations x and small change ϵ. In other words, if we know a good answer for an input x (for example, if x is a labeled training example) then that answer is probably good in the neighborhood of x. If we have several good answers in some neighborhood we would combine them (by some form of averaging or interpolation) to produce an answer that agrees with as many of them as much as possible.

An example of the local constancy approach is the k-nearest neighbors family of learning algorithms. These predictors are literally constant over each region containing all the points x that have the same set of k nearest neighbors in the training set. For $k = 1$, the number of distinguishable regions cannot be more than the number of training examples. While the k-nearest neighbors algorithm copies the output from nearby training examples, most kernel machines interpolate between training set outputs associated with nearby training examples. An important class of kernels is the family of local kernels where $k(u, v)$ is large when $u = v$ and decreases as u and v grow farther apart from each other. A local kernel can be thought of as a similarity function that performs template matching, by measuring how closely a test example x resembles each training example x^i. Much of the modern motivation for deep learning is derived from studying the limitations of local template matching and how deep models are able to succeed in cases where local template matching fails [141].

The smoothness assumption and the associated non-parametric learning algorithms work extremely well so long as there are enough examples for the learning algorithm to observe high points on most peaks and low points on most valleys of the true underlying function to be learned. This is generally true when the function to be learned is smooth enough and

varies in few enough dimensions. In high dimensions, even a very smooth function can change smoothly but in a different way along each dimension. If the function additionally behaves differently in different regions, it can become extremely complicated to describe with a set of training examples.

▷ AI tasks have structure that is much too complex to be limited of simple specified properties such as periodicity. So, we want learning algorithms that embody more general purpose assumptions. Many different deep learning algorithms provide implicit or explicit assumptions that are reasonable for a broad range of AI tasks in order to capture the advantages. The core idea in deep learning is that we assume that the data was generated by the composition of factors or features, potentially at multiple levels in a hierarchy. Many other similarly generic assumptions can further improve deep learning algorithms. These assumptions allow a big gain in the relationship between the number of examples and the number of regions that can be distinguished. The exponential advantages conferred by the use of deep learning representations overcome the exponential challenges posed such as: the curse of dimensionality, manifold and local constancy, etc.

D.4 Structure of Deep Network

Human brain is a very powerful machine where we see multiple images every second and process them without realizing how the processing is done, but that is not the case with machines. A human's brain neuronal activity is incredibly complex and simulating it at a 100% ratio is impossible with current technology. Achieving just a 10 % simulation rate, it was impossible that the supercomputers run this limited simulations in the past. This is because, the act of neurons network (crucial for every activity that happens in the brain) requires more power than today's hardware (show Fig. D.4).

But, these recent developments and the increased capacity of processing units, as well as the recent advances in machine learning, signal processing and the increasing volume of information used in training have allowed the spread of a deep learning algorithms which are basically inspired by the brain. The network seen in Fig. D.5, is a neural network made of interconnected neurons.

D.5 Convolutional Neural Networks (CNNs)

A convolutional neural networks (CNNs) are a class of deep learning, most of this networks commonly applied to analyzing visual imagery. CNNs are very similar to ordinary neural

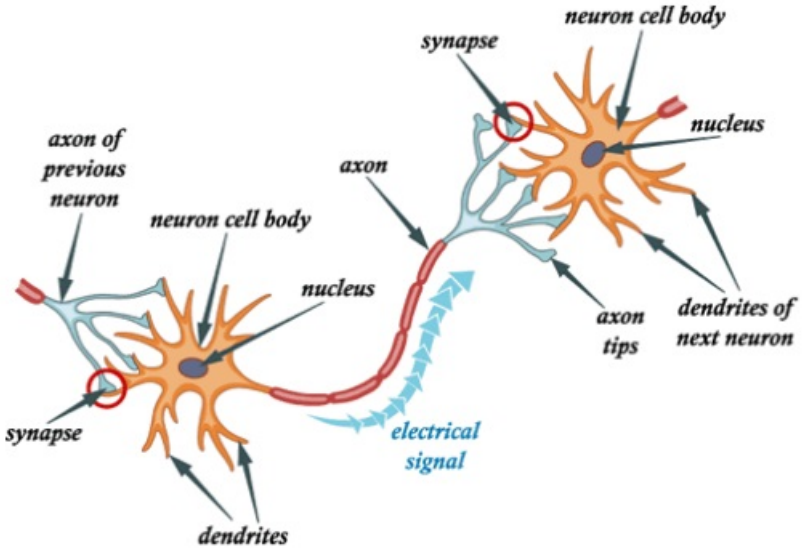


FIGURE D.4: The neurons in a human brain's.

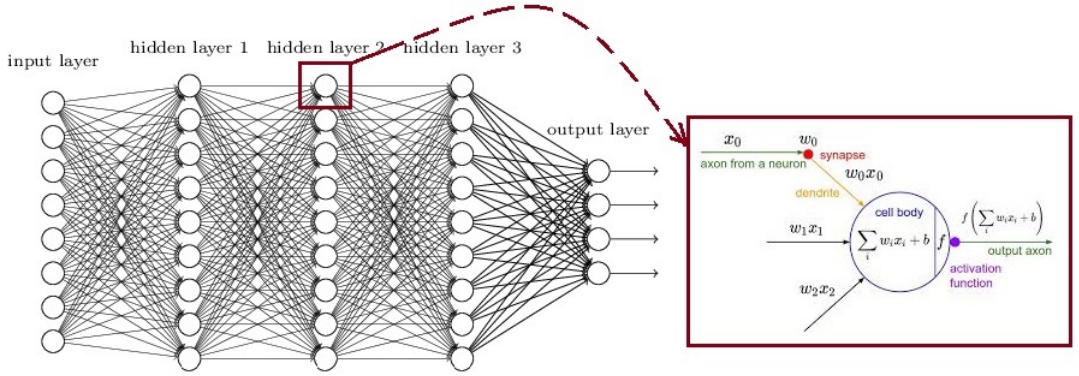


FIGURE D.5: A example of deep neural network.

networks, they are made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity [64]. Convolutional Neural Networks (CNNs) are a specific form of neural networks that explicitly assume the inputs to the network be structured samples, such as audio signals or image pixels which can be filtered [84]. These architectures typically focus on the solutions for computer vision applications, like classification, localization and segmentation of images and videos.

D.5.1 Architecture of CNNs

Neural Networks receive an input, and transform it through a series of hidden layers. Each hidden layer is made up of a set of neurons, where each neuron is fully connected to

all neurons in the previous layer, and where neurons in a single layer function completely independently and do not share any connections. The last fully-connected layer is called the "output layer" and in classification settings it represents the class scores.

In image processing, every image is an arrangement of dots (a pixel) arranged in a special order. If you change the order or the color of a pixel, the image would change as well. The machine will basically break this image into a matrix of pixels and store the color code for each pixel at the representative location. Supervised learning based deep image recognition CNN architectures are composed of multiple convolutional stages stacked [84] on top of each other to learn hierarchical visual features as captured in Fig. D.6. Regularization approaches such as stochastic pooling, dropout, data augmentation have been used to enhance the recognition accuracy [78]. Recently, the faster convergence of these architectures is attributed to the inclusion of Rectified Linear Units (ReLU) nonlinearity into each of the layer with weights.

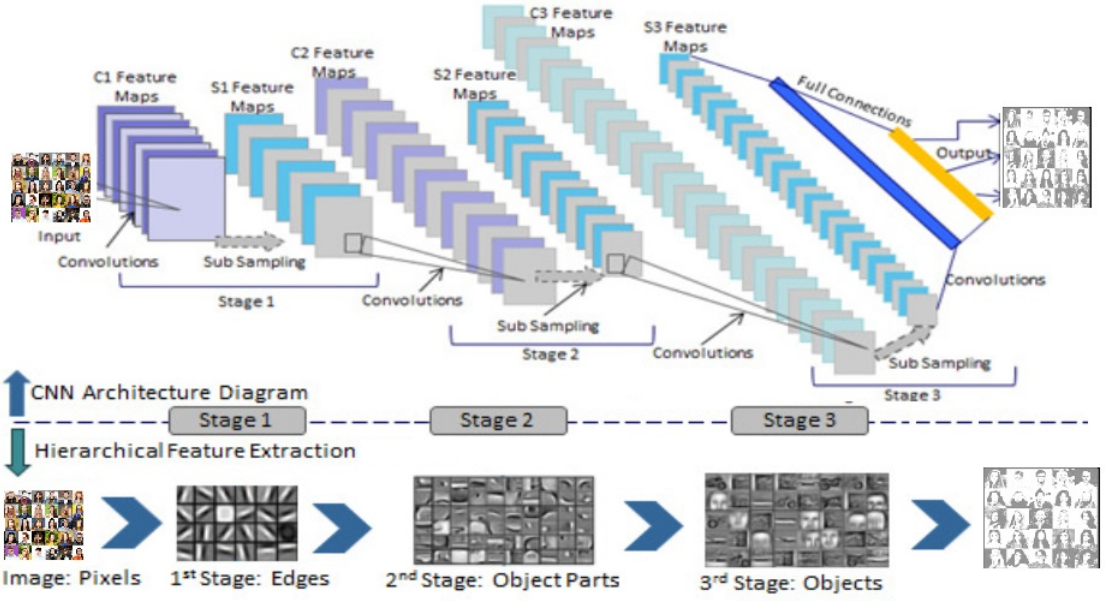


FIGURE D.6: Learning hierarchy of image features in CNN architecture.

D.5.2 The Convolution Layer

To prevent the networks from having too many parameters, the fully-connected layers are replaced by convolutional layers in a neural networks, leading to CNN models. In convolutional layers (CONV), the hidden neurons are replaced with convolutional filters. Instead of solving of neuron weights, we solve with a family of filters, each filter having its own weights.

The convolutional layers arrange the neurons in a 3D fashion using the height, width and depth for the signal being processed. In the depth dimension, the CONV layer is analogous to a filtered signal used for digital image processing, where each filtered signal came from a learned filter, whose weights shall be learned during the training process. Fig. D.7 shows a fully-connected conventional neural networks [143].

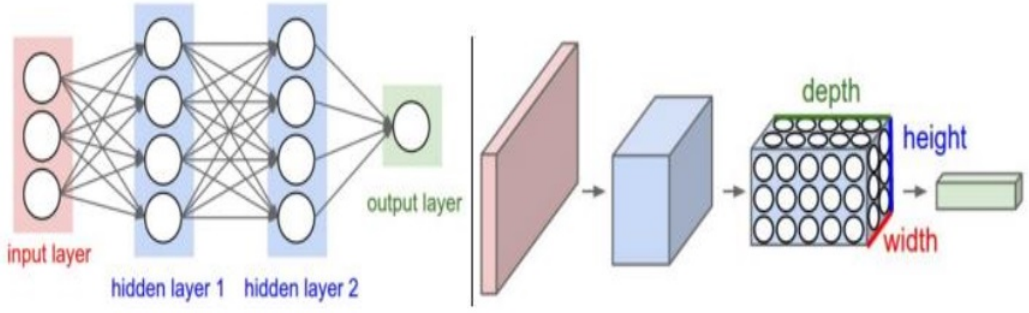


FIGURE D.7: A conventional neural networks [143].

• **The Concept of Stride and Padding**

The filter or the weight matrix, was moving across the entire image moving one pixel at a time. We can define it like a hyperparameter, as to how we would want the weight matrix to move across the image. If the weight matrix moves 1 pixel at a time, we call it as a stride of 1 (see Fig. D.8). Let's see how a stride of 2 would look like [143].

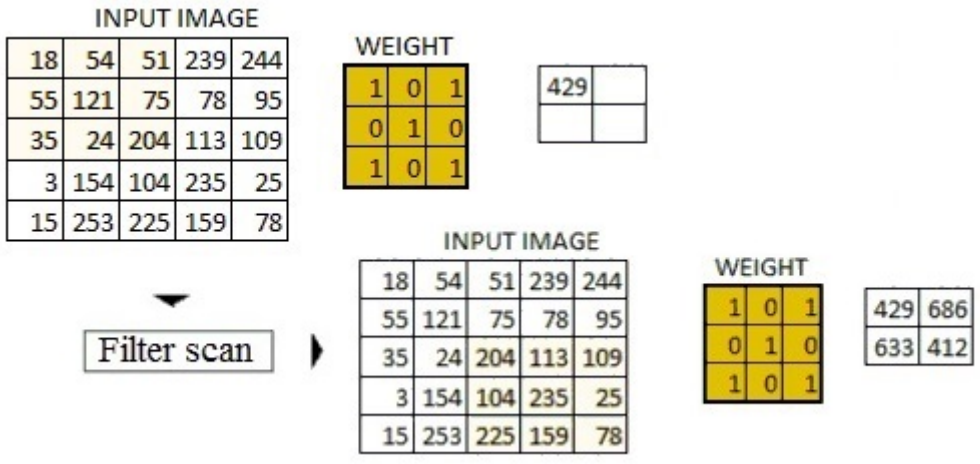


FIGURE D.8: Example of stride and padding with out zero..

Padding with a greater value of zero is a helpful method to preserve the information on the borders of the image from vanishing through multiple convolutions. It also preserves the spatial dimensions of the output from the convolutional layers, often called the output volume. Padding the input image with zeros across it solves this problem for us. We can also add more than one layer of zeros around the image in case of higher stride values (see Fig. D.9). As you can see that the size of image keeps on reducing as we increase the stride value [143].

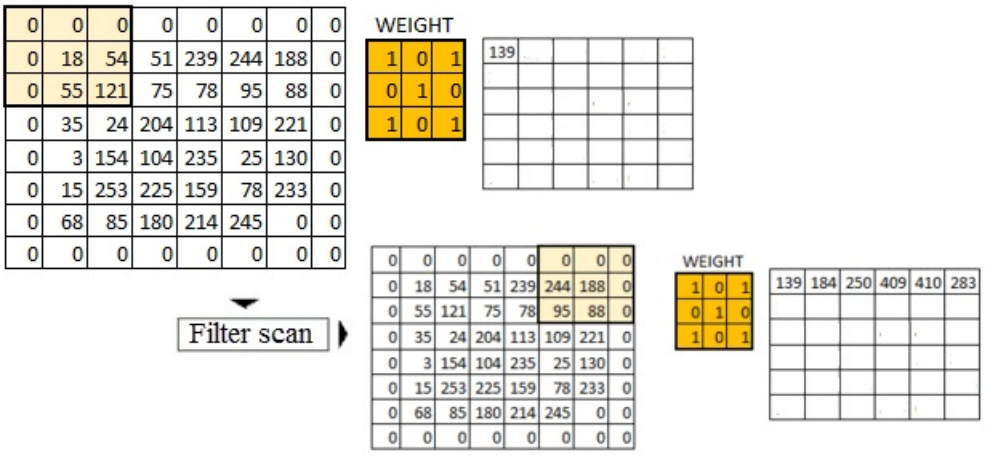


FIGURE D.9: Example of stride and padding with zero.

- **Multiple Filters and Activation Map**

One thing to keep in mind is that the depth dimension of the weight would be a same as the depth dimension of the input image. The weight extends to the entire depth of the input image. Therefore, convolution with a single weight matrix would result into a convolved output with a single depth dimension. In most cases, instead of a single filter (weight matrix), we have multiple filters of the same dimensions applied together [143].

The output from the each filter is stacked together forming the depth dimension of the convolved image. Suppose, we have an input image of size $32 \times 32 \times 3$ and we apply 10 filters of size $5 \times 5 \times 3$ with valid padding. The output would have the dimensions as $28 \times 28 \times 10$ (see Fig. D.10).

D.5.3 The Pooling Layer

Sometimes when the images are too large, we would need to reduce the number of trainable parameters. It is then desired to periodically introduce pooling layers between subsequent

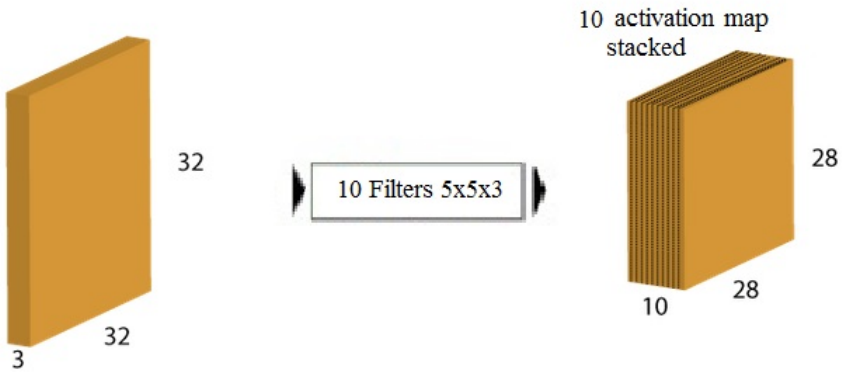


FIGURE D.10: Example of activation map stacked.

convolution layers. Pooling is done for the sole purpose of reducing the spatial size of the image. Pooling is done independently on each depth dimension, therefore the depth of the image remains unchanged. The most common form of pooling layer generally applied is the max pooling (see Fig. D.11). Similarly, other forms of pooling can also be applied like average pooling or the L2 norm pooling [143].

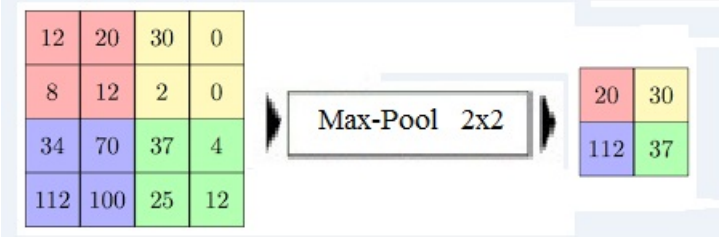


FIGURE D.11: Example of max pooling.

For output dimensions, It might be getting a little confusing to understand the input and output dimensions at the end of each convolution layer. Three hyper-parameter would control the size of output volume [143].

- The number of filters: the depth of the output volume will be equal to the number of filter applied. it had stacked the output from each filter to form an activation map. The depth of the activation map will be equal to the number of filters.
- Stride: when it has a stride of one, we move across and down a single pixel. With higher stride values, we move a large number of pixels at a time and hence produce smaller output volumes.

- Zero padding: this helps us to preserve the size of the input image. If a single zero padding is added, a single stride filter movement would retain the size of the original image.

We can apply a simple formula to calculate the output dimensions. The spatial size of the output image can be calculated as: $\frac{(w - f + 2p)}{s} + 1$. Here, W is the input volume size, F is the size of the filter, P is the number of padding applied and S is the number of strides.

D.5.4 The Output layer

After multiple layers of convolution and padding, it would need the output in the form of a class. The convolution and pooling layers would only be able to extract features and reduce the number of parameters from the original images. However, to generate the final output, it needs to apply a fully connected layer to generate an output equal to the number of classes we need [143]. It becomes tough to reach that number with just the convolution layers. Convolution layers generate activation maps that the output need these maps for classification of images. The output layer has a loss function like categorical cross-entropy to compute the error in prediction. Once the forward pass is complete the back propagation begins to update the weight and biases for error and loss reduction.

Appendix E

PERSONAL CONTRIBUTIONS

A.1 Publications

1. Evolving Systems Journal. "Deep learning for finger-knuckle-print identification system based on PCANet and SVM classifer", ISSN 1868-6478, April 2018.

A.2 Conferences

1. The 2nd International Conference on Artificial Intelligence and Information Technology, University of Ouargla, 4-6 March 2019. ICA2IT 2019 "Can Handwriting Style Help Strengthen the Person Identity?".
2. The 10th Conference in Electrical Engineering, Ecole Militaire Polytechnique Bordj El Bahri, 17 - 18 April 2017. CGE'10 2017 "Hyperspectral Vs Multispectral Palmprint Recognition Using Random Forest Tree".
3. The 2nd International Conference on Information Technology for Organizations Development, University of USMBA, Fez, Morocco, March 30- April 1st, 2016. IT4OD'16 "Visible spectrum bands of palmprint image for a robust biometric identification system".
4. The International Conference in Electrical Engineering, University of Bechar, 17-19 Nov 2013. CIGE'2013 "Fusion of Palm-Vein and Finger-Vein for Personal Identification Using Principal Component Analysis".

Bibliography

[1] T. Best, "Expanding the Fusion Concept ", *Biometric Technology Today*, Vol. 14, Issue 6, pp. 7-8, June 2006.

[2] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics : A Tool for Information Security ", *IEEE Transactions on Information Forensics and Security*, Vol. 1, No. 2, pp. 125-143, 2006.

[3] A. K. Jain, A. Ross, and S. Prabhakar. "An Introduction to Biometric Recognition ", *IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics*, Vol. 14(1), pp. 4-20, 2004.

[4] A. Ross, K. Nandakumar, A. K. Jain. "Handbook of Multibiometrics ", *Springer*, 2006.

[5] N. Belgacem, R. Fournier, A. Nait-Ali, F. Bereksi-Reguig. "A novel biometric authentication approach using ECG and EMG signals ", *Jornal Med. Eng. Tech.* Vol. 39(4), pp. 226-238, 2015.

[6] R. Upadhayay, R. K. Yadav. "Kernel principle component analysis in face recognition system: a survey ", *International J Adv Res Comput Sci Softw Eng*, Vol. 3(6), pp. 2348-2353, 2013.

[7] Z. Bochang. "Local derivative pattern versus local binary pattern:face recognition with high-order local pattern descriptor ", *IEEE Transactions Image Process*, Vol. 19(2), pp. 533-544, 2010.

[8] A. Kumar, Ch. Ravikanth. "Personal Authentication Using Finger Knuckle Surface ", *IEEE Transactions On Information Forensics and Security*, Vol. 4, No. 1, pp. 98-110, 2009.

[9] G. Koltzsch. "Biometrics-Market Segments and Applications ", *Journal of Business Economics and Management*, Vol. 8(2), pp. 119-122, 2007.

[10] T. Dunstone, N. Yager. "Biometric system and data analysis: Design, evaluation, and data mining ", *Springer*, 2006.

[11] Biometrics History. NSTC. Home page, http://www.biometrics.gov/Documents/BioHistory.pdf, 2011.

[12] History of Fingerprinting. FINGERPRINTING. Home page, http://www.fingerprinting.com/history-of-fingerprinting.php, 2011.

[13] J. Meaney. "History of Fingerprints Timeline ", home page, http://www.fingerprintamerica.com/fingerprinthistory.asp, 2011.

[14] H. T. F. Rhodes, Alphonse Bertillon. "Father of Scientific Detection ", AbelardSchuman, New York. 1956.

[15] A. K. Jain, R. Bolle, S. Pankanti. "Biometrics, Personal Identification in Networked Society: Personal Identification in Networked Society ", *Kluwer Academic Publishers*, Norwell, MA, USA, 1998.

[16] D. Maltoni, D. Maio, A. K. Jain, S. Prabhakar. "Handbook of Fingerprint Recognition ", *Springer-Verlag*, 2003.

[17] A. Kong, D. Zhang, G. Lu. "A study of identical twins' palmprints for personal verification ", *Pattern Recognition*, Vol. 39, pp. 2149-2156, 2006.

[18] F. Prokoski. "Disguise detection and identification using infrared imagery ", *Proceedings of SPIE, Optics, and Images in Law Enforcement II*, pp. 27-31, 1982.

[19] V. Nalwa. "Automatic on-line signature verification ", *Proceedings of the IEEE*, Vol. 85(2), pp. 215-239, 1997.

[20] J. Campbell. "Speaker recognition: a tutorial ", *Proceedings of the IEEE*, Vol. 85(9), pp. 1437-1462, 1997.

[21] F. Monrose, A. Rubin. "Authentication via keystroke dynamics ", *In Proceedings of the 4th ACM conference on Computer and communications security*, pp. 48-56, 1997.

[22] A. K. Jain, P. Flynn, A. Ross. "Handbook of biometrics ", *Springer*, 2007.

[23] R. M.Bolle, J. H. Connell, S. Pankanti, N. K. Ratha, A. W.Senior. "Guide to biometrics ", *Springer*, 2004.

[24] J. Egan. "Signal detection theory and ROC analysis ", *Academic Press*, 1995.

[25] H. Moon, P. J. Phillips. "Computational and performance aspects of PCA-based face recognition algorithms ", *Perception*, Vol. 30(5), pp. 303-321, 2001.

[26] A. Meraoumia, S. Chitroub, A. Bouridane. "Multimodal Biometric Person Recognition System based onFingerprint & Finger-Knuckle-Print Using Correlation Filter Classifier ", *IEEE International Conference On Communications-ICC10*, Canada, 2012.

[27] NIST Report to the United States Congress. "Summary of NIST Standards for Biometric Accuracy, Tamper Resistance, and Interoperability ", Available at: http://sequoyah.nist.gov/pub/nist.internal.reports, 2011.

[28] Y. Chen, S. C. Dass, A. K. Jain. "Fingerprint Quality Indices for Predicting Authentication Performance ", *Proceedings of the Fifth International Conference on Audio and Video-Based Person Authentication, (AVBPA'05)*, , pp. 160-170, 2005.

[29] S. Prabhakar, S. Pankanti, A. Jain, "Biometric Recognition: Security and Privacy Concerns ", *IEEE Security and Privacy*, Vol. 1(2), pp. 33-42, 2003.

[30] D. W. Aha, D. Kibler, M. K. Albert. "Instance based learning algorithms ", *Machine Learning*, Vol. 6, pp. 37-66, 1991.

[31] H. Kang, B. Lee, H. Kim, D. Shin, J. Kim. "A study on performance evaluation of the liveness detection for various fingerprint sensor modules ", *Proceedings of KES*, pp. 1245-1253, 2003.

[32] L. Thalheim, J. Krissler, P. Ziegler. "Body Check: Biometrics Defeated ", *c't Magazine article*, 2002. English translation is available at: http://www.extremetech.com/article2/0,2845,13919,00.asp.

[33] E. Bigun, J. Bigun, S. Fisher. "Expert conciliation for multimodal person authentication systems using baysian statistics ", *Proceedings of the International Conference on Audio and Video-Based Biometric Person Authentication*, Vol. 12(6), pp. 291-300, 1997.

[34] R. Brunelli, D. Falavigna. "Person identification using multiple cues ", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17(10), pp. 955-966, 1995.

[35] J. Kittler, M. Hatef, R. Duin, J. Matas. "On combining classifiers ", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*,Vol. 20(3), pp. 226-239, 1998.

[36] A. Ross, A. K. Jain. "Multimodal biometrics: An overview ", *In proc. of 12th European Signal Processing Conference (EUSIPCO)*, pp.1221-1224, 2004.

[37] R. Frischholz, U. Dieckmann. "BioID: A multimodal biometric identification system ", *IEEE Computer*, Vol. 33(2), pp. 64-68, 2000.

[38] K. Nandakumar, Y. Chen, S. C. Dass, A. K. Jain. "Likelihood ratio-based biometric score fusion ", *IEEE Transactions on Pattern Analysais andMachine Intelligence*, Vol. 30(2), pp. 342-347, 2009.

[39] R. Singh, M. Vatsa, A. Noore, "Integrated Multilevel Image Fusion and Match Score Fusion of Visible and Infrared Face Images for Robust Face Recognition ", *Pattern Recognition*, Vol. 41(3), pp. 880-893, 2008.

[40] J. Kludas, E. Brunio, S.Marchand-Maillet. "Information fusion in multimedia information retrieval ", *In proc. of 5th International Workshop on daptive Multimedia Retrieval: Retrieval, User, and Semantics*, pp. 147-159, 2008.

[41] J. Linas, C. Bowman, G. Rogova, A. Steinberg, E. Waltz, F. White. "Revisiting the JDL data fusion model II ", *In proc. of 7th International Conference on Information Fusion*, Stockholm, Sweden, 2004.

[42] C. Sanderson, K. K. Paliwal. "Information fusion for robust speaker verification ", *In proc. of Seventh European Conference on Speech Communication and Technology*, pp. 755-758, 2001.

[43] D. Zhang. "Automated Biometrics Technologies and Systems ", *Originally published by Kluwer Academic Publishers*, 2000.

[44] X.Liu, T.Chen. "Geometry-assisted statistical modeling for face mosaicing ", *In proc. of IEEE International Conference on image Processing (ICIP)*, Vol. 2, pp. 883-886, 2003.

[45] A. Ross and A. K. Jain. "Multimodal biometrics: An overview ", *in Proc. 12th Eur. Signal Process. Conf.*, pp.1221-1224, 2004.

[46] A. Ross and A. K. Jain. "Information fusion in biometrics ", *Pattern Recognition Letters*, Vol. 24, pp. 2115-2125, 2003.

[47] R. Snelick, M. Indovina, J. Yen, A. Mink. "Multimodal biometrics: Issues in design and testing ", *In Proceedings of Fifth International Conference on Multimodal Interfaces*, pp. 68-72, 2003.

[48] Combining multiple biometrics, at: http://www.cl.cam.ac.uk/jgd1000/combine/combine.html, 2000.

[49] L. I. Kuncheva. "That elusive diversity in classifier ensembles ", *Pattern Recognition and Image Analysis*,Vol. 26(52), pp. 1126-1138, 2003.

[50] L. Kuncheva. "Combining Pattern Classifiers - Methods and Algorithms ", Wiley, 2004.

[51] L. Xu, A. Krzyzak, C. Suen. "Methods of combining multiple classifiers and their applications to handwriting recognition ", *IEEE Transactions on Man and Cybernetics Systems*, Vol. 22(3), pp. 418-435, 1992.

[52] A. Kong, D. Zhang, G. Lu. "A study of identical twins palmprints for personal verification ", *Pattern Recognition*, Vol. 39, pp. 2149-2156, 2006.

[53] D. Maltoni, D. Maio, A. K. Jain, S. Prabhakar. "Handbook of Fingerprint Recognition ", *Springer-Verlag*, 2003.

[54] K. Nandakumar. "Multibiometric Systems: Fusion Strategies and Template Security ", *PhD thesis, Department of Computer Science and Engineering*, Michigan State University, 2008.

[55] R. Clarke. "Biometrics and Privacy ", 2001.

[56] M.M. Ghazi, H.K. Ekenel. "A Comprehensive Analysis of Deep Learning Based Representation for Face Recognition ", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Las Vegas, USA, pp. 34-41, 2016.

[57] Y. Taigman, M. Yang, M.A. Ranzato, L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification ", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, USA, pp. 1701-1708, 2014.

[58] L. Tian, C. X. Fan, Y. Ming, Y. Jin. "Stacked PCA network (SPCANet):an effective deep learning for face recognition ", *Proc. IEEE Int. Conf. Digital Signal Process*, pp. 1039-1043, 2015.

[59] Z. Y. Feng, L. W. Jin, D. P. Tao, S. P. Huang. "DLANet: A manifold-learning-based discriminative feature learning network for scene classification ", *Neurocomputing*, Vol. 157, pp. 11-21, 2015.

[60] L. Zhang, Lei. Zhang, D. Zhang. "Finger-Knuckle-Print Verification Based on Band-Limited Phase-Only Correlation ", *Springer-Verlag*, Berlin Heidelberg, pp. 141-148. 2009.

[61] J. Donahue, Y. Jia, O. Vinyals. "DeCAF: a deep convolutional activation feature for generic visual recognition ", *Comput. Sci.*, Vol. 50(1), pp. 815-830, 2013.

[62] X. Gu, P. P. Angelov, C. Zhang, P. M. Atkinson. "A Massively Parallel Deep Rule-Based Ensemble Classifier for Remote Sensing Scenes ", *IEEE Geoscience and Remote Sensing Letters*, Vol. 15(3), pp. 345-349, 2018.

[63] A.S. Razavian, H. Azizpour, J. Sullivan. "CNN features off-the-shelf: an astounding baseline for recognition ", *Computer Vision and Pattern Recognition Workshops*, pp. 24-29, 2014.

[64] A. Krizhevsky, I. Sutskever, G.E. Hinton. "Imagenet classification with deep convolutional neural networks ", *Advances in Neural Information Processing Systems 25; Curran Associates*, New York, USA, pp. 1097-1105, 2012.

[65] G. E. Hinton, S. Osindero, Y.-W. Teh. "A fast learning algorithm for deep belief nets ", *Neural computation*, Vol. 18(7), pp. 1527-1554, 2006.

[66] A. Meraoumia, L. Laimeche, H. Bendjenna, S. Chitroub. "Do We Have to Trust the Deep Learning Methods for Palmprints Identification? ", *Proceedings of the Mediterranean Conference on Pattern Recognition and Artificial Intelligence*, Tebessa, Algeria, pp. 85-91, 2016.

[67] T. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma. "PCANet: a simple deep learning baseline for image classification ", *IEEE Trans. Image Processing 24*, 5017, 2015.

[68] A. Meraoumia, M. Korichi, H. Bendjenna, S. Chitroub. "Multispectral palmprint identification method using rotation invariant variance measures ", *IEEE International Conference on Information Technology for Organizations Development (IT4OD)*, Fez, Morocco, pp. 1-6, 2016.

[69] X. Gu, P. P. Angelov. "Self-organising fuzzy logic classifier ", *Information Scinces*, Vol. 447, pp. 36-51, 2018.

[70] N. Kasabov. "Evolving Connectionist Systems: The Knowledge Engineering Approach ", *Second edn. Springer*, New York, USA, 2007.

[71] S. Pang, T. Ban, Y. Kadobayashi, N. Kasabov. "Personalized mode transductive spanning SVM classification tree ", *Information Sciences*, Vol. 181(11), pp. 2071-2085, 2011.

[72] A. Esposito, M. Marinaro, D. Oricchio, S. Scarpetta. "Approximation of continuous and discontinuous mappings by a growing neural RBF-based algorithm ", *Neural Networks*, Vol. 12, pp. 651-665, 2000.

[73] Y. Chang, W. Li; Z. Yang. "Network Intrusion Detection Based on Random Forest and Support Vector Machine ", *IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, Vol. 1, pp. 635-638, 2017.

[74] I. Nakanishi, Y. Sodani. "SVM-Based Biometric Authentication Using Intra-Body Propagation Signals ", *Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 561-566, 2010.

[75] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez. "Target dependent score normalization techniques and their application to signature verification ", *IEEE Trans. Syst. Man Cybern. C Appl. Rev.*, Vol. 35(3), pp. 418-425, 2005.

[76] M. V. Karki, S. S. Selvi. "Multimodal biometrics at feature level fusion using texture features ", *International Journal of Biometrics and Bioinformatics*, Vol. 7(1), pp. 58-73, 2013.

[77] M.J. Sudhamani, M.K. Venkatesha, K.R. Radhika. "Revisiting Feature level and Score level Fusion Techniques in Multimodal Biometrics System ", *Proceedings of International Conference on Multimedia Computing and Systems (ICMCS)*, pp. 881-885, 2012.

[78] The Hong Kong Polytechnic University (PolyU) Finger-Knuckle-Print Database. Available at: http://www.comp.polyu.edu.hk/∼biometrics/FKP.htm.

[79] J. Schmidhuber. "Deep learning in neural networks: An overview ", *Neural Networks*, Vol. 61, pp. 85-117, 2015.

[80] 8 E. Marchi, F. Vesperini, F. Eyben, S. Squartini, B. Schuller. "A novel approach for automatic acoustic novelty detection using a denoising autoencoder with bidirectional LSTM neural networks ", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'15)*, p. 5, 2015.

[81] G. Kaur, A. K. Yadav, S. Chaudhary. "An improved approach to multibiometrics security ", *Int. J. Comput. Sci. Commun.*, Vol. 5(1), pp. 181-187, 2014.

[82] M.J. Sudhamani, M.K. Venkatesha, K.R. Radhika. "Revisiting Feature level and Score level Fusion Techniques in Multimodal Biometrics System ", *Proceedings of International Conference on Multimedia Computing and Systems (ICMCS)*, pp. 881-885, 2012.

[83] P. Angelov, X. Gu. "MICE: Multi-layer multi-model images classifier ensemble ", *3rd IEEE International Conference on Cybernetics (CYBCONF 2017)*, UK, pp.1-8, 2017.

[84] M. Oquab, L. Bottou, I. Laptev, J. Sivic. "Learning and transferring mid-level image representations using convolutional neural networks ", *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1717-1724, 2014.

[85] L. Deng. "A deep convolutional neural network using heterogeneous pooling for trading acoustic invariance with phonetic confusion ", *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, pp. 6669-6673, 2013.

[86] L. Zhang, Lei Zhang, D. Zhang, H. Zhu. "Ensemble of local and global information for finger-knuckle-print recognition ", *Pattern Recognition*, Vol. 44(9), pp. 1990-1998, 2011.

[87] L. Zhang, Lei Zhang, D. Zhang, H. Zhu. "Online finger-knuckle-print verification for personal authentication ", *Pattern Recognition*, Vol. 43(7), pp. 2560-2571, 2010.

[88] L. Zhang, Lei Zhang, D. Zhang. "Finger-knuckle-print: a new biometric identifier ", *Proceedings of the IEEE International Conference on Image Processing*, pp. 1981-1984, 2009.

[89] L. Zhang, Lei Zhang, D. Zhang, Z. Guo. "Phase congruency induced local features for finger-knuckle-print recognition ", *Pattern Recognition*, Vol. 45(7), pp. 2522-2531, 2012.

[90] A. Kumar, Y. Zhou, "Personal identification using finger knuckle orientation features ", *Electronic Letters*, Vol. 45, No. 20, pp.1023-1025, 2009.

[91] S.S. Kulkarni, R.D. Rout. "Secure biometrics: Finger knuckle print", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 1(10), pp. 852-854, 2012.

[92] I. Jain, C. Stroka, J. Yan, W.M. Huang, M.K. Iovine. "Bone growth in zebrafish fins occurs via multiple pulses of cell proliferation ", *Developmental dynamics : an official publication of the American Association of Anatomists*, Vol. 236(9), pp. 2668-2674, 2007.

[93] L. Damon, J. Patrick. "Finger surface as a biometric identifier ", *Computer Vision and Image Understanding*, 2005.

[94] M. Choras, R. Kozik. "Knuckle Biometrics Based on Texture Features ", *International Workshop on Emerging Techniques and Challenges for Hand-Based Biometrics*, 2010.

[95] L. Zhang, Li. Hongyu, Y. Shen. "A Novvel Reisz Transforms based Coding Scheme for Finger-Knuckle-Print Recognition ", *Proceedings of International Conference on Hand Based Biometrics*, 2011.

[96] C.Ravikanth, A. Kumar. "Biometric authentication using finger-back surface ", *Proc. CVPR*, Vol. 45, pp. 1-6, 2007.

[97] L. Gail, M. Jungbluth, T. Pasko William, J. Jusko. "Factors affecting ceftriaxone plasma protein binding during open heart surgery ", *J. of Pharm. Sci.*, 1989.

[98] A. Kumar, Ch. Ravikanth. "Personal Authentication Using Finger Knuckle Surface ", *IEEE Transactions On Information Forensics and Security*, Vol. 4, No. 1, pp. 98-110, 2009.

[99] L. Zhang, Lei Zhang, D. Zhang. "Monogenic Code: A Novel Fast Feature Coding Algorithm with Applications to Finger-KnucklePrint Recognition ", *Proceedings of the International Workshop on Emerging Techniques and Challenges for Hand Based Biometrics*, 2010.

[100] A. Meraoumia, S. Chitroub, A. Bouridane. "An Efficient Hand-Based Biometric Recognition System Using Finger-Knuckle-Print Data ", *Recent Patents on Telecommunication*, Vol. 1 No. 1 pp. 151-162, 2012.

[101] V. Morales, C. Regnard, A. Izzo, I. Vetter, P.B. Becker. "The MRG domain mediates the functional integration of MSL3 into the dosage compensation complex ", Mol. Cell. Biol, Vol. 25(14), pp. 5947-5954, 2005.

[102] L. Zhang, Lei Zhang, D. Zhang. "'Finger-knuckle-print verification based on band-limited phase-only correlation ", *The 13th International Conference on Computer Analysis of Images and Patterns-CAIP*, Germany, pp. 141-148, 2009.

[103] T. Mitchell. "Machine Learning ", *McGraw Hill*, ISBN 978-0-07-042807-2, 1997.

[104] C. M. Bishop. "Pattern Recognition and Machine Learning ", *Springer*, ISBN 978-0-387-31073-2, 2006.

[105] M. Mohri, A. Rostamizadeh, A. Talwalkar. "Foundations of Machine Learning ", *The MIT Press*, ISBN 9780262018258, 2012.

[106] P. Langley. "The changing science of machine learning ", *Machine Learning*, Vol. 82(3), pp. 275-279, 2011.

[107] N. Vapnik, Vladimir. "Support-vector networks ", *Machine Learning*, Vol. 20(3), pp. 273-297, 1995.

[108] A. Aizerman, Braverman, M. Emmanuel,Lev I. Rozonoer. "Theoretical foundations of the potential function method in pattern recognition learning ", *Automation and Remote Control*, Vol. 25, pp. 821-837, 1994.

[109] Boser, E. Bernhard, Guyon, M. Isabelle, N. Vapnik, Vladimir. "A training algorithm for optimal margin classifiers ", *Proceedings of the fifth annual workshop on Computational learning theory*, p. 144, 1992.

[110] P. K. Atrey, M. A. Hossain, A. El Saddik, M. S. Kankanhalli. "Multimodal fusion for multimedia analysis: A survey ", *Multimedia Syst.*,Vol. 16, No. 6, pp. 345-379, 2010.

[111] C. R. Huang, Y.-T. Chen, W.-Y. Chen, H.-C. Cheng, and B.-S. Sheu. "Gastroesophageal reflux disease diagnosis using hierarchical heterogeneous descriptor fusion support vector machine ", *IEEE Trans. Biomed. Eng.*, Vol. 63, No. 3, pp. 588-599, 2016.

[112] D. S. Broomhead, D. Lowe. "Multivariate functional interpolation and adaptative networks ", *Complex Systems*, Vol. 2, pp. 321-355, 1988.

[113] J. Hertz, A. Krough, R. G. Palmer. "Introduction to the Theory of Neural Computation ", *Addison Wesley*, Redwood City, Canada 1991.

[114] L. Breiman. "Random Forests ", *Machine Learning*, Vol. 45(1), pp. 5-32, 2001.

[115] H. Tin Kam. "Random Decision Forests (PDF) ", *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, pp. 14-16, 1995.

[116] H. Trevor, T. Robert, J. Friedman. "The Elements of Statistical Learning (2nd ed.) ", *Springer*, ISBN 0-387-95284-5, 2008.

[117] R. Caruana, N. Karampatziakis,A. Yessenalina. "An empirical evaluation of supervised learning in high dimensions ", *Proceedings of the 25th International Conference on Machine Learning*, pp. 96-103, 2008.

[118] A. Ethem. "Introduction to Machine Learning ", *MIT Press.*, p. 9. ISBN 978-0-262-01243-0, 2010.

[119] H. Tin Kam. "A Data Complexity Analysis of Comparative Advantages of Decision Forest Constructors ", *Pattern Analysis and Applications*, pp. 102-112, 2002.

[120] G. James, D. Witten, T. Hastie, R. Tibshirani. "An Introduction to Statistical Learning ", *Springer*, pp. 316-321, 2013.

[121] G. Seymour. "Predictive Inference: An Introduction ", *New York: Chapman. Hall.*, ISBN 0-412-03471-9, 2016.

[122] M. Faundez-Zanuy. "Data fusion in biometrics ", *IEEE Aerospace and Electronic Systems Magazine*, Vol. 20, pp. 34-38, 2005.

[123] J. Fierrez-Aguilar. "Adapted Fusion Schemes for Multimodal Biometric Authentication ", *PhD thesis, Universidad Politecnica de Madrid*, 2006.

[124] T.Sim, S. baker, M. Bsat. "The CMU pose, illumination, and expression database ", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25(12), pp. 1615-1618, 2003.

[125] R.Raghavendra, A. Rao, G. H. Kumar. "Multisensor biometric evidence fusion of face and palmprint for person authentication using Particle Swarm Optimization (PSO) ", *International Journal of Biometrics*, Vol. 2(1), pp. 119-33, 2010.

[126] P. J.Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek. "Overview of the face recognition grand challenge ", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 947-954, 2005.

[127] X. Y. Jing, Y. F. Yao, J. Y. Yang, M. Li, D. Zhang. "Face and palmprint pixel level fusion and kernel DCV-RBF classifier for small sample biometric recognition ", *Pattern Recognition*, Vol. 40, pp. 3209-3224, 2007.

[128] A. Ross, R. Govindarajan. "Feature level fusion using hand and face biometrics ", *SPIE Conf. on Biometric Tech. for Human Identification II*, pp. 196-204, 2005.

[129] G.Feng, K. Dong, D. Hu, D. Zhang. "When faces are combined with palmprint: A novel biometric fusion strategy ", *First International Conference on Biometric Authentication (ICBA)*, pp. 701-707, 2004.

[130] A. Rattani, D. R. Kisku, M. Bicego, M. Tistarelli. "Feature level fusion of face and fingerprint biometrics ", *1st IEEE International Conference on Biometrics: Theory, Applications and Systems*, pp. 1-6, 2010.

[131] E. Bailly-Baillire, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Marithoz, J. Matas, K. Messer, V. Popovici, F. Pore, B. Ruiz, J. P. Thiran. "The BANCA database and evaluation protocol ", *International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pp. 625-638, 2003.

[132] M. He, et al. "Performance evaluation of score level fusion in multimodal biometric systems ", *Pattern Recognition*, Vol. 43(5), pp. 1789-1800, 2010.

[133] A. K. Jain, L. Hong, R. Bolle. "On-line fingerprint verification ", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19(4), pp. 302-314, 1997.

[134] A. K. Jain, K., Nandakumar, A. Ross. "Score normalization in multimodal biometric systems ", *Pattern Recognition*, Vol. 38, pp. 2270-2285, 2005.

[135] R. Frischholz, U. Dieckmann. "A multimodal biometric identification system ", *IEEE Computer*, Vol. 33(2), pp. 64-68, 2000.

[136] P. Yu, D. Xu, H. Zhou, H. Li. "Decision fusion for hand biometric authentication ", *IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS)*, Vol. 4, pp. 486-490, 2009.

[137] P. Huber. "Robust Statistics ", *John Wiley, Sons*, 1981.

[138] X. Yuan, W. Gan. "A statistical approach towards performance analysis of multimodal biometric systems ", *IEEE International Conference on Robotics and biometrics*, pp. 877-882, 2008.

[139] S. Haykin. "Neural network - A comprehensive foundation ", *(2nd edition) Prentice Hall*, 1999.

[140] I. Goodfellow, Y. Bengio, A. Courville. "Deep Learning ", *Boock* at: http://www.deeplearningbook.org..

[141] Y. Bengio, O. Delalleau, N. Le Roux. "The curse of highly variable functions for local kernel machines ", *In NIPS'2005*, 2006.

[142] L. Deng, Y. Dong. "Deep Learning: Methods and Applications ", *Signal Processing*, Vol. 7, Nos. 3-4, 2014.

[143] F. F. Li, A. Karpathy. "Convolutional Neural Networks for Visual Recognition ", 2015.

[144] S. B. Imandoust, M. Bolandraftar. "Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background ", *Int. Journal of Engineering Research and Applications*, Vol. 3(5), pp. 605-610, 2013.

[145] P. Mulak, N. Talhar. "Analysis of Distance Measures Using K-Nearest Neighbor Algorithm on KDD Dataset ", *International Journal of Science and Research (IJSR)*, Vol. 4(7), pp. 2101-2104, 2015.

[146] S. M. Tamura, T. Yamawaki. "Textural Features Corresponding to Visual Perception ", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 8(6), pp. 460-473, 1978.

[147] J. K. Kim, H. W. Park. "Statistical Textural Features for Detection of Microcalcification in Digitized Mammograms ", *IEEE Transaction on Medical Imaging*, Vol. 18(3), 1999.

[148] A. A. Kadir, X. Xu, E. Hammerle. "Virtual Machine Tools and Virtual Machining-A Technological Review ", *Journal of Robotics and Computer-Integrated Manufacturing*, Vol. 27(3), pp. 494-508, 2011.

[149] Ch. Zhang, P. Hu, J. Fan. "The Research on Line Features Extraction from Images of Industrial Parts Based on CAD Data ", *IEEE International Conference on Audio, Language and Image Processing*, pp. 453-457, 2010.

[150] M. Zahedi, P. Dreuw, D. Rybach, T. Deselaers, H. Ney. "Geometric Features for Improving Continuous Appearance-based Sign Language Recognition ", *Journal of British Machine Vision Conference (BMVC)*, Vol. 3, pp. 1019-1028, 2006.

[151] A. Meraoumia. "Modèle de Markov caché appliqué à la multi-biométrie ", *Thèse de Doctorat (USTHB)*, 2014.

[152] C. Jie Ng, A. B. Jin Teoh. "DCTNet : A Simple Learning-free Approach for Face Recognition ", *IEEE Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 16-19 Dec, 2015.

[153] J. Han. "Data Mining: Concepts and Techniques ", *Morgan Kaufmann Publishers*, 2005.

[154] S. Y. Kung, M. W. Mak. "Machine Learning in Bioinformatics, Chapter 1: Feature Selection for Genomic and Proteomic Data Mining ", *John Wiley - Sons, Hoboken*, 2009.

[155] B. E. Boser, I. M. Guyon, N. Vapnik. "A training algorithm for optimal margin classifiers ", *Proceedings of the fifth annual workshop on Computational learning theory, ACM*, pp. 144-152, 1992.