PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

Akli Mohand Oulhadj University
-Bouira-
Faculty of Science and Applied Science

ACADEMIC MASTER

Presented at:

Department: *Electrical Engineering*

Domain: *Science and Technology*

Branch: *Telecommunications*

Speciality: *Telecommunications Systems*

Through:

BEN-ADJALI Azzeddine
and
BOUCHENAK Zine el abidine

Theme

Automatic Visual Kinship Verification

Presented on: ...../..../**2023**

Before the commission composed of :

Mr :........................	Prof.	Univ. Bouira	President
ALIMOHAD Abdennour	M.C.A	Univ.Bouira	Supervisor
BELABBACI El ouanas	Doc	Univ. Bouira	Co-Framer
....	M.C.B	Univ. Bouira	Examiner
....	M.A.A	Univ. Bouira	Examiner

Academic Year 2022/2023

Dedication

All the words cannot express our gratitude, love, respect and recognition, it is simply that I dedicate this modest work:

To my dearest mother, for the love she has always given me, for her support, her encouragement, her sense of sacrifice and her prayers, without you I would not be here.

And to my dearest father who always encouraged and motivated me during my studies, I thank him infinitely for his precious advice which always guided my steps towards success and for the noble values he taught me, the education and the moral support from him.

To my big sister, you have been present at my side during all the stages of this work, I am so grateful to have a sister like you.

No words, no dedication could express my respect, consideration, and love for the sacrifices you have made for my education and welfare. I love you.

To my partner **Bouchenak Zine el abidine** I had a great pleasure to work with you on all the projects we did in our student career. Our supervisor **Mr. ALIMOHAD Abdennour** thank you for all that you have done for us, we could not have achieved this work has this quality without your efforts. Our co-supervisor **BELABBACI El ouanas**, you have been helping us during this work with enormous efforts thank you for all what you did.

BEN-ADJALI Azzeddine.

Dedication

With the expression of my gratitude, I dedicate this modest work .

To my dear father , my precious offer from the god, who owes my life, my success and all my respect.

To my adorable mother.. who suffered without letting me suffer, who never said no to my demands and who spared no effort to make me happy.

To my dear wife who never stopped advising, encouraging and supporting me throughout my studies.

To my dear little son **Nazim**.

To all my familly and my neighber **Djamal Chaibi** Thank you for your love and encouragement. Without forgetting my buddy **BEN ADJALI Azzeddine** for his moral support, his patience and his understanding during this project, It is a great honor for me to work with you.

To our supervisor, **Mr. ALIMOHAD Abdenour**, I express my gratitude to him for his precious advice and his commitment throughout this work.

To our co-supervisor **BELABACI El ouanas**. Thank you for your help, advice, and everything you have done for us.

Bouchenak Zine el abidine.

Acknowledgements

First of all, we thank ALLAH from the bottom of our hearts for having enlightened us on the right path.

We want to thank our familys for their emotional contribution and their sacrifices. We cannot thank enough our supervisors **Dr. ALIMOHAD Abdennour**, **Mr. BELABBACI El ouanas** for their precious time, help and advice. Their availability, their patience throughout thiswork has been beneficial to us.

We are deeply grateful to all members of the jury for agreeing to read the manuscript and judge this thesis.

Our thanks also extend to all our teachers Akli Mohand Oulhadj University – Bouira during the years of study. And all those who are dear to us and who helped us to realize this work.

and gave us the will and the desire to do everything to succeed.

Abstract

Verification of kinship from facial images is attracting more and more attention from the research community, is an emerging research topic in computer vision. Checking whether two people are from the same family or not can be automatically checked by facial images. Many potential applications: such as creating family trees, organizing family albums, annotating images; the search for missing children and forensic medicine, are targeted by the verification of kinship.

This paper presents a successful kinship verification system, which utilizes two consecutive methods (MSR+NDM) in the image preprocessing stage to enhance image quality and overcome issues relating to contrast, lighting, and noise. Additionally, we propose a new descriptor based on the histograms of a Two dimensional Discrete Wavelet Transform (Hist-2D DWT). We further investigate the complementarity of handcrafted (LPQ, Hist-2D DWT) and deep features (VGG16, ResNet50) by fusing them at the score level using the Logistic Regression method.

Extensive experiments conducted on two kinship datasets, verification accuracies of **95.18%** and **91.81%** have been reached under Cornell KinFace and TS KinFace datasets.

Key words: Kinship Verification, Hist-2D DWT descriptor,Deep Features, ShallowFeatures, MSR+NDM, LR Fusion.

Contents

List of Figures

List of Tables

Lists of Acronyms

ACC	Accuracy
AFR	Automatic Face Recognition
BSIF	Binarized Statistical Image Features
CNN	Convolutional Neural Network
CS	Cosine Similarity
DNA	Deoxyribonucleic acid
DL	Deep Learning
FD	Father-daughter
FS	Father-son
GFGD	Grandfather-granddaughter
GFGS	Grandfather-grandson
GMGD	Grandmother-granddaughter
GMGS	Grandmother-grandson
Hist-2D DWT	Histogram of Two Dimensional Discrete Wavelet Transform
HOG	Histogram Oriented Gradients
HH	High-High
HL	High-Low
KNN	K Nearest Neighbors
KinFaceW	Kinship Face in the Wild
LBP	Local Binary Pattern
LDA	Linear Discriminant Analysis
LOOP	Local optimal oriented patterns
LPQ	Local Phase Quantization

LR	Logistic Regression
LH	Low-High
LL	Low-Low
MD	Mother-daughter
MS	Mother-son
MPCA	Multilinear Principal Component Analysis
MSR	Multi Scale Retinex
MSIDA	Multilinear Side Information Discriminant
NDM	Normal Distribution Mapping
NRML	Neighborhood Repelled Metric Learning
PCA	Principal Component Analysis
ReLu	Rectified Linear Unit
ResNet	Residual Network
SVM	Support vector machines
TXQDA	Tensor Cross-View Discriminant Analysis
VGG	Visual Geometry Group
WCCN	Within-Class Covariance Normalization

General Introduction

General Context:

Automatic kinship verification from facial images is a relatively new and challenging research problem in computer vision [1]. The most elementary kinship verification is 1-vs-1 verification, it aims to automatically check whether two people are from the same family or not by examining their facial attributes [2]. This is a field with high potential impact in research and application, it has attracted more and more attention over the years, this research is motivated by the psychological findings that stipulate that the facial image can be used to measure genetic similarity. We can divide kinship relationships into three groups:[3]

- Same-generation pairs: Brother-Brother(BB), Brother-Sister (SIBS) and Sister-Sister(SS).
- first generation pairs: Father-Son (FS), Father-Daughter (FD), Mother-Son (MS) and Mother-Daughter (MD).
- second generation pairs: Grandfather-grandson (GFGS), Grandfather-granddaughter (GFGD), Grandmother-grandson (GMGS) and Grandmother-granddaughter (GMGD).

The research community is paying more attention to automatic kinship verification, several experiments have been conducted, we classify them into three categories: those based on feature extraction (handcrafted features and deep features), others based on deep learning [4] and finally those based on metric learning [5]. Large-scale datasets' availability has allowed deep learning to advance significantly in a number of kinship recognition tasks [6]. Our overall goal in this work is to create and evaluate new computational models in order to establish an automatic kinship verification system that ensures performance close to the state of the art, based on the application of computer vision and deep learning techniques.It should be noted that we focus on improving the accuracy of the system and not on optimizing resources.

Problematic and motivation:

In many situations, such as organizing and resolving identities in photo albums, finding relatives in public databases, identifying the relationship of a victim or suspect by law enforcement, reviewing asylum applications when kinship ties need to be established, etc., automatic kinship verification using facial images can be very helpful. Similarly , automatic kinship verification from videos can be used in security, surveillance, and border control applications. By using surveillance cameras to confirm the relationship between an adult and his or her child, it is possible to stop illicit child trafficking.

The most reliable method for establishing a relationship between two people right now is through DNA testing, which is also the most expensive, time-consuming, and complex. However, an automatic kinship verification algorithm can deliver cost-effective results in circumstances that call for real-time processing with uncooperative users.

However, the high degree of variability in appearance influences, such as genetic, age, and gender differences, as well as inheriting the challenges of looking for face verification from images captured in nature under unfavorable pose, expression, or lighting, make kinship verification from faces a relatively challenging task despite recent advances.

Contribution:

Among the modalities in the system of automatic recognition, the "automatic recognition of kinship based on facial images" by the fact that it is permanent and unique. Researchers are still trying to develop recognition systems through mathematical tools usually complex to discriminate between classes. The objective followed in this dissertation proposes an approach to improve the performance of the kinship verification system by using several methods with a set of operations. For this purpose, we have made the combination between different methods of extraction of deep and shallow features , which allowed us to obtain a better adaptation for the realization of an automatic kinship recognition system and the improvement of its robustness.

Plan of the dissertation:

This dissertation contains a general introduction in which we discuss the problem of automatic verification of kinship links, we also emphasize our modest contribution to this vast field that has motivated a large number of researchers over the past ten years.

Four main chapters are presented in this work as follows:

- **Chapter I:** In this first chapter we give an overview of the automatic kinship verification from faces by presenting some definitions and terminology necessary to understand this topic.
- **Chapter II:** In this chapter we introduce the available public databases and some related works. we also discuss the kinship approches such as face pre-processing, feature extraction methods (the handcrafted feature-based method and the deep feature-based method).
- **Chapter III:** This chapter is dedicated to the design of the proposed solution, we explain the different steps used for the implementation of our system(face pre-processing (with MSR+NDM) Features extraction using deep (VGG16 and ResNet50) and shallow features (LPQ, and the proposed descriptor Hist-2D DWT)
- **Chapter IV:** The experimental results are presented and discussed in this last chapter. We study the task of automatic kinship verification using several experiments.

Chapter 1

Basic Concepts for Kinship Verification

1.1 Introduction:

Kinship authentification is a procedure that verifies a person's relationship using facial characteristics. An excellent example of this ability is being able to identify people in photographs. This chapter discusses the background knowledge and ideas required to comprehend the kinship verification subject.The discussion of kinship terms is presented after the description of kinship verification in computer vision.It was made clear how different a kinship authentification system is from a facial recognition system. We also provide a summary of the issues with kinship verification using human facial pictures.

1.2 Automatic kinship recognition:

People have frequently been forced to decide whether there should be or should not be a family relationship between two or more people. Sometimes this can be quickly identified by human eyes due to physical characteristics. The presence of distinguishable characteristics, such as the color and shape of the eyes, snout, nasolabial folds, and mouth, is helpful in determining the relationship, but it is not always clear. Currently, the most reliable method for determining whether a parent-child relationship exists is the implementation of a DNA-test. However, this is challenging due to the high cost, the length of the testing procedure, and the difficulty in understanding the results.

Even though it was first introduced in 2010 by Fang et al. [7], automatic recognition of the parent-child relationship is still in its infancy. There is still no practical application using an automatic visual verification system after all these years. This delay is caused by the need for a large and varied set of data in order to accurately describe the distribution of families in the real world, as well as the difficulties in verifying faces using images taken in the natural world with variable location, brightness, face expressions and aging. Which makes this problem extremely difficult to resolve. So genetic information about kinship relationships is much less discriminatory in the visual domain than in more conventional problems, also we can't justify why two people appear to us in a kinship relationship or transfer our knowledge and abilities to algorithms [8].

1.3 Psychological Aspects of Kinship:

Humans have the ability to recognise whether a person belongs to their own family or not, they are even able to guess whether a couple of strangers are related, and this is due to their perception of similarity as a result of allocentric behaviour. This has led to several psychological studies to assess the performance of humans in kinship recognition by asking people to evaluate easy images of people belonging to the same family, from the results of which we have drawn the following points [9]:

- The ability to recognise kinship improves with time, the older the person gets the more his or her abilities increase.
- The capacity for both sexes is equivalent.
- There is no general rule that can be generalised for all relationships.
- Facial likeness analysis can only be performed if the full face is available.
- Human facial similarity analysis is usually done by "patching".

1.4 Kinship Verification in computer Vision:

Recently, a new field of computer vision called kinship verification has emerged. This field requires the use of information about face shapes in place of other types of information like voice, imprint, or even iris, and it does so by employing a set of characteristics that are stored in a vector based on physiological characteristics. Kinship verification using the face modality consists of an essential biometric system in the various fields of application thereof [10] the fact that this operation can be done remotely can represent an enormous privilege because verification without direct interaction between users and the system can be useful as an example to avoid contamination (covid-19 pandemic which put the world on alert in 2020), also verification can be done remotely , reliably and in real time, it reinforces security systems and minimizes costs .

This field is very interesting but not yet mature. There are indeed limitations in existing systems due to different variations in age, brightness, position and other inter- class and intra-class variations.

This operation is embodied in a system that takes 2 input images corresponding to the face of any two people and returns an output result that classifies them according to whether they are related to each other (KIN) or not (NO KIN).

1.5 Face verification and kinship verification:

The kinship verification procedure may initially be mistaken for the facial verification procedure. These two systems are actually distinct from one another, but they share some characteristics. The highest degree of face verification could be said to be kinship verification [11].

The basic structure is what kinship verification and face verification systems have in common, and Tab I.1 compares how they vary from one another.

Kinship verification	**Face verification**
• Extract features from different person • Verify the relationship • Different trait of query image 1 and query image 2 • Highest level system • In decision stage Kin or not Kin • Accuracy is around 90	• Extract features from same person • Verify or identify • Same trait of query image 1 and query image 2 • Height level system • In decision stage matched or not matched • Performance of the machine is very roughly as accurate as human

Table 1.1: Difference between kinship and face Recognition system.

1.6 Motivations and applications :

Automatic kinship verification using facial images has several applications such as locating relatives in public databases, determining the parentage of a victim or suspect by law enforcement, screening asylum applications where family ties must be determined, the organization and resolution of identities in photo albums. Kinship verification has several security aspects: relatives of people identified as a security threat can be identified using an automatic kinship verification framework. Automatic determination of kinship information can also be used to enhance automatic face recognition capabilities by using kinship characteristics as soft biometrics. On the other hand, automatic kinship verification in videos is a relatively unexplored area of research and can be very useful in various contexts such as security, surveillance and immigration control. For example, during the investigation of surveillance videos of the Boston Marathon bombing, two male suspects were identified as the perpetrators. Later it was established that the two men were brothers which led to their identification. An automatic kinship verification system that determines kinship in a video could have expedited this investigation. Another application of kinship verification is border control. using surveillance videos that can be used to validate the relationship between an adult and a child, thus preventing the illegal trafficking of children. In addition, video kinship verification can validate or invalidate the parentage claims of refugees and asylum seekers. Currently, as part of its reunification program, the US State Department performs DNA tests to allow people who have relatives in the United States to enter the United States as refugees. Rapid-DNA is used for this purpose, but an automatic kinship verification algorithm can produce cost-effective results in real time. Relationship information may also be used to manage multimedia on social media websites such as Facebook and Youtube. In many cases, family members have different Youtube channels where they upload daily videos. Kinship information can be applied to automatically tag these videos and identify family members present in these videos. Related context in videos can also be used for automatic indexing and organization of videos, making them easily searchable.

1.7 Kinship challenges:

Kinship verification through facial images is challenging due to the high degree of appearance variability of influences such as genetic difference, age gap and gender difference [12]. In summary, the following two factors have significant influence on problem-solving: Unique challenges which appear only in the kinship verifications systems, and common challenges which already exist in the automatic face recognition systems, both challenges are discussed in details below.

1.7.1 Unique Challenges:

Humans inherit facial features from their ancestors, and kinship verification is different from face recognition or verification. In kinship verification, the problem is to extract similar features from different persons that share some traits in common. Three unique major challenges in kinship verification are identified, which will help to understand the problem and provide a general guide to develop a computational system [13].

- Familial traits (features) are measured across age and gender for people of different ages, genders, or both (father-son, brother-sister, father-daughter) [13].

- Familial traits have "special" properties for each family pair, with sons and daughters inheriting traits differently from their parents even if they're with same sex [13].

- Kinship has a stochastic combination of familial traits, which must be measured by a "stochastic" (rather than a fixed) combination. Familial traits are the building blocks for kinship measure, and recent approaches have processed it similar to an automatic face verification problem [2] [13]

Figure 1.1: Illustration of Unique Challenges: a)Across age Father-Son, b) Across gender Brother-Sister, c) Across age and gender Father-Daughter.

1.7.2 Common challenges:

Kinship verification is a subset of automatic face verification, and its appearance is sensitive to changes in facial expressions, occlusion and pose. Additionally, lighting, blurring and low resolution can also affect the appearance of kin faces. These factors can affect the appearance of kin faces in different ways.

- **Pose variations:** Head movements, such as pitch, roll and yaw, or camera changing point of views, can lead to significant changes in face appearance and/or shape, making automated face recognition across pose a difficult task. Pose correction is essential and can be achieved by using efficient techniques to rotate the face and/or align it to the image's axis, as detailed in reference [14] [15].

Figure 1.2: Illustration of pose variations [15].

- **Presence/absence of structuring elements/occlusions:** Face images taken in an unconstrained environment often require effective recognition of faces with disguise or altered by accessories and/or occlusions. This is illustrated in Fig I.3, where elements such as hats, glasses or beard can represent a factor for occlusion. Texture-based algorithms can help [15].

Figure 1.3: Illustration of Presence/absence of structuring elements/occlusions [15].

- **Facial expression changes:** Human expressions are composed of macro-expressions, such as anger, disgust, fear, happiness, sadness or surprise, and other involuntary, rapid facial patterns, such as micro-expressions. These expressions generate non-rigid motion of the face, which is important for both the evaluation of emotional states and automated face recognition. Fig I.4shows the variability in face appearance caused by changes in emotional states [15].

Figure 1.4: Illustration of Facial expression changes [15].

- **Ageing of the face:** Face appearance changes can be caused by ageing, which can have a significant impact on the face recognition process [16]. To overcome this issue, methods need to take into account facial ageing patterns [15].

Figure 1.5: Illustration of aging [15].

- **Varying illumination conditions:** Large variations of illuminations can degrade the performance of AFR systems, with low levels of lighting making face detection and recognition difficult. Too high levels of lighting can lead to overexposure and indiscernible facial patterns. Image processing techniques such as illumination normalization and machine learning are used to deal with these variations[15].

Figure 1.6: Illustration of illumination variation [15].

- **Image resolution and modality:** AFR performance is influenced by the quality and resolution of the face image, the set-up and modalities of digital equipment, and the use of different photographic hardware. Faces acquired in real-world conditions can lead to further challenges due to multiple modalities[15].

Figure 1.7: Illustration of variations of the image scale and resolution [15].

1.7.3 Kinship modes

Any relationship between two individuals is based on the degree of closeness or distance of that relationship. Most kinship verification researches dealt with 11 types of kin ties combined in 3 categories:

- **Parent-Child:** i.e. father-daughter (F-D), father-son (F-S) mother-daughter (M-D) and mother-son (M-S).
- **Brothers and Sisters:** i.e. brother-brother (B-B), sister-sister (S-S) and brother-sister (SIBS).
- **Grandparents-Grandchildren:** i.e. grandfather-grandson (GF-GS), grandfather-grand daughter (GF-GD), grandmother-grandson (GM-GS), grandmother-granddaughter (GM-GD).

1.8 Kinship verification system structure

Because of the difficulty and complexity of kinship verification, a framework must be set in place to deal with different approaches. The framework is divided into four components, Preprocessing, Features extraction, Similarity measurements and Verification.

- **Preprocessing stage:** The goal of preprocessing stage is to make it easy to measure kinship similarity by extracting kin faces from photos and reducing the influences of various variations.
- **Features extraction:** This stage aims to extract certain caracterstics from face image to do the training with. These features can be either handcrafted-based features or learningbased features. Most existing kinship verification methods have opted for handcraftedbased features, where each face image is first divided into several blocks and then a certain characteristic is extracted from these blocks, most common descriptors are Local Binary Patterns (LBP), Local Phase Quantization (LPQ), Binarized Statistical Image Features (BSIF), Gabor wavelets... etc. The features can be typed as low-level, high, mild-level, high-level features.
- **Similarity measurement:** The goal of this stage is to examine the similarity between features vectors that were obtained from the previous stage. There are

many methods to measure similarity based on distance calculation or statistical model learning.

- **Verification:** the final stage aims to finally verify the kin relations between data samples using various classifiers like Support Vector Machine (SVM), or K-Nearest Neighbor (KNN).

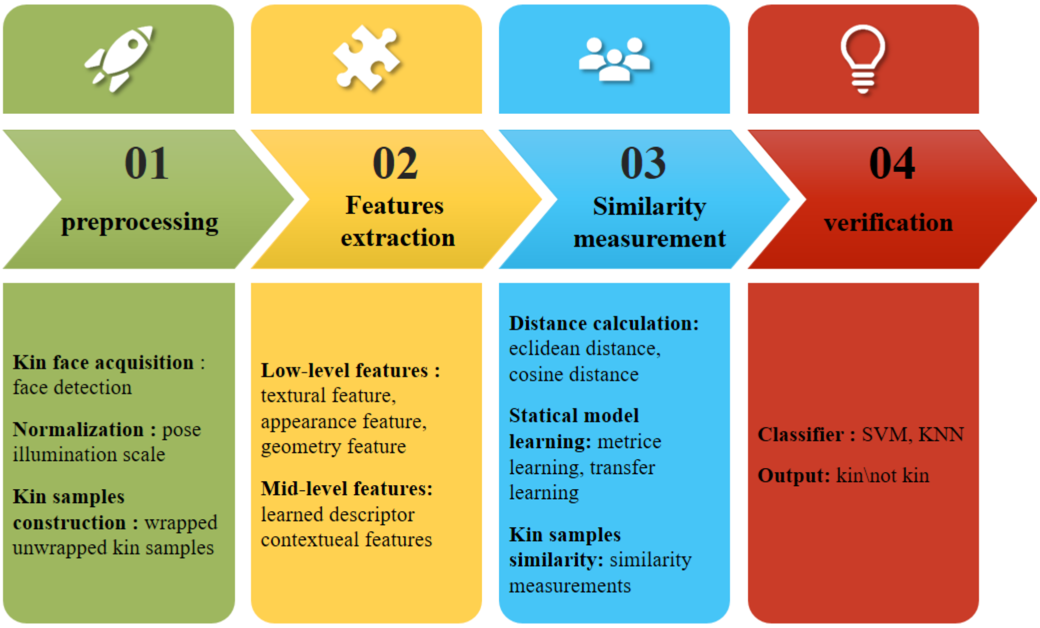

Figure 1.8: A global system framework for kinship verification.

1.9 Conclusion

In this chapter, we gave an overview of automatic kinship verification from faces. We also talked about some applications that can use this type of systems and the challenges which are facing the researchers.

In the next chapter, we will present some notions and several definitions of the design of kinship verification systems

Chapter 2

Kinship Verification Methods

2.1 Introduction:

Automatic kinship verification is a new area of research that has seen wide interest in recent years. In this chapter we summarize some of known databases.We also present related works and an overview of automatic kinship verification approaches. We discuss some of the methods used, divided into two categories: methods based on handcrafted characteristics and those based on deep characteristics,in order to know the state of the art in this field.

2.2 Databases for kinship verification:

Cornell Kinship[2]**:** It consists of 286 images pertaining to 143 subject pairs. The facial images in this database are frontal pose and have a neutral expression.

KinFaceW-I [17]: This database consists of 1066 images corresponding to 533 kin pairs. It has 156 Father-Son, 134 Father-Daughter, 116 Mother-Son, and127 Mother-Daughter kin pair images.

KinFaceW-II [17]: This database has been created such that images belonging to the kin pair subjects are acquired from the same photograph. It consists of 1000 kin pair images with an equal number of images belonging to the four kinship relationships: Father-Son, Father-Daughter, Mother-Son, and Mother-Daughter.

UB KinFace [18]: This database consists of 200 groups consisting of 600 images. Each group has one image of the child and one image belonging to the corresponding parent when they were young and when they were old. The database has 91 Father-Son, 79 Father-Daughter, 15 Mother-Son, and 21 Mother-Daughter kin pair images.

Family 101 [19]: The Family 101 dataset contains 101 different families with distinct family names, including 206 nuclear families, 607 individuals, with 14,816images. The database is particularly useful for family classification where the problem is identifying which family a particular image belongs to.

TSKinFace[20]: The Tri-subject kinship face database consist of images be-longing to the child, mother, and father. The database consists of 513 images of Father, Mother and Son group and 502 images of Father, Mother and Daughter group.

	Number of familys	Number of persons	Number of faces	Resolution	Age Variation	Family tree
CornellKin	150	300	300	100*100	No	No
UB Kinface	200	400	600	89*96	Yes	No
KinfaceW-I	-	533	1066	64*64	No	No
KinfaceW-II	-	1000	2000	64*64	No	No
TS kinface	787	2589	-	64*64	Yes	Yes
Family101	101	607	14816	100*100	Yes	Yes

Table 2.1: Summary of kinship datasets in the literature.

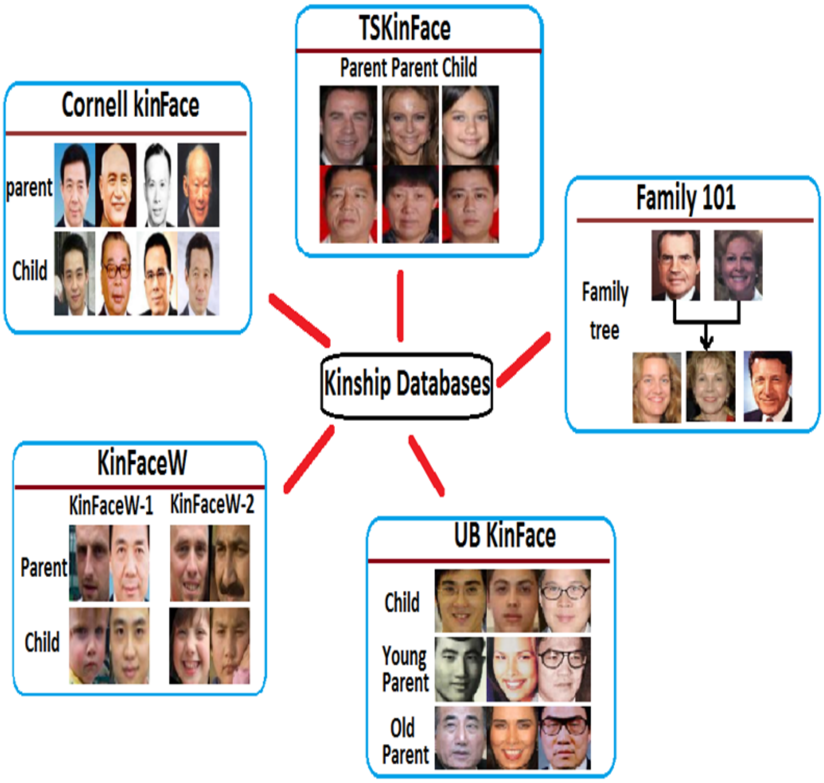

Figure 2.1: Sample images from six different data sources: Cornell KinFace, TSKinFace, KinFaceW, Family 101 and UBKinFace.

2.3 Related works:

The first research on kinship verification was published in 2010 [2]. The authors compared an automatic kinship verification technique to human performance. However, they made use of the "Cornell KinFace" database, which only has 143 parent-child pairs worth of information. They made an initial effort to confirm the similarity between parent-child pairs.

They made an initial effort to confirm the similarity between parent-child pairs. They built their classification on the extraction of 22 features of the face, including the skin color, the eyes, the mouth, the detection of distance features, and statistical features like the Histogram of Gradients (HOG). Then, either the Support Vector Machine (SVM) with a radial basis function or the KNN (K nearest neighbor) classifier with an Euclidean metric are used to identify the pairs of faces. Although this method has produced positive results, it cannot be used to prove kinship verification because of physical and genetic differences, such as the age gap between the father and son or the gender (brother/sister, for example).

Statistical-based genetic studies have demonstrated a critical observation that the faces of parents when they were young resemble those of their children more than the images captured when they are old, the latter prompted the creation of the database UB KinFace, composed of images of children's faces, young parents and elderly parents, at the end of this last Xia et al. proposed the transfer learning method (TSL) [21] which appeared in the hope of decreasing the huge divergence of distribution between children and elderly parents, and this by using an intermediate distribution close to the two distributions as well as Gabor wavelets for feature extraction. This approach improved the overall accuracy of kinship verification and made the task more discriminatory [22].

Afterwards, two KinFaceW-I KinFaceW-II databases were collected by Lu et al [17] in order to use them to guide larger research, the availability of these further motivated researchers to contribute on this topic. They also proposed the neighborhood repelled metric learning (NRML) method, metric learning allows to learn a good distance metric in order to minimize the distances between the pairs of positive images with kinship links, while by pushing further the pairs of images of those who do not have it. This method has been tested using different local feature descriptors like Local Binary Pattern (LBP), Histogram of Oriented Gradients (HOG) and Scale-Invariant Feature Transform (SIFT) [17].

Fang et al proposed a new approach for kinship verification on their "Family101" dataset [19], modeling this problem as that of reconstructing a face from the shuffled parts of a set of families , this model is inspired by the biological process of inheritance. Their approach is to segment the face into parts (eyes, nose, mouth..) instead of taking the whole face and reconstructing each part as a linear combination of a set of parts from the database, and to evaluate this approach they used a dense SIFT descriptor on resized facial images of size 61 x 49 pixels [19].

2.4 Kinship approaches:

During the last ten years, several physiological researches concerning the human capacity in the facial recognition and verification were established which led the researchers in computer vision and machine learning to define different automatic approaches with variant performances, these last cannot be directly compared since they are used on different datasets [22].

2.4.1 Face pre-processing:

Face pre-processing is the set of techniques used to prepare facial images for further analysis[23]. Some of the most common techniques used in face pre-processing include:

2.4.1.1 Face Detection (Viola and Jones method):

The Viola and Jones is a method of object detection in digital image, proposed by Paul Viola and Michael Jones in 2001 [24]. It is one of the first methods that can efficiently detect objects in an image in real time. It allows finding multiple faces in an image with low processing times, Viola and Jones based their algorithm on numerous simple features and classifiers cascaded. To do this the method uses a Haar base for the extraction of characteristics and Ada-boost for the selection and classification of characteristics. The flowchart is as shown in Fig II.2.

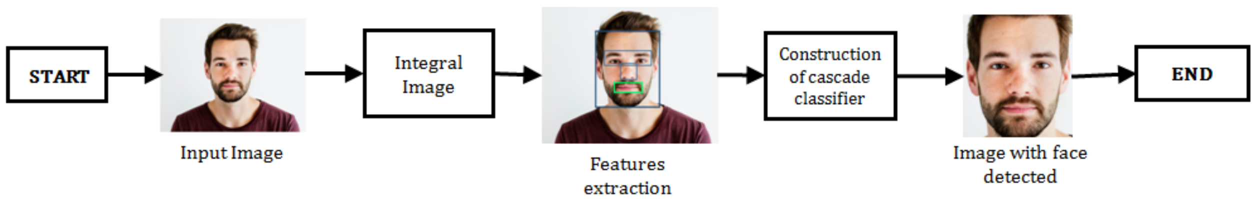

Figure 2.2: Flowchart of Face Detection (Viola and Jones).

1- **Integral Image:** Generates a new image representing the sum of pixels in rectangles "left-behind". Integral images are used to compute the projections in step 2 in a computationally effective manner.

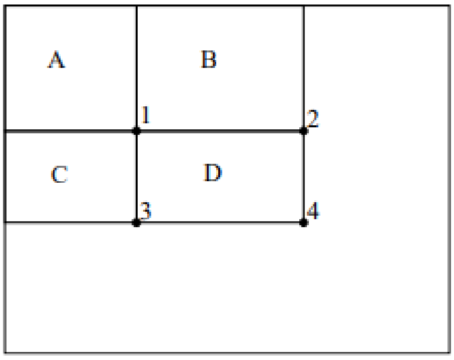

Figure 2.3: The sum of the pixels within rectangle D can becomputed with four array references.

2- **Feature extraction:** Local Haar feature images are computed by projecting the original image neighborhood on the Haar filters. These projections are features producing feature-images.

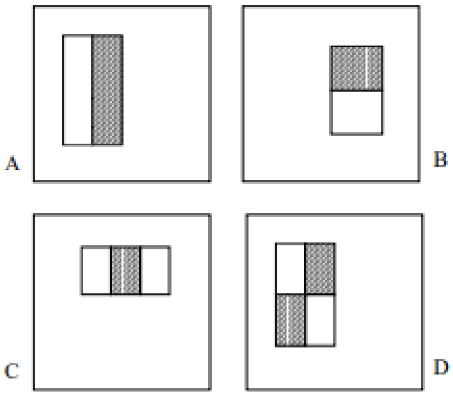

Figure 2.4: Example rectangle features shown relative to the enclosing detection window. Two-rectangle features are shown in (A) and (B). Figure (C) shows a three-rectangle feature, and (D) a four-rectangle feature.

3- **Construction of cascade classifiers:** Using AdaBoost a series of weak classifiers are constructed and combined in cascades.

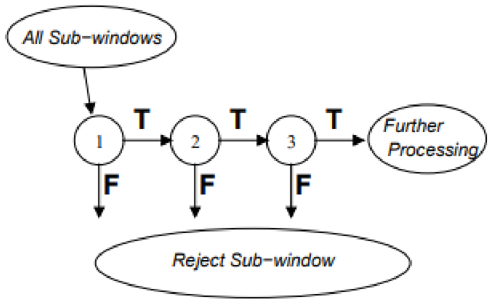

Figure 2.5: Schematic depiction of a the detection cascade.A series of classifiers are applied to every sub-window.

2.4.1.2 Normalization and enhanced the quality of image:

2.4.1.2.1 Multi-Scale Retinex (MSR): The multi scale retinex (MSR) algorithm is an extension of the single scale retinex (SSR) algorithm again proposed by Rahman et al., in 1996 [25]. MSR was developed to overcome the limitations of the SSR, cause if the dynamic range of a scene is much larger than that of image capturing device then, in such case the unrecoverable information loss can occur. MSR combines the quality of different surround space to provide a resultant image with good compression of both dynamic range and total rendition [26].

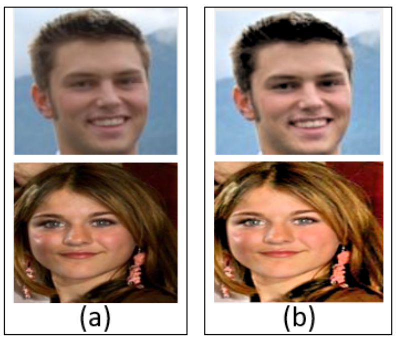

Figure 2.6: Sample image explains the results of using retinex Algorithm : a)original image - b)MSR image.

2.4.1.2.2 Normal Distribution Mapping (NDM): The Normal Distribution Mapping (NDM) provide enhanced 2D face recognition performance compared to unprocessed facial images [27]. The NDM normal curve is defined as follows:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{\frac{-(x-\mu)^2}{\sigma^2}} \tag{2.1}$$

Where $\sigma \succ 0$ stands for the standard deviation and μ stands for the mean value. μ and σ are two parameters that must be chosen when applying the histogram mapping approach with the normal distribution as the aim.

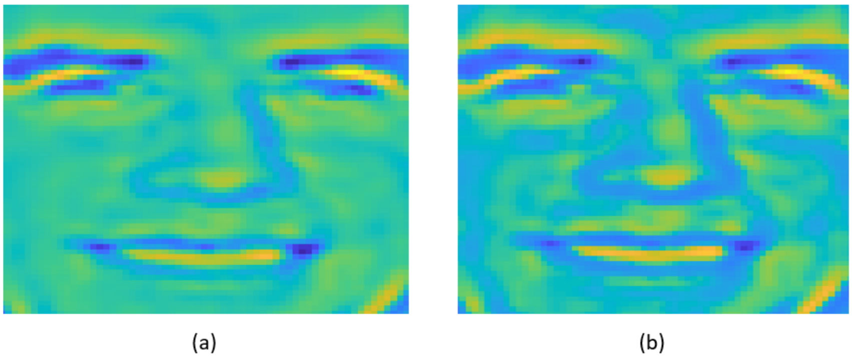

Figure 2.7: Sample image with scaled colors explains the results of using (NDM) Algorithm: a) input image b) output using NDM Algorithm.

2.4.2 Features Extraction:

We classify these methods into two categories: Those based on the extraction of Shallow features (Handcrafted) and those based on deep learning. Deep learning models were not widely applied in the field of automatic kinship verification due to insufficient data.

2.4.2.1 Shallow Features:

Shallow features have been used for over a decade in a number of computer vision applications including object detection and image classification. There are different descriptors that can be used for the parentage verification problem such as:

2.4.2.1.1 Local Phase Quantization (LPQ): The Local Phase Quantization (LPQ) operator was originally proposed by Ojansivu and Heikkila as a texture descriptor [28]. LPQ is based on the blur invariance property of the Fourier phase spectrum. It uses the local phase information extracted using the 2-D short-term Fourier transform (STFT) computed over a rectangular neighborhood at each pixel position of the image, as shown in the equation:

$$F_u(X) = \sum_{m \in N_x} h(m-x) f(m) \exp^{-2j\pi u^T m} \quad = E_u^T f_x \tag{2.2}$$

where, E_u of size $= 1M^2$, is a basic vector of 2DWFT with frequency u, and f_x, size= $M^2 * N$, is a vector containing the image pixel values in N_x at each position x. The window function, $h(x)$ is a rectangular function.

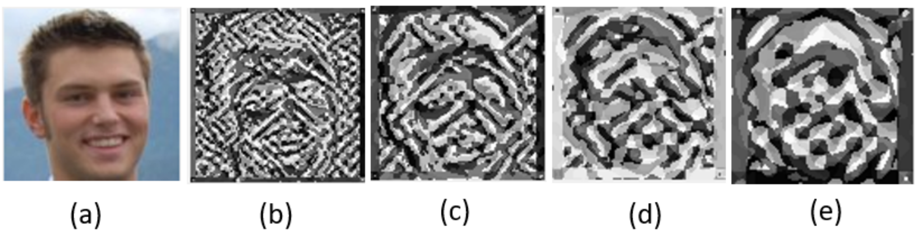

Figure 2.8: Sample images of LPQ descriptor with: a)Input Image, b-c-d-e) Image result of LPQ coefficients with rayon =3,5,7 and 9 respectively.

In LPQ only four complex coefficients are considered, corresponding to 2-D frequencies. In our experiments, we use the original code shared by the inventors of LPQ. The LPQ method can be summarized in four distinct steps. First, the (LPQ) operator is applied to the input image to obtain the labeled image. Then, the resulting image is divided into small regions. For each of them, a histogram of the labels is constructed in order to obtain vectors of the characteristics. The global representation (global feature vector that represents the entire image) is obtained by combining all the vectors.

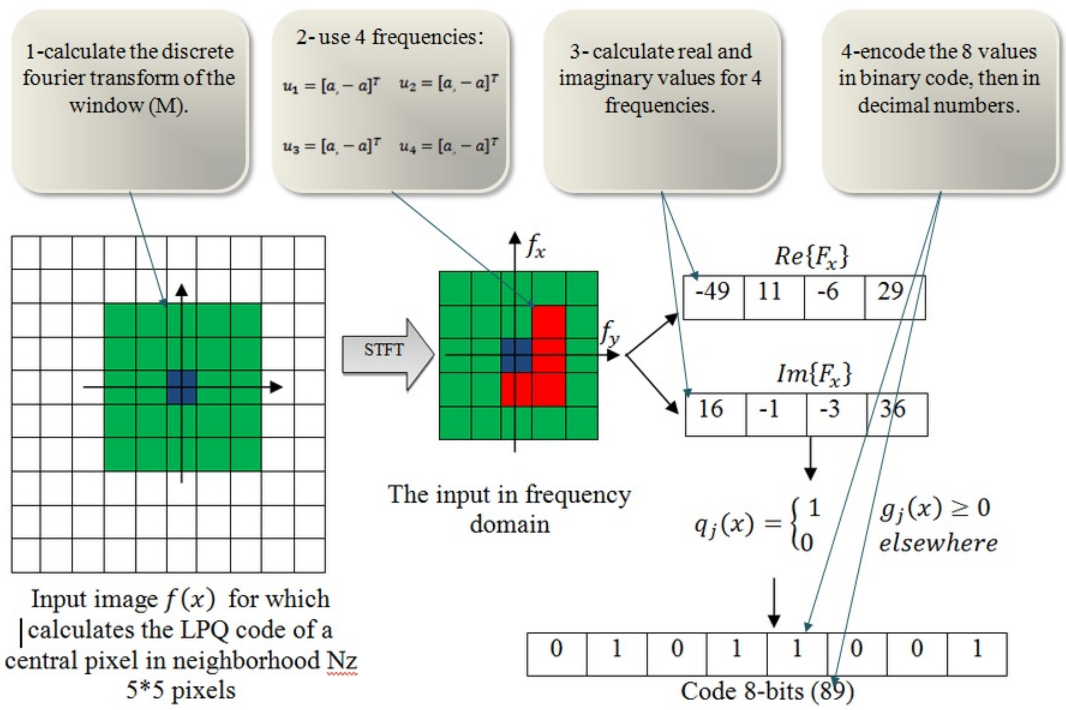

Figure 2.9: Flowchart of all the steps necessary for the generation of the vector of characteristics by the LPQ method.

2.4.2.1.2 Local binary pattern (LBP) Local binary pattern is powerful texture descriptor introduced in the 90s by ojala et.al.[29] describe the relationship of a pixel to its neighborhood. The idea of this texture operator is to give each pixel a code depending on the gray levels of its neighborhood. The gray level of the central pixel (i_c) is compared to those of its neighbors (i_n) according to the following formula:

$$\mathcal{LBP}(x_c, y_c) = \sum_{n=0}^{p} s(i_n - i_c)2^n \tag{2.3}$$

$$s(x) = \begin{cases} \mathbf{0} & \mathbf{if} \quad \mathbf{x} < 0 \\ \mathbf{1} & \mathbf{if} \quad x \geq 0 \end{cases} \tag{2.4}$$

LBP is used to calculate local texture feature of image and often used for texture classification problem. LBP work as, it describe the eight neighborhood pixel in binary code and summaries all code into histogram which serve as texture feature. In simple word, LBP label each pixel of an image with decimal number called local binary pattern or LBP code. Figure shows the example of basic LBP descriptor.

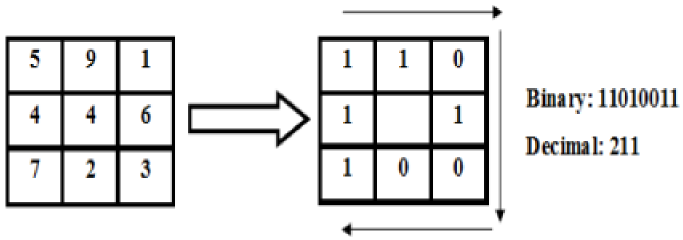

Figure 2.10: An example of basic LBP descriptor.

As shown in above figure each pixel is compare with its eight neighbors. The central pixel value is subtracting from its neighboring pixel value. The resulting negative value is encoded with 0 and other with 1. A binary number is obtained by concatenating all these binary code in a clockwise direction starting from the top left one and its corresponding decimal value is used for labeling the pixel. The 256-bin histogram of LBP labels is computed and then used as texture descriptor of an image [30].

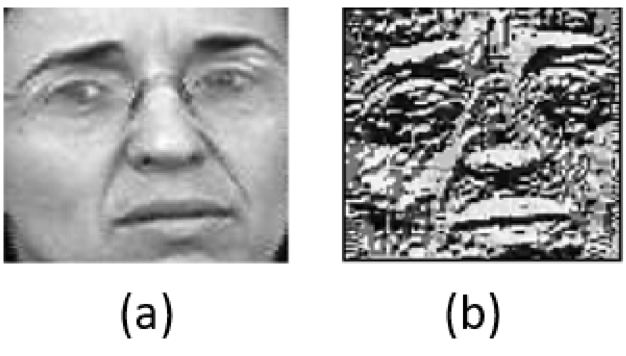

Figure 2.11: Sample image of Local Binary Pattern (LBP) descriptor: a) input image, b) result of LBP.

2.4.2.1.3 Binarized Statistical Image Features (BSIF) Unlike LBP and LPQ which can be used to calculate label statistics in local pixel neighborhoods, the local descriptor called BSIF (Binarized Statistical Image Features), which was recently proposed by Kannlaand Rahtu, uses a predefined set manually linear filters and binarization of filter responses [31].

Given a patch image X of size $l * l$ pixels and a linear filter W_i of the same size, the filter response Si obtained by:

$$s_i = \sum_{u,v} W_i(u,v)X(u,v) = w_i^T x \tag{2.5}$$

Given n linear filters W_i, we can stack them on a matrix W and calculate all the answers at once.
Where the vector notation is introduced in the last step, namely, the vectors w and x contain the pixels of W_i and X. The binarized function b_i is obtained by:

$$b_i = \begin{cases} 1 & if \quad s_i > 0 \\ 0 & otherwise \end{cases} \tag{2.6}$$

Where b_i is the ith element of b. In this way, an n-bit serial b binary code can be calculated for each pixel and subsequently the image region can be represented by binary code histograms of the pixel.

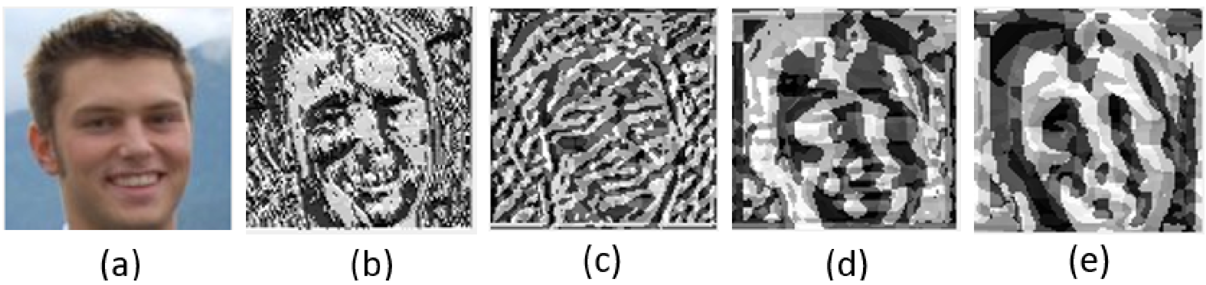

Figure 2.12: Sample images of Binarized statistical image features (BSIF) descriptor.With: a) Input Image, b-c-d-e)result of BSIF with filterof 3x3, 7x7, 11x11 and 15x15 dimensionrespectively.

2.4.2.1.4 Local optimal oriented patterns (LOOP) The Local optimal oriented patterns (LOOP) [32] as a recent texture descriptor encodes local structures and repeated local patterns of images. Compared to other commonly used feature descriptors, such as local directional pattern (LDP) and local binary pattern (LBP), the LOOP descriptor has demonstrated superior performance in numerous image recognition tasks. The LOOP descriptor possesses a notable advantage in capturing intricate local information from images, making it a suitable choice for applications such as facial recognition. To obtain the final codes, three steps must be taken. First, the edge responses of pixels with gray values $gi(i = 0, 1, ..., 7)$ in eight directions are determined using the Kirsch masks. Second, binarization weights wi are assigned to pixels based on the rank of the mask response value. Finally, the obtained weights are incorporated into the LBP formula to compute the final code relative to the center pixel.

$$LOOP(x_x, y_c) = \sum_{I=0}^{7} \xi(g_i - g_c) \cdot 2^w i \tag{2.7}$$

where, function ξ is expresses by

$$\xi(\eta) = \begin{cases} 1, & if\ \eta \geq 0 \\ 0, & if\ \eta < 0 \end{cases} \tag{2.8}$$

In this regard, g_c refers to the gray level of the center pixel located at (x_x, y_c)

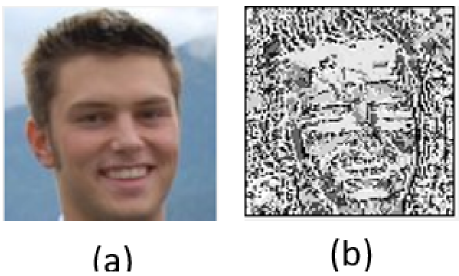

Figure 2.13: Sample images of Local optimal oriented patterns (LOOP) descriptor: a) input image, b) image of LOOP descriptor.

2.4.2.1.5 Two dimensionalDiscrete Wavelet Transform(2D DWT) Feature extraction is one of the most important parts of any face verification system. The feature extraction stage provides a feature vector or a matrix for each subject in the dataset. These features can be considered as the biometric signature of this subject. 2D DWT is a commonly used transform in image processing [33]. It converts an image from a spatial domain to a frequency domain. As a result, a 2D face image is decomposed into four sub-bands, also called scales. These are: approximate coefficient $\phi(a,b)$, horizontal coefficient$\psi H(a,b)$, vertical coefficient $\psi V(a,b)$ and diagonal coefficient $\psi D(a,b)$ [34].Defined by:

$$\phi(a,b) = \phi(b)\phi(b) \tag{2.9}$$

$$\psi H(a,b) = \psi(a)\phi(b) \tag{2.10}$$

$$\psi V(a,b) = \phi(a)\psi(b) \tag{2.11}$$

$$\psi D(a,b) = \psi(a)\psi(b) \tag{2.12}$$

The two-dimensional discrete scaling and translation basis functions are given by:

$$\phi_{j,m,n}(a,b) = 2^{\frac{j}{2}}\phi(2^j a - m, 2^j b - n) \tag{2.13}$$

$$\psi_{j,m,n}(a,b) = 2^{\frac{j}{2}}\psi^i(2^j a - m, 2^j b - n), i \in \{H, V, D\} \tag{2.14}$$

Where m, n are the translation quantities and j represents a scale.

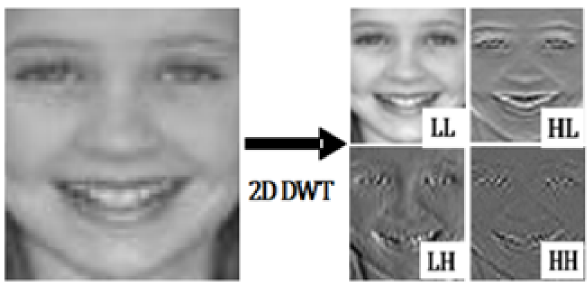

Figure 2.14: Different sub-bands after first decomposition level of 2D-DWT implementation on an image.

2.4.2.2 Deep Features based on Convolutional neural network (CNN)

Modern deep learning models in computer vision use convolutional neural networks, also called ConvNets or CNNs, introduced by LeCun [35],it is a particular type of deep neural networks specially designed to process input images. They are very popular nowadays because of their high performance. In recent years convolutional neural networks (CNNs) have revolutionized image processing and have achieved great success in various computer vision tasks, such as image classification, image segmentation, face recognition [36],[37],[38], object recognition [39]etc.

- CNN Architecture:

Over the last 10 years, several CNN architectures have been pressented [40],[41]. Model architecture is a critical factor in improving the performance of different applications. Various modifications have been achieved in CNN architecture from 1989 until today. Such modifications include structural reformulation, regularization, parameter optimizations, etc. Conversely, it should be noted that the key upgrade in CNN performance occurred largely due to the processing-unit reorganization, as well as the development of novel blocks.

Studying these architectures features (such as input size, depth, and robustness) is the key to help researchers to choose the suitable architecture for their target task. CNN architecture is composed of different layers such as convolution, pooling, ReLU correction layer and fully connected layers. In the most classical CNN, a convolution layer is followed by a pooling layer several times and then fully connected layers are added at the end. CNNs provide both feature extraction and classification.

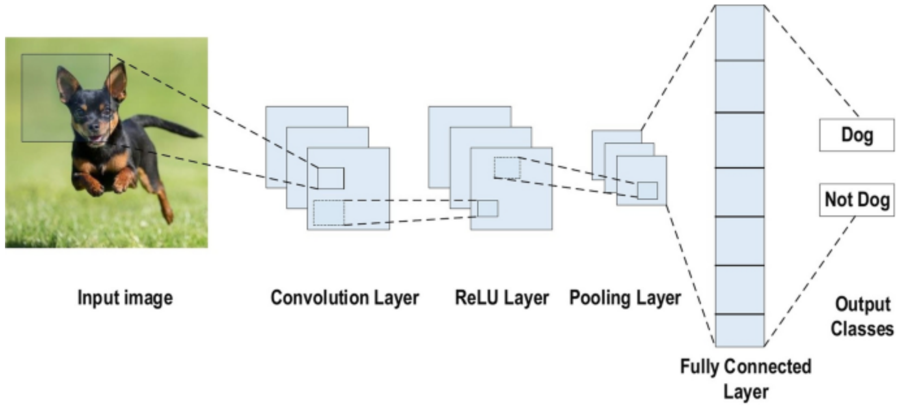

Figure 2.15: An example of CNN architecture for image classification.

1 **The Input layer:** Each image is of dimension $[W_i, H_i, C_i]$, where (W_i is its width in pixels), (H_i its height in pixels) and(C_i the number of channels), (1 for a grayscale image, 3 for a color image) [42].

2 **The convolution layer:** The convolution layer is the particularity of the CNN networks since it works as a feature extractor, they consist in dragging a convolution kernel on the image and manage to detect low level features like edges and curves, this is done with the help of filters. The kernels of the filters designate the weights of the convolution layer, unlike traditional methods, they are not pre-defined according to a particular formalism, but learned by the network during the training phase (initialized and then updated by the backpropagation of the gradient) [42].

The convolution layer has four hyperparameters [43]:

- **The number of filters K:** defines the depth of the output volume.
- **The size F of the filters:** each filter has dimensions $F \cdot F \cdot D$ pixels.
- **The step(Stride) S:** step used when convolving a filter through the image, controls the overlap between windows.
- **The zero-padding P:** we add to the input image of the layer a black outline of thickness P pixels.

Produces a matrix of dimension $[W_0, H_0, C_0]$:

$$W_0 = \frac{W_i - F + 2P}{S} + 1.H_0 = \frac{H_i - F + 2P}{S} + 1 \tag{2.15}$$

C_0 corresponds tp the number oh K filters.

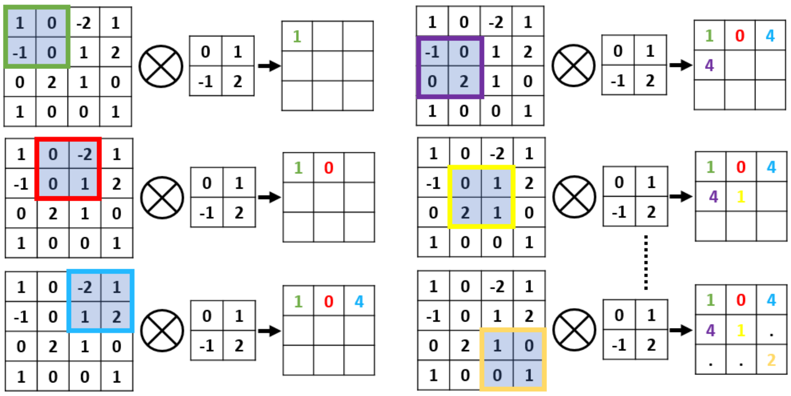

Figure 2.16: The primary calculations executed at each step of convolutional layer.

3 **The ReLu activation layer:** The mostly commonly used function in the CNN context. It converts the whole values of the input to positive numbers. Lower computational load is the main benefit of ReLU over the others. Its mathematical representation is in the following formula:

$$F_{ReLU}(x) = max(0, x) \tag{2.16}$$

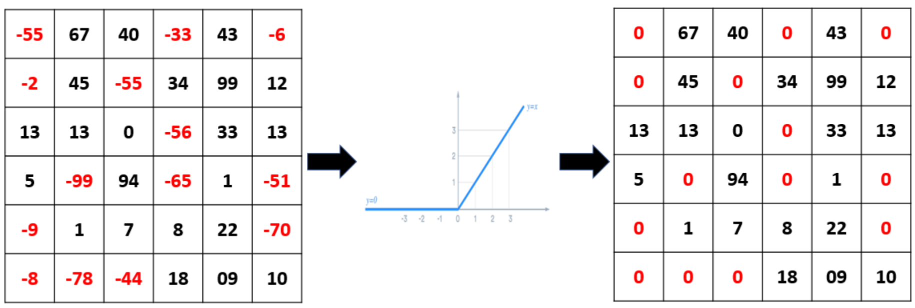

Figure 2.17: Activation function ReLU.

4 **Pooling layer:** The main task of the pooling layer is the sub-sampling of the feature maps. These maps are generated by following the convolutional operations. In other words, this approach shrinks large-size feature maps to create smaller feature maps. Concurrently, it maintains the majority of the dominant information (or features) in every step of the pooling stage. In a similar manner to the convolutional operation, both the stride and the kernel are initially size-assigned before the pooling operation is executed. Several types of pooling methods are available for utilization in various pooling layers. These methods include tree pooling, gated pooling, average pooling, min pooling, max pooling, global average pooling (GAP), and global max pooling. The most familiar and frequently utilized pooling methods are the max, min, and GAP pooling. The most common choice is Max-pooling with filters of size 22 pixels that do not overlap (stride=2). The pooling layer has only two hyperparameters[43].

- The size F of the patches: the image is sliced into square patches of size $F * F$ pixels.
- The pitch S: the patches are separated from each other by S pixels.

Accepts a volume of size [L1,H1,C1] and product matrix of dimensions $[W_2, H_2, C_2]$ where

$$W_2 = \frac{W_1 - F}{S} + 1, H_2 = \frac{H_1 - F}{S} + 1, C_2 = C_1 \tag{2.17}$$

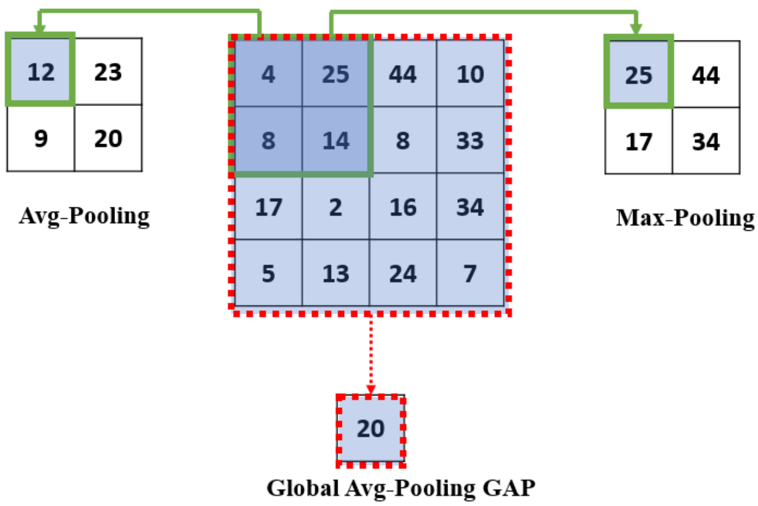

Figure 2.18: Three types of pooling operation.

5 **Fully connected layer:** Commonly, this layer is located at the end of each CNN architecture. Inside this layer, each neuron is connected to all neurons of the previous layer, the so-called Fully Connected (FC) approach. It is utilized as the CNN classifier. It follows the basic method of the conventional multiple-layer perceptron neural network, as it is a type of feed-forward ANN. The input of the FC layer comes from the last pooling or convolutional layer. This input is in the form of a vector, which is created from the feature maps after flattening. The output of the FC layer represents the final CNN output.

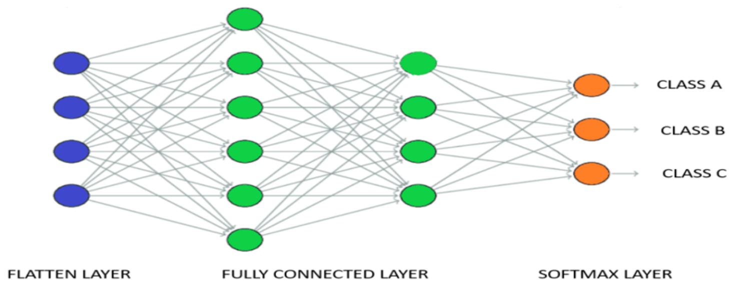

Figure 2.19: Fully connected layer.

- **Pretrained model:**

Deep features are extracted by convolutional neural networks (CNNs) which have significantly improved the state of the art in many applications, this success would not have been possible without the availability of the ImageNet organisation of large labelled datasets which have been used to learn high quality features [44]. We cite some deep learning models that can be used :

VGG CNN : A CNN model among the most known feature extractors can be used. This network is revolutionary in its inherent simplicity and structure, it uses a very deep architecture with very small convolutional kernels (33) with a stride of 1 pixel, this model has been trained on 2.6 million images of 2622 different celebrities [45].

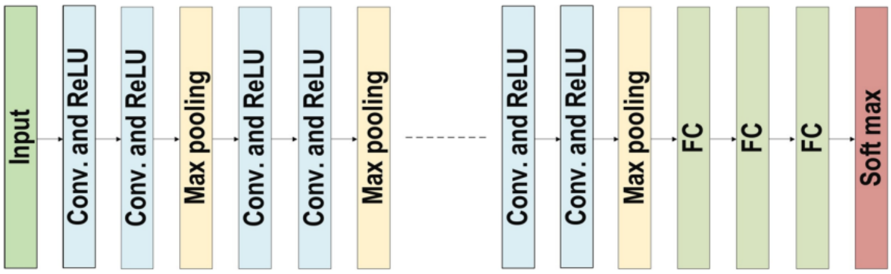

Figure 2.20: Fully connected layer.

FaceNet : A neural network for face recognition, verification and clustering. Uses an architecture with 22 layers. During training, the deep network extracts and learns various facial features, these features are then converted directly into 128D embeddings [36].

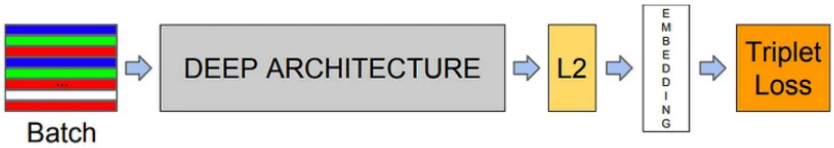

Figure 2.21: High Level Modal Structure ofFaceNet.

Resnet : The abbreviation for residual network. It is a type of deep neural network with a different network topology than VGG, the main idea of this network is the residual block. The network allows the development of extremely deep neural networks, which can contain 100 or more layers in order to extract the best possible features, with several architecture: resnet-22, resnet-50, resnet101, resnet-150, etc...[46].

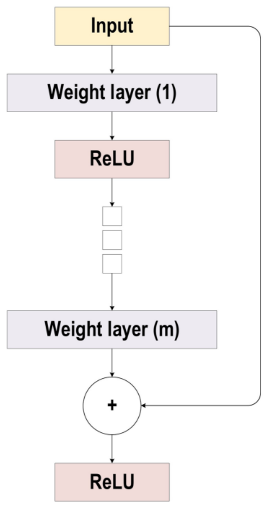

Figure 2.22: The block diagram for ResNet.

All these approaches extract facial features and form a kinship verification classifier, where facial similarity is represented by the difference between facial features. Extracted features include facial colour, position and shape of facial parts [47].

2.4.3 Tensor Design and compensation for the effect of variability

The training 3^{rd} order tensors $X, Y \in RI_1 * I_2 * I_3$ are constructed using the histograms of different feature descriptors extracted from the training face images. The three modes of the tensors X and Y are defined as follows: I_1 corresponds to the feature descriptors extracted at different scales, I_2 represents the histograms, and I_3 face samples in the database.

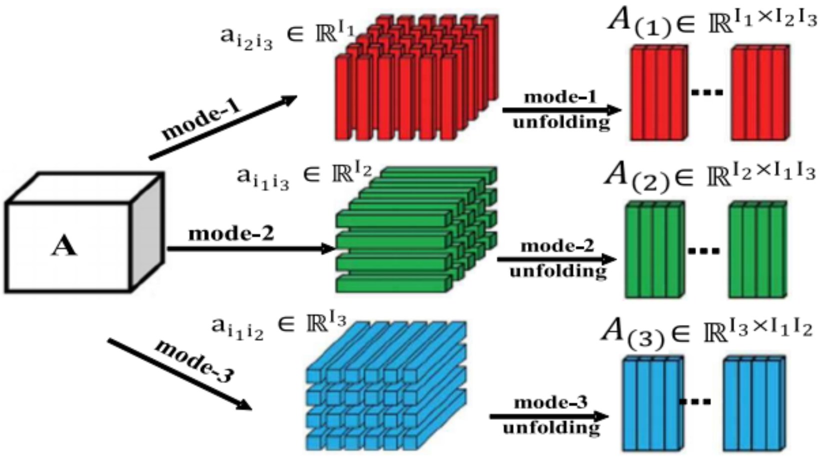

Figure 2.23: Example of tensor unfolding.

2.4.3.1 Dimensionality reduction stage

main problem of the vector is its high dimensionality. Therefore, it is required to project the feature vector into a lower space that contains only discriminant information. The dimensionality reduction stage stimulates the efficiency of the calculation of the recognition system and avoids technical problems, such as the curse of dimensionality. Dimensionality reduction can circumvent this problem by reducing the number of features in the data set before the training process. This can also reduce the computation time, and the resulting classifiers take less space to store. The main drawback of dimensionality reduction is the possibility of information loss. When done poorly, dimensionality reduction can discard useful instead of irrelevant information. No matter what subsequent processing is to be performed, there is no way to recover this information loss[48].

There are many dimensionality reduction methods used in data science for different types of applications namely Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Side Information Discriminant Analysis (SIDA), Multilinear PCA (MPCA), Multilinear SIDA (MSIDA), Cross-View Quadratic Discriminant Analysis XQDA ... etc.

2.4.3.2 Cross-view Quadratic Discriminant Analysis (XQDA)

The XQDA (Cross-view Quadratic Discriminant Analysis) is a method used in kinship verification to analyze and compare facial images of individuals to determine their degree of relatedness[49]. The XQDA method is a variation of Quadratic Discriminant Analysis (QDA), which is a statistical technique used for classification and prediction. XQDA extends QDA by incorporating a metric learning approach, which learns a transformation of the data to a new feature space that maximizes the class separation while minimizing the within-class variation. In kinship verification, XQDA is used to learn a metric that maps facial features to a space where kinship verification can be performed. This allows for accurate classification of whether two individuals are related or unrelated based on their facial features. XQDA has been shown to be effective in kinship verification tasks, achieving high accuracy rates in various studies.

2.4.3.3 Within-Class Covariance Normalization (WCCN)

This method is used as an additional session variability compensation technique to scale the subspace to reduce the dimension of high intra-class variance[50].The WCCN matrix (B) is calculated using the following Cholesky decomposition:

$$BB^T = W^{-1} \tag{2.18}$$

Where the intra-class covariance matrix W is calculated using :

$$W = \frac{1}{s}\sum_{s=1}^{s}\sum_{i=1}^{n_s}(A^T(w_i^s - \bar{w}_s))(A^T(w_i^s - \bar{w}_s))^T \tag{2.19}$$

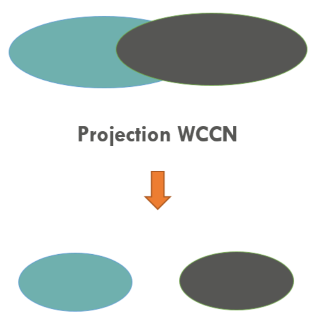

Figure 2.24: Normalization of the intra-class covariance.

2.4.4 Decision

The decision task in verification systems is to determine whether a given example belongs to a specific class or not. To make this decision, verification systems often use a decision threshold. The decision threshold is a value that is used to distinguish between positive and negative examples. If the value of the similarity measure or probability exceeds the decision threshold the example is considered positive, otherwise it is considered negative.

The final stage aims to finally verify the kin relations between data samples using various classifiers like Support Vector Machine (SVM)[51], or K-Nearest Neighbor (KNN)[52].

2.5 Metric Learning

The choice of the right metric is crucial when evaluating machine learning models.Various metrics are proposed for the evaluation of models, concerning our work we use the following metrics:

Confusion matrix: Summary of prediction results on a classification problem. The correct and incorrect predictions are highlighted and divided by class. The results are then compared with the real values. This matrix allows to measure the quality of a classification system

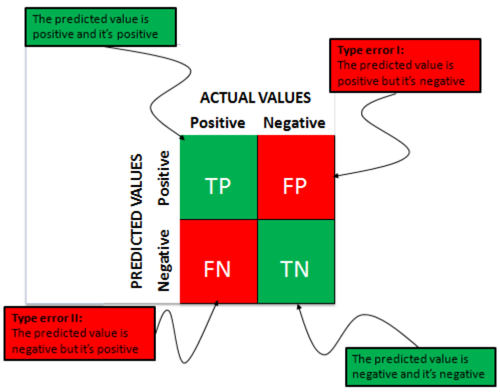

Figure 2.25: Confusion Matrix.

From this matrix, many measures can be extracted such as: the global accuracy, the precision, the AUC-ROC curve...etc.

Precision: Corresponds to the quality of the class. We divide the number of well ranked items in the class by the total number of items assigned to the class. For the classes (KIN and NON-KIN) , we calculate the precision as follows:

$$Precision_{KIN} = \frac{TP}{TP + FP} \tag{2.20}$$

$$Precision_{NON-KIN} = \frac{TN}{TN + FN} \tag{2.21}$$

Global accuracy: Is the ratio of the correctly classified samples to the total number of classified samples. this measure is used to evaluate the overall performance of a proposed model.

$$acc_i = \frac{TP + TN}{TP + FP + TN + FN} \tag{2.22}$$

Then each global accuracy will be multiplied by its respective weight. according to the following formula: for each relationship i :

$$weight_i = \frac{number\ of\ parirs\ of\ ther\ relation}{total\ number\ of\ pairs} \tag{2.23}$$

Sensitivity /recall: is the ratio of the true-positive samples to all infected samples (true-positive and false-negative). This measure is used to evaluate the performance of a proposed model in predicting true-positive cases [53], [54].

$$Sensitivity = \frac{TP}{TP + FN} \tag{2.24}$$

Specificity: is the ratio of the true-negative samples to all healthy samples (true-negative and false-positive). This measure is used to evaluate the performance of a proposed model in predicting true-negative cases.

$$Specifity = \frac{TN}{TN + FP} \tag{2.25}$$

F1 Score: is the consistency mean of sensitivity and precision, in the case where the imbalance of false positive/negative samples is important to be measured.

$$F1Score = \frac{2 \ * \ (Sensitivity \ * \ Precision)}{(Sensitivity + Precision)} \tag{2.26}$$

Coefficient of Quartile Deviation: measures the variability of among the image samples themselves and around the average. low coefficient value means low dispersion. Whereas, $\mathbf{Q}_3$ represents the observations that have upper quartile, $\mathbf{Q}_1$ represents the observations that have lower quartile [55].

$$\mathbf{QCoD} = \frac{(\mathbf{Q}_3 - \mathbf{Q}_1)}{(\mathbf{Q}_3 - \mathbf{Q}_1)} \tag{2.27}$$

ROC curve: The ROC curve (ROC stands for "receiver operating characteristic," the term comes from radar engineering). It is a graph representing the performance of a classification model for all thresholds. This curve plots the true positive rate against the false positive rate:

True positive rate (TPR) defined as follows:

$$TPR = \frac{TP}{TP + FN} \tag{2.28}$$

The false positive rate (FPR) is defined as follows:

$$FPR = \frac{TN}{TN + FP} \tag{2.29}$$

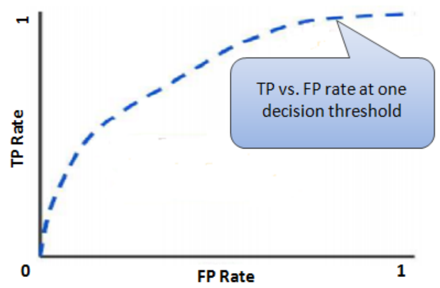

Figure 2.26: Classic ROC curve.

AUC: AUC stands for "area under the ROC curve". This value measures the entire two dimensional area under the entire ROC curve (by integral calculations) from (0.0) to (1.1). The AUC provides an aggregate measure of performance for all possible classification thresholds.

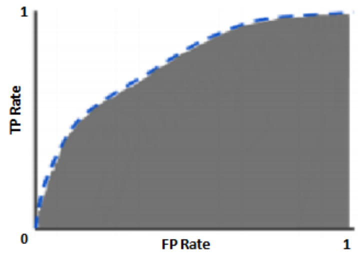

Figure 2.27: AUC (area under the ROC curve).

2.6 Conclusion

The objective of this chapter was to give a global overview on the solutions realized so far for automatic kinship verification, we started by mentioning the different public databases available and defining several approaches that have been used by researchers in this field.As the first title, we saw the methods of face preprocessing and feature extraction and we closed the chapter by defining the feature reduction methods, the learning and classification algorithms.

In the next chapter we will discuss our system design and the proposed approaches that we will use.

Chapter 3

Proposed solution

3.1 Introduction

Currently deep learning models can achieve human-scale precision in image analysis and segmentation. Motivated by the impressive success of deep learning approaches in the representation and classification of various images, we have proposed a contribution for the automatic verification of kinship links, our work consists of combine deep features and shallow features.

In this chapter we are going to detail the various stages useful for the implementation of our automatic system of kinship verification.

3.2 Proposed solution

The importance of robust facial features in identifying and verifying relationships between individuals is widely acknowledged. In image classification tasks, the quality of the representational encoding of images is a key factor that affects the effectiveness of the approach. These encodings can be local textural details or learned features. We propose a novel approach to extract efficient and discriminative features from face images by leveraging prior knowledge and fusing deep features with shallow features using powerfull technique (LR fusion), such as VGG-16, ResNet-50, LPQ, and our proposed descriptor named Hist-2D DWT. By combining these features, we can extract complementary information that aids in determining kin relations .

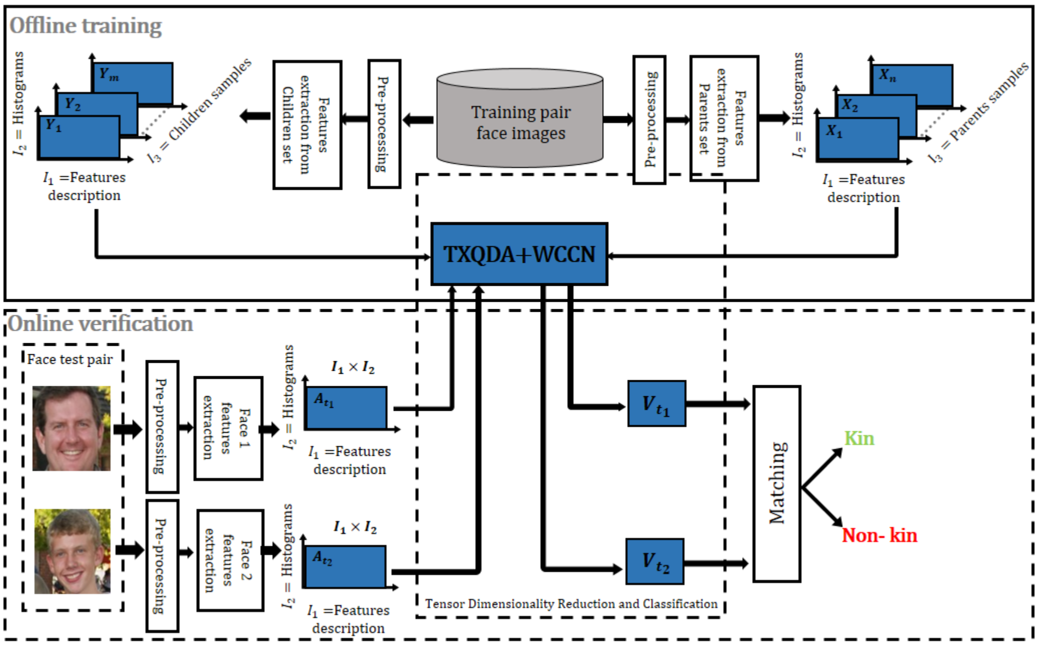

Figure 3.1: The proposed pipeline for face kinship verification system.

3.2.1 Face preprocessing using MSR+NDM

We applied the Multiscale Retinex (MSR) [25], as an image enhancement technique to enhance the quality of digital images by increasing their dynamic range, and preserving their color accuracy. Next, we used the Viola and Jones [24] method for detected the face region, followed by the utilization of the Normal Distribution Mapping (NDM). The NDM algorithm is known for its effectiveness in extracting illumination-insensitive features for face images under varying lighting conditions, as well as its ability to mitigate the effects of image noise.

3.2.2 Features extraction using deep and shallow features

- **Deep features:** For the deep features, we extract them from the original face image with a size of 224 ∗ 224 ∗ 3. We use four layers of the VGG-16 network, namely "fc6, relu6, fc7, and relu7" and one layer of the Resnet50 called "fc1000", to extract features.

- **Shallow Features:** To extract shallow features, we employ the LPQ, and the proposed descriptor Hist-2D DWT on the facial image, which is then partitioned into 12 blocks. Each block is summed into a histogram comprising 256 bins, and the resulting histograms are concatenated to form a feature vector .

- **The proposed Hist-2D DWT descriptor:** To improve the performance of the proposed kinship verification, a new discriminant face descriptor called Hist-2D DWT is presented. After extracting the four coefficients LL, HL, LH, and HH images using 2D DWT, each component is divided into k sub-blocs. Each of these sub-blocs is represented by a histogram of 256 bins. Then, to characterize the feature matrix, a single vector with a size of $ktimes256$ is created by concatenating all the histograms of all sub-blocs. Finally, these feature vectors are inserted into a new matrix to form the Hist-2D DWT face descriptor. The Hist-2D-DWT face feature extraction process is illustrated in Fig III.2.

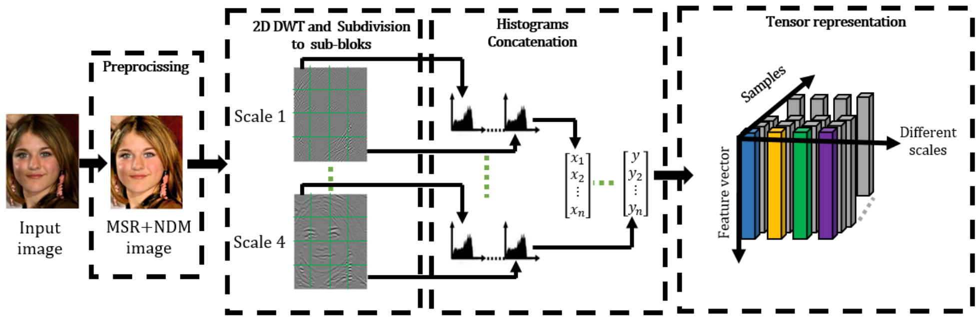

Figure 3.2: Feature extraction using Hist-2D DWT.

3.2.3 Multilinear subspace learning and dimensionality reduction based TXQDA+WCCN

In our work we use a new dimensionality reduction technique developed for kinship verification called Tensor Cross-view Quadratic Discriminant Analysis (TXQDA) proposed by Laiadi et al. in [56]. Which was developed over the XQDA method. this new technique had achieved better results than the previous ones in dimensionality reduction for kinship verification datasets. The organigram for Tensor Cross-view Quadratic Discriminant Analysis method is proceeds in Fig III.3. In order to bolster method, we leverage the within-class covariance normalization (WCCN) technique [57] for feature learning, which minimizes the expected variances between training features of the same class. As a result, we arrive at an optimized version of the method, referred to as TXQDA+WCCN.

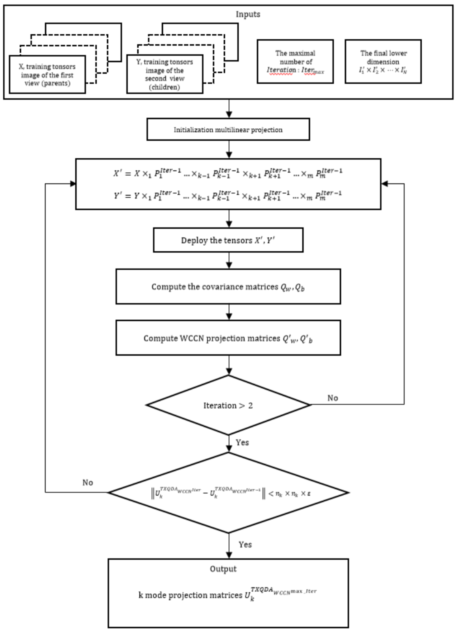

Figure 3.3: Flowchart of the TXQDA+WCCN approach.

3.2.4 Matching and logistic regression fusion

We utilized the reduced features projected onto the TXQDA+WCCN subspace, which were concatenated to form a single feature vector, for comparing pairs of facial images. For each testing pair consisting of two facial images, we applied cosine similarity [58] the cosine similarity between two vectors, $V_(t_1)$ and $V_(t_2)$, is defined by the Eq III.1.

$$cos(V_{t1}, V_{t2}) = \frac{v_{t1}^T, v_{t2}}{||v_{t1}||.||v_{t2}||} \tag{3.1}$$

Our system implements a score-level fusion strategy to leverage the complementary nature of two distinct types of similarity scores, thereby improving overall accuracy. Specifically, we use Logistic Regression (LR) [59] within our framework. We chose this method as it has demonstrated significant improvement in previous fusion studies. Given the input scores a_i the output probability b_i is defined by the Eq III.2:

$$b_i = (1 + \exp(xa_i + y))^{-1} \tag{3.2}$$

Where, x repesents a scalar factor, while y denotes a bias term.

3.3 Conclusion

This chapter gave a clear vision on our work by showing our contribution through the conceptual aspect of our system and the stages established in order to concretize this last. We have described our proposed solution by explaining precisely the architectures we used to build our final system.

The next chapter will present the tests carried out, and the results obtained as well as their interpretations.

Chapter 4

Results and discussions

4.1 Introduction

Computer vision researchers around the world are always trying to optimize the performance of kinship verification systems with trying and developing different approaches and methods.
In this last chapter, we are going to implement our kinship verification system using four different types of parameter extraction: shallow features using Hist-2D DWT and LPQ descriptors ,deep features using VGG16 and ResNet50.In order to get the best performance scorefor a kinship verification system, we will conduct a series of experiments on two kinship datasets (Cornell and TS) with various settings and parameters. We will use a new dimensionality reduction technique called Tensor Cross-View Quadratic Discriminant Analysis.we will do this by employing LR Fusion to merge each type's top score.

4.2 Environment used

The programming language used in this work is MATLAB, emulated by the programming environment of the same name (in our case MATLAB 2021a) and developed by The Math Works. Matlab allows for the simple and fast implementation of algorithms, the implementation of tasks requiring high computing power, the manipulation and display of curves, and the creation of graphical interfaces. The experiments were carried out on a PC with an 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz with 16 GB of RAM. We used Deep Learning Toolbox Model[60] environment for CNNs.

4.3 Working protocol

We use Cross-Validation or K-fold Cross Validation method, the model is trained and tested in the same time. Cross-validation is primarily used in applied machine learning to estimate the skill of a machine learning model on unseen data. That is, to use a limited sample in order to estimate how the model is expected to perform in general when used to make predictions on data not used during the training of the model. It is a popular method because it is simple to understand and because it generally results in a less biased or less optimistic estimate of the model skill than other methods, such as a simple train/test split. The process of testing is determined by the chosen dataset's protocol. Fig IV.1illustrates the cross validation process using 5 folds, hence it's named 5-fold cross validation [[2]- [9]].

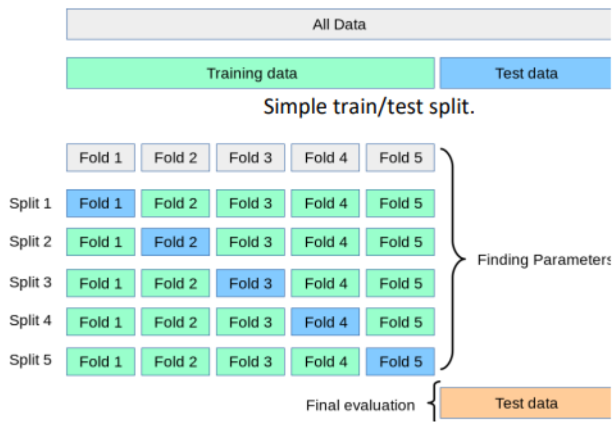

Figure 4.1: K-fold Cross Validation, with K = 5.

4.4 Experiments and results

We do several experiments to assess the proposed method:

a. For shallow features we use two descriptors (Hist-2D DWT and LPQ)

1. The number of bloc histograms is chosen to: 4,8,12,16,20.
2. After experimenting the results with Hist-2D DWT Coefficients_LL the best performing number of blocs is fixed, and coefficients are changed iteratively to: LL,LH,HL,HH.

   3. We set the number of blocs at 12 for LPQ and the rayon scale is changed iteratively to:3,4,5,6,7,8,9.

b. For deep features we use two pretrained models (VGG16 and ResNet50)

   1. VGG16 there are only 4 scales of features to be extracted and trained with (fc6, relu6, fc7, and relu7).

   2. ResNet50 there is 1 scale of features to be extracted and trained with (fc1000).

c. The best setting in Hist-2D DWT and LPQ is fixed and fused with the VGG16 and ResNet50 score.

4.4.1 Experiments on Cornell kinface database

Here we present the results of our experiments on Cornell kinface dataset, in which Tab IV.1 illustrate the accuracy of inserting the original images to our system with/without histogram,we didn't do any parameter extraction. Tab IV.2 shows the results of using Hist-2D DWT descriptor with and without prepoccessing. Tab IV.3 illustrates the results of bloc number histograms variations which is chosen to: 4,8,12,16,20 in order to fix it at the best accuracy. Tab VI.4 illustrated the mean accuracy of Hist-2D DWT with his different coefficients (LL,LH,HL,HH) and LPQ approach which we variate the coefficient rayon (R=3,4,5,6,7,8,9) with presprocessing inwhich we fixed the bloc number of histograms at 12. After that the experiments were conducted with the extraction of deep features from original images using VGG16 and ResNet50 the results are illustrated in Tab IV.5. Finally, the fusion at score level between the best previous performing results illustrated in Tab IV.6. On the other hand, the ROC (Receiver Operating Characteristics) is illustrated in Fig IV.1. showing us the results of different settings.

Settings	**Mean accuracy (%)**
Without histogram	54.51
With histogram	**59.34**

Table 4.1: The mean accuracy (%) using originale image with and without histogram in cornell kinface dataset.

Settings	**Mean accuracy (%)**
Without MSR+NDM	76.58
With MSR+NDM	**92.01**

Table 4.2: The mean accuracy (%) using Hist-2D DWT LL with and without preprocessing in cornell kinface dataset.

Bloc number	**Mean accuracy(%)**
4	77.89
8	88.89
12	**92.01**
16	86.05
20	87.78

Table 4.3: The mean accuracy (%), on Cornell Kin Face dataset using Hist-2D DWT LL and different blocs.

Shallow Features	**Scales**	**Mean accuracy (%)**
Hist-2DDWT	**Coefficients_LL**	**92.01**
	Coefficients_LH	84.62
	Coefficients_HL	79.36
	Coefficients_HH	79.03
LPQ	**R=3**	**94.16**
	R=4	92.72
	R=5	93.43
	R=6	93.39
	R=7	92.06
	R=8	93.09
	R=9	92.78

Table 4.4: The mean accuracy (%) for Sallow features on Cornell KinFace dataset.

Deep Features	**Mean accuracy (%)**
VGG16	91.02
ResNet 50	73.15

Table 4.5: The mean accuracy (%) for Deep features on Cornell Kin Face dataset.

Fusion	**Mean accuracy (%)**
LR fusion method	**95.18**

Table 4.6: Fusion of best performing features for Hist-2D DWT (LL) and LPQ (rayon 3), and deep features(VGG16 and ResNet50) on Cornell Kin Face dataset.

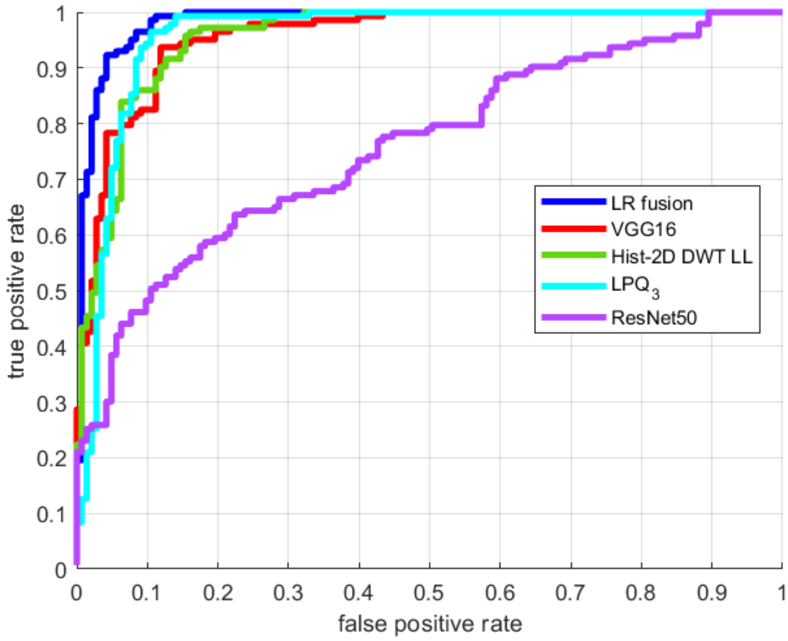

Figure 4.2: ROC curves of Shallow features and Deep features on Cornell KinFace dataset.

4.4.2 Experiments on TS kinFace database:

In this section we present the results of our experiments on TS kinface dataset, in which Tab IV.7 illustrate the accuracy of inserting the original images to our system with/without histogram,we didn't do any parameter extraction. Tab IV.8 shows the results of using Hist-2D DWT descriptor with and without prepoccessing. Tab IV.9 illustrates the results of bloc number histograms variations which is chosen to: 4,8,12,16,20. Tab VI.10 illustrated the mean accuracy of Hist-2D DWT with his different coefficients (LL,LH,HL,HH) and LPQ approach which we variate the coefficient rayon (R=3,4,5,6,7,8,9) with presprocessing in which we fixed the bloc number of histograms at 12. After that the experiments were conducted with the extraction of deep features from original images using VGG16 and ResNet50 the results are illustrated in Tab IV.11. Finally, the fusion at score level between the best previous performing results illustrated in Tab IV.12. On the other hand, the ROC (Receiver Operating Characteristics) is illustrated in Fig IV.2 of various settings for different relations.

The relations are as follows **FD** : Father-Daughter, **FS** : Father-Son, **MD** : Mother-Daughter, **MS** : Mother-Son.

Settings	FD	FS	MD	MS	Mean accuracy (%)
Without histogram	54.26	54.46	52.97	53.76	53.86
With histogram	**64.13**	**67.09**	**66.97**	**66.91**	**66.27**

Table 4.7: The mean accuracy (%) using original image with and without histogram in TS kinface dataset.

Settings	FD	FS	MD	MS	Mean accuracy (%)
Without MSR+NDM	77.03	79.41	81.98	81.68	80.03
With MSR+NDM	**87.13**	**85.14**	**88.02**	**86.83**	**86.78**

Table 4.8: The mean accuracy (%) using Hist-2D DWT LL, with and without preprocessing in TS kinface dataset.

Bloc number	**FD**	**FS**	**MD**	**MS**	**Mean accuracy (%)**
4	70.45	71.43	74.13	73.93	72.48
8	78.14	80.02	82.42	82.98	80.89
12	87.13	**85.14**	88.02	**86.83**	**86.78**
16	**87.27**	84.90	**89.29**	86.25	**86.93**
20	88.18	87.17	85.34	85.02	86.42

Table 4.9: The mean accuracy (%), on TS Kin Face dataset using Hist-2D DWT LL and different blocs.

Shallow Features	**Scales**	**FD**	**FS**	**MD**	**MS**	**Mean accuracy (%)**
Hist-2DDWT	**Coefficients_LL**	**92.01**	**85.14**	**88.02**	**86.83**	**86.78**
	Coefficients_LH	79.31	78.61	80.99	78.02	79.23
	Coefficients_HL	74.46	72.87	75.84	74.85	74.51
	Coefficients_HH	76.04	75.94	79.01	78.02	77.25
LPQ	R=3	86.93	85.94	88.22	87.23	87.08
	R=4	87.43	86.13	87.81	86.92	87.07
	R=5	86.83	86.24	86.83	86.53	86.60
	R=6	**87.43**	**86.63**	**88.91**	**87.52**	**87.62**
	R=7	85.14	85.01	87.97	86.26	86.09
	R=8	84.93	82.89	84.91	85.13	84.46
	R=9	85.36	82.34	83.56	84.50	83.94

Table 4.10: The mean accuracy (%) for Sallow features on TS Kin Face dataset.

Deep Features	**FD**	**FS**	**MD**	**MS**	**Mean accuracy (%)**
VGG16	78.12	77.38	79.70	79.32	78.63
ResNet 50	71.49	71.19	74.85	73.07	72.65

Table 4.11: The mean accuracy (%) for Deep features on TS Kin Face dataset.

Fusion	**FD**	**FS**	**MD**	**MS**	**Mean accuracy (%)**
LR fusion Method	90.30	91.49	93.17	92.28	**91.81**

Table 4.12: Fusion of best performing features for Hist-2D DWT (LL) and LPQ (rayon=3), and deep features(VGG16 and ResNet50) in TS KinFace dataset.

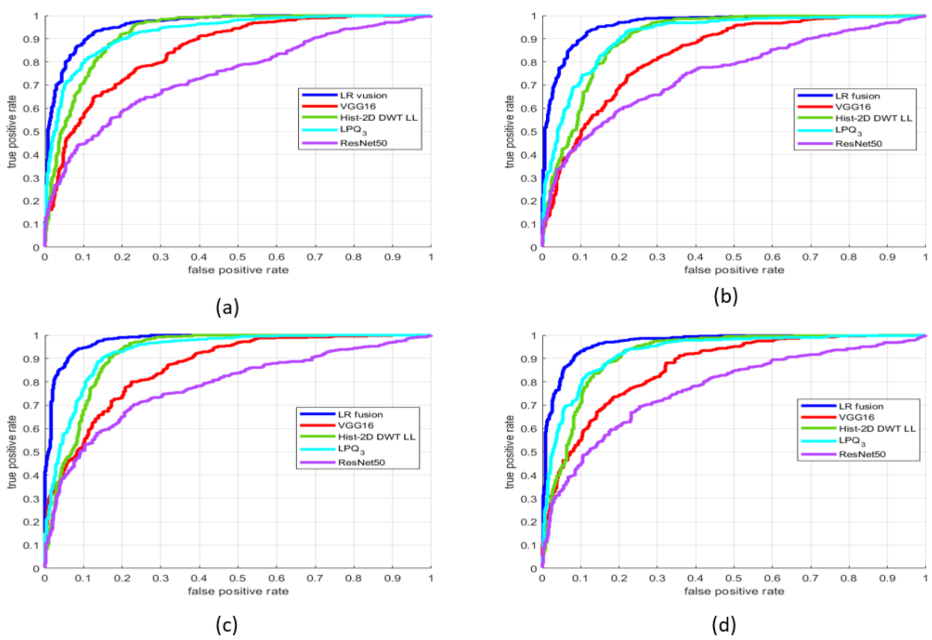

Figure 4.3: ROC curves of Shallow features and Deep features on TS KinFace dataset, (a) FD set, (b) FS set, (c) MD set and (d) MS set.

4.5 Discussion

The experiments that we conducted using the proposed approach TXQDA based LR fusion of four types of features (Hist-2D DWT, LPQ, VGG16, ResNet50) on two datasets (Cornell KinFace and TSKinFace),We conclude the following:

1. **Influence of the image histogram:** In order to get more precision for the metricevaluation, histogram is an essentialfactor that will increase the accuracy, for that our system gained **4.83%** for Cornell dataset and **12.41%** for TS dataset.

2. **The benefit of the preprocessing based MSR+NDM:** we found that preprocessing added **15.43%** and **6.75%** in the accuracy for Cornell and TS KinFace datasets respectively.

3. **Effect of bloc number variation:** For the variation of bloc number, we found that it affects the accuracy of our proposed descriptor but we fixed it at 12 blocs which was the best performing number for Cornelle KinFace dataset **(92.01%).**

4. **The powerful of Shallow features:** We applied two descriptors Hist-2D DWT and LPQ, for the proposed descriptor we got the best accuracy **92.01%,86.78%** in LL coefficient and for LPQ the best accuracy **94.16%** (R=3), **87.62%** (R=6) with Cornelle KinFace and TS KinFace respectively.

5. **The powerful of Deep features:** We employed two pretrained CNN Models (VGG16,ResNet50), after the simulation we found that the results were not as good as shallow features results, like for VGG16 the accuracy was **91.02%** and **78.63%**, for ResNet50 was **73.15%** and **72.65%** with Cornelle KinFace and TS KinFace respectively. But we know that features extracted are not the same so we keep them to aim the complementarity.

6. **The power of LR fusion at the score level for two types of features:** We use the LR approach to combine the top scores produced by different feature types (Hist-2D DWT, LPQ, VGG16 and ResNet50) at the score level. It is impressive that the LR fusion method performs well, scoring **95.18%** and **91.81%** in with Cornell KinFace and TSKinFace, respectively (see Tab IV.6 and IV.12) The Receiving Operating Characteristic (ROC) Curves of several approaches are presented in

Fig IV.1 and Fig IV.2 using Cornell KinFace and TSKinFace, respectively, to better illustrate the performances of various characteristics. These numbers demonstrate that LR fusion may significantly boost accuracy.

4.6 Comparison with state-of-the-art methods

The proposed method most effective results Hist-2D DWT, LPQ, VGG16 and ResNet50 score fusion using TXQDA is compared to more modern methods in **Tab IV.13** for Cornell KinFace and TS KinFace databases. The related works are cited according to the algorithm used. The comparison demonstrates that, on the two datasets, Cornell KinFace and TSKinFace, our proposed technique exceeds the recent state of the art.

Work	**Year**	**Algorithm**	**Cornell Dataset**	**TS Dataset**
Bessaoudi & al [56]	2019	MSIDA	86.87	85.18
Goyal & Meenpal [57]	2021	FMRE2	84.16	90.85
Zhang& al [58]	2021	AdvKin	81.40	-
Mukherjee& Meenpal [59]	2022	BC2DA	83.07	83.55
Serraoui& al [60]	2022	TXQEDA	93.77	90.68
Proposed	**2023**	**Fusion**	**95.18**	**91.81**

Table 4.13: Performance comparison (verification accuracy %) of kinship verification state of the art on Cornell KinFace and TS KinFace datasets.

4.7 Conclusion

In this last chapter, we have presented the results given by the implementation of our architectures presented in the previous chapter. In the first we presented the implementation of our work, where we presented our development environment on which the system was realized, after that we analyzed the results for each experiment we made by calculating evaluation metrics and presenting graphs in order to facilitate the comparison of the experiments.

With the presented results we have shown that the fusion technique allows to design a better performing hybrid system.

General conclusion

The work done in this dissertation consists in implementing an automatic kinship verification system. Then, to improve its performance by using new techniques preprocessing, for parameter extraction, score calculation and fusion.

In this kind of system, the environment and the different types of variability have a huge influence on its performance. The basic system, which we started with, using Haar filters for face detection and Multi-scale Retinex (MSR) with Normal Distribution Mapping (NDM) for the enhancement of images' quality in the preprocessing task, For parameter extraction based on shallow features we used our proposed descriptor (Hist-2D DWT). We employed an optimized method refferd as TXQDA-WCCN for dimensionality reduction and variability compensation. The evaluation focused on Cornell KinFace and TS KinFace datasets. After several experiments in which we implemented our descriptor with and without preprocessing and diffrent blocs, we managed to reach correct verification accuracy of **92.01%** , **86.78%** for Cornell and TS KinFace datasets.

After that we replaced our descriptor with LPQ descriptor. We resumed the same experiments as before. We found that the results have improved. In terms of correct verification accuracy, the result has become equal to **94.16%** for Cornell Kinface and **87.62%** for TS KinFace.

Another parameter extraction method based on deep features (VGG16 and ResNet50) was introduced with set of configurations for our system.We kept the same protocol to make an objective comparison with the shallow features. This method has been applied on the same Cornell Kinface and TS KinFace datasets. The results of VGG16 and ResNet50 were respectively **91.02%,73.15%** for Cornell Kinface dataset and **78.63%,72.65%** for TS KinFace dataset.

In order to benefit from the advantages of the two methods (deep and shallow)we have applied a fusion of the scores on the system. Using Logistic Regression (LR) fusion with four extractors (VGG16, ResNet50, Hist-2D DWT, LPQ), we have achieved an excellent result in correct verification accuracy of **95.18%** and **91.81%** for Cornell and TS datasets respectively.

At the end of this work, we believe that we have achieved a system that meets the objective that we set for ourselves. Thus, the use of score fusion between two parameter extraction methods (Deep and Shallow) for automatic kinship verification system allows to have a better robustness by improving the performance of this system.

Perspectives

The perspectives of this work revolve around two main points:

- Use of other and newer pretrained CNN models for features extraction such as AlexNet [66], ResNet101 [45], VGG-19 [67] along with BSIF,LBP, LOOP and the proposed Hist-2D DWT approach.
- CNN models perform poorly on significantly small datasets, thus the need for bigger datasets for kinship verification is crucial to achieve a significant improvement in terms of accuracy using a raw CNN model.

Bibliography

[1] Bordallo Loprz,M., Hadid, A., Boutellaa, E., Goncalves, J., Kostavos, V., & Hosio, S. (2018). Kinship verification from facial images and videos: human versus machine. *Machine Vision and Application, 29,873-890.*

[2] Qin, X., Liu, D., & Wang, D. (2020). A literature servey on kinship verification through facial images. Neurocomputing, 377, 213-224.

[3] Horman,S., Knoche, M., & Rigoll, G. (2020). *A Multi-Task Comparator Framework for Kinship Verification. 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020).*

[4] Wu, X., Feng, X., Li,L., Boutellaa, E., & Hadid, A. (2019). Kinship verification based on deep learning. *Deep Learning in Object Detection and Recognition, 113-132.*

[5] Jiwen Lu, Xiuzhuang Zhou, Yap-Pen Tan, Yuanyuan Shang. & Jie Zhou (2014).*Neighborhood Repulsed Metric Learning for Kinship Verification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(2),331-345.*

[6] Xiaoting Wu , Boutellaa, E., Feng,X., & Hadid, A.(2016). Kinship verification from faces: *Methods, databases and challenges. 2016 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC).*

[7] Fanf, R., Tang, K. D., Dnavely, N., & Chen, T. (2010, September). Towards computational models of kinship verification. In 2010 IEEE International conference on image processing (pp. 1577-1580). IEEE

[8] Yan, H., Lu, J., Deng, W., and Zhou, X. (2014). Discriminative multimetric learning for kinship verification. Information Forensics and Security, IEEE Transactions on, 9 :1169–1178.

[9] Bordallo Lopez, M., Hadid, A., Boutellaa, E., Goncalves, J., Kostakos, V., & Hosio, S. (2018). Kinship verification from facial images and videos: human versus machine. Machine Vision and Applications, 29, 873-890

[10] Prabhakar, S., Pankanti, S. & Jain, A. K. (2003). Biometric recognition: Security and privacy concerns. IEEE security & privacy, 1(2), 33-42.

[11] Almuashi, M., Mohd Hashim, S. Z., Mohamad,D., Alkawaz, M. H., & Ali, A. (2017). Automated kinship verification and identification through human facial images: a survey. *Multimedia Tools and Application*, 76, 265-307.

[12] Kohli, N. (2019). *Automatic kinship verification in unconstrained faces using deep learning. West Virginia University.*

[13] Guo, G., & Wang, X. (2012). Kinship measurement on salient facial features. *IEEE Transactions on Instrumentation and Measurement, 61(8), 2322-2325*

[14] Nash, S., Rhodes, M., & Olszewaska, J. I. (2016, February). Ifr: Interactively pose corrected face recognition. In International Conference on Bio-insipred System and Signal Processing (Vol.5,pp.106-112). SCITEPRESS.

[15] V.I. B. R. A. T. N. A. L. Spectroscopies et al., "Pattern Recognition, Analysis and Application," in intechopen, Dr. Mahali., vol. i, no. tourism, S. Ramakrishnan, Ed. 2016, p.13.

[16] Liu, L., Xiong, C., Zhang, H., Niu, Z., Wang, M., & Yan, S. (2015). Deep aging face verification with large gaps. *IEEE Transactions on Multimedia, 18(1), 64-75.*

[17] Lu, J., Zhou, X., Tan, Y. P., Shang, Y., & Zhou, J.(2013). Neighborhood repulsed metric learning for kinship verification. *IEEE transaction on pattern analysis and machine intelligence, 36(2),331-345.*

[18] Shao, M., Xia, S., & Fu, Y.(2011, June). Genealogical face recognition based on ub kinface database. In *CVPR 2011 workshops (pp. 60-65).* IEEE.

[19] Fang, R., Gallagher, A. C., Chen, T., & Loui, A. (2013, September). Kinship classification by modeling facial feature heredity. in *2013 IEEE international conference on image processing* (pp. 2983-2987).

[20] X. Qin,X. Tan, and S. Chen, "Tri-subject kinship verification: Understanding the core of a family," IEEE Transactions on Multimedia, vol. 17, no. 10, pp.1855-1867,2015.

[21] S. Xia, M. Shao, J. Luo, and Y. Fu, "Understanding kin relationship in a photo," IEEE Transactions on Multimedia, vol.14, no.4, pp.1046-1056,2012.

[22] Lu, J., Hu, J., Zhou,X., Zhou,J., Castrillo-Santana, M.,Lorenzo-Navarro,J., ... & Vieira, T.F. (2014, Septembre). *In IEEE international joint conference on biometrics (pp. 1-6)* IEEE.

[23] Han, H., Shan, S., Chen, X., & Gao, W. (2013). A comparative study on illumination preprocessing in face recognition. Pattern Recognition, 46(6), 1691-1699.

[24] Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision, 57, 137-154.*

[25] Rahman, Z. U., Jobson, D. J., & Woodell, G.A. (1996, Septembre). Multi-scale retinex for color image enhancement. *In Preceedings of 3rd IEEE international conference on image precessing (Vol. 3,pp. 1003-1006).* IEEE.

[26] Parihar, A. S., & Singh, K. (2018, January). A study on Retinex based method for image enhancement. In *2018 2en International Conference on Inventive Systems and Control (ICISC)* (pp. 619-624). IEEE.

[27] Struc, V., Zibert, J., & Pavesic, N. (2009). Histogram remapping as a preprocessing step for robust face recognition. image, 7(8),9.

[28] Ojansivu, V., & Heikkila, J. (2008). Blur insensitive texture classification using local phase quantization. In *Image and Signal Processing: 3rd International Conference, ICISP 2008. Cherbourg-Octeville, France, July 1-3, 2008. Proceedings 3* (pp. 236-243). Springer Berlin Heidelberg.

[29] Olaja, T., Pietikainem, M., & Harwood, D.(1996). A comparative study of texture measures with classification based on featured distribution. Pattern recognition, 29(1),51-59.

[30] Dochi, N.P., & Schaefer, G. (2012, September). Rotation-invariant local bunary pattern texture classification. In *Preceedings ELMAR-2012* (pp. 71-74). IEEE.

[31] Kannala, J., & Rahut, E. (2012, November). Basif: Binarized statistical image features. In *Proceedings of the 21st international conference on pattern recognition (ICPR2012).*(pp.1363-1366). IEEE.

[32] Chakraborti, T., mcCane, B., Mills, S., & Pal, U. (2018). LOOP descriptor: local optimal-oriented pattern. IEEE Signal Processing Letters, 25(5), 635-639.

[33] Chang, W. H., Lee, Y. S., Peng, W. S., & Lee, C. Y. (2001, May). A line-based, memory efficient and programmable architecture for 2D DWT using lifting scheme. In ISCAS 2001. The 2001 IEEE International Symposium on Circuits and Systems (Cat. No. 01CH37196) (Vol. 4, pp. 330-333). IEEE

[34] Huang, Z. H., Li, W. J., Shang, J., Wang, J., & Qhang, T.(2015). Non-uniform patch based face recognition via 2D-DWT. Image and Vision Computing, 37,12-19.

[35] Le Cue, Y., Jackel, L. D., Boser, B. Denker, J.s, Graf, H.P., Guyon, I.,... & Hubbard, W.(1989). Handwritten digit recognition: Applications of neural network chips and automatic learning. *IEEE Communications Magazin, 27(11),* 41-46.

[36] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceeding of the IEEE conference on computer vision and pattern recognition*(pp.815-823)

[37] Wen, Y., Ahang, K., Li, Z., & Qiao, Y.(2016). A discriminative feature learning approach for deep face recognition. In Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14,2016, Proceeding, Part VII 14 (pp. 499-515). Springer International Publishing.

[38] Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2017). Sphereface: Deep hypersphere embedding for face recognition. In *Proceeding of the IEEE conference on computer vision and pattern recognition* (pp. 212-220).

[39] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceeding of the IEEE conference on computer vision and pattern recognition (pp.770-778).

[40] Khan, A., Sohil, A., Zahoora, U., & Qureshi, A. S. (20200). A survey of the recent architectures of deep convolutional neural network. *Artificial intelligence review, 53,* 5455-5516.

[41] Shrestha, A., & Mahmood, A. (2019). Review of deep learning algorithms and architectures. IEEE access, 7,53040-53065.

[42] Fei-Fei Li, Ranjay Krishna, D.X. (2020). Convolutional Neural Networks for Visual Recognition. https://cs23In.github.io/convolutional-networks/. accessed on March 20,2023.

[43] Monasse at al., (2020) Pascal Monasse, K.N. (2020). Clasez et segmentez des données visuelles, Part 3. https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-données-visuelles/. accessed on March 20,2023.

[44] Robinson, J.P., Shao,M., Wu, Y., & Fu, Y.(2016). Family in the wild(FIW): A large-scale kinship recognition database. *arXiv preprint: 1604.02182.*

[45] Parkhi, O., Vedaldi, A., and Zisserman, A. (2015). Deep face recognition. volume 1, pages 41.1–41.12.

[46] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceeding of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[47] Robinson, J. P., Yin, Y., Khan, Z., Shao, M., Xia, S., Stopa, M., ... & FU, Y. (2020, November). Recognizing families in the wild (RFIW): The 4th edition. In *2020 15 IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)* (pp. 857-862). IEEE.

[48] Van Der Maaten, L., Postma, E., & Van den Herik, J. (2009). Dimensionality reduction: a comarative. J Mach Learn Res, 10(66-71),13.

[49] Liao, S., Hu, Y., Zhu, X., & Li, S.Z. (2015). Person re-identification by local maximal occurrence representation and metric learning. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2197-2206).

[50] Kanagasundaram, A., Vogt, R., Dean,D., Sridharan, S., & Mason, M.(2011). I-vector based speaker recognition on short utterances. In Proceedings of the 12th annual conference of the international speech communication association (pp. 2341-2344). International Speech Communication Association.

[51] Schuldt, C., Laptev, I., & Caputo.B . (2014, August). Recognizing human actions: a local SVM approach. In Proceeding of the 17th International Conference on Pattern Recognition,2004. ICPR 2004. (Vol. 3, pp. 32-36). IEEE.

[52] Gue, G., Wang, H., Bell, D., Bi, Y., & Greer,K. (2003). KNN model-based approach in classification. In ON the Move to Meaningful Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7,2003. Proceedings (pp. 986-996).

[53] Elsalmony, H. A. (2016). Healthy and unhealthy reed blood cell detection in human blood smears using neural networks. *Micron,83,*32-41.

[54] Das, P.K.; Meher, S.; Panda,R.; Abrahman, A. (2020). A Review of Automated Methods for the Detection of Sckle Cell Disease. IEEE Rev. Biomed.Eng., 13,309-324.

[55] Sharma, P., Berwal, Y.P.S, & Ghai, W. (2020). Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. Information Precessing in Agriculture, 7(4), 566-574.

[56] Bessaoudi, M., Chouchane, A., Oumane, A., & Boutellaa, E. (2021). Multilinear subspace learning using handcrafted and deep features for face kinship verification in the wild. *Applied Intelligence, 51,* 3534-3547.

[57] Hatch, A. O., Kajarekar, S., & Stolcke, A. (2006). Within-class covariance normalization for SVM-based speaker recognition. In *Ninth international conference on spoken language precessing.*

[58] Shum, S., Dehak, N., Dehak, R., & Glass, J.R. (2010, June). Unsupervised speaker adaptation based on the cosine similarity fot text-independent speaker verification. IN *Odyssey* (p.16).

[59] pigeon, S., Druyts, P., & Verlinde, P. (2000). Applying logistic regression to the fusion of the NIST'99 1-speaker submissions. Digital Signal Processing, 10(1-3), 237-248.

[60] Deep Learning Toolbox, Design, train, and analyze deep learning networks: https://www.mathworks.com/products/deep-learning.html?s_tid=FX_PR_info

[61] Bessaoudi, M., Ouamane, A., Belahcene, M., Chouchane, A., Boutellaa, E., & Bourennane, S.(2019). Multilinear side-information based discriminant analysis for face and kinship verification in the wild. *Neurocomputing, 239,*267-278.

[62] Goyal, A., & Meenpal, T. (2021). Eccentricity based kinship verification from facial images in the wild. *Pattern Analysis and Applications, 24,*119-144.

[63] Zhang, L., Duan, Q., Zhang, D., Jia, W., & Wang, X. (2020). AdvKin: Adversarial convolutional network for kinship verification. *IEEE transactions on cybernetics, 51*(12, 5883-5896.)

[64] Mukherjee, M., & Meenpal, T.(2022). Binary cross coupled discriminant analysis for visual kinship verification. *Dignal Processing : Image Communication, 108,*116829.

[65] Serraoui, I., Laiadi, O., Ouamane, A., Dornaika, F ., & Taledb-Ahmed, A. (2022). Knowledge-based tensor subspace analysis system for kinship verification. *Neural Networks, 151* 222-237.

[66] Gu S, Ding L, Yang Y, Chen X. A new deep learning method based on AlexNet model and SSD model for tennis ball recognition. In2017 IEEE 10th International Workshop on Computational Intelligence and Applications (IWCIA) 2017 Nov 11 (pp. 159-164). IEEE.

[67] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014 Sep 4.

Appendix A

Annex

1.1 Graphical user interface:

In our project, we created a graphical interface that ensures easy communication between the user and the Automatic Kinship Verification system by:

- simplifying the reading and understanding of results by optimizing the way they are presented by the system.
- making it easier for the user to run the system by proposing a list of pre-established choices.

Our interface is simple and realised using GUIDE in Matlab 2021

1.1.1 interface home window:

The home window illustrated in Fig A.1 offers an easy-to-use experience through several steps :

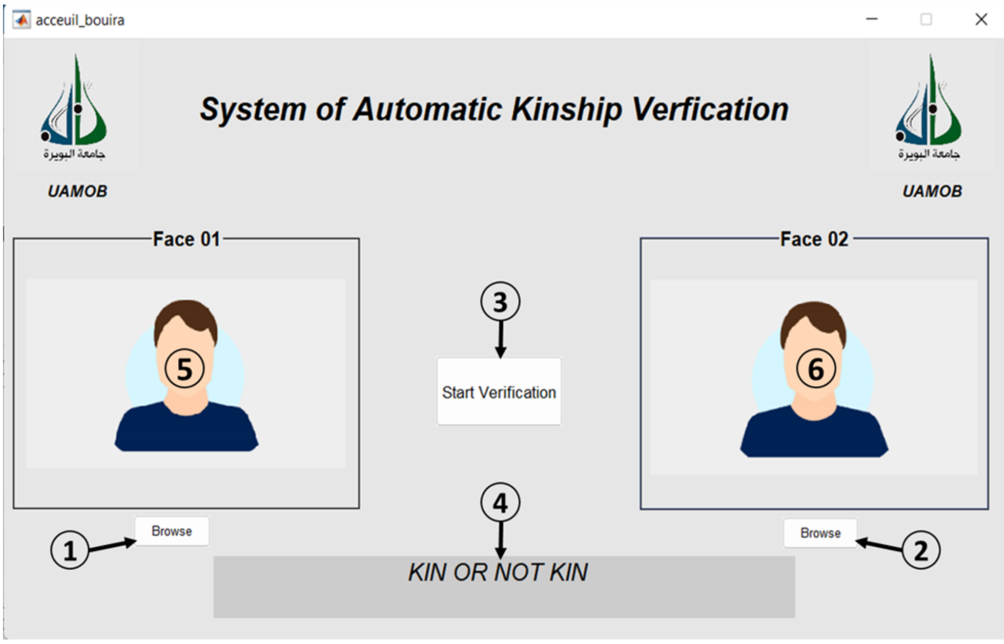

Figure 1.1: interface home window.

(1)- Browse button for face 01 : which allows us to load face image 1 from any directory on the machine and display it in **(5)**.

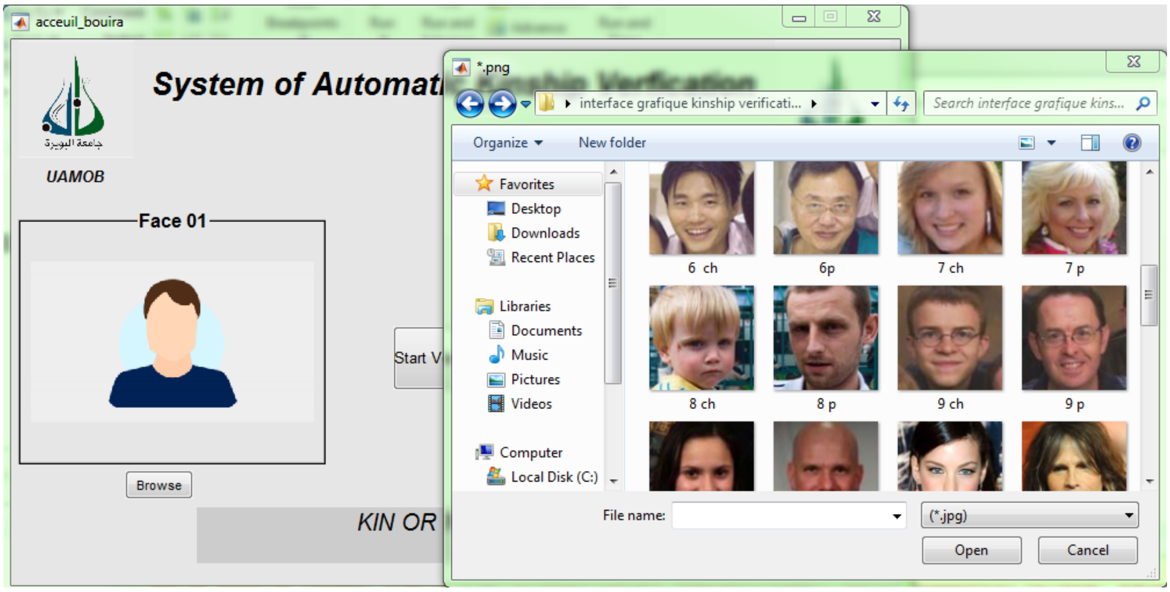

Figure 1.2: Load Face image 1

(2)- Browse button for face 02 : which allows us to load face image 2 that is likely to be related to face image 1 from any directory on the machine and display it in **(6)**.

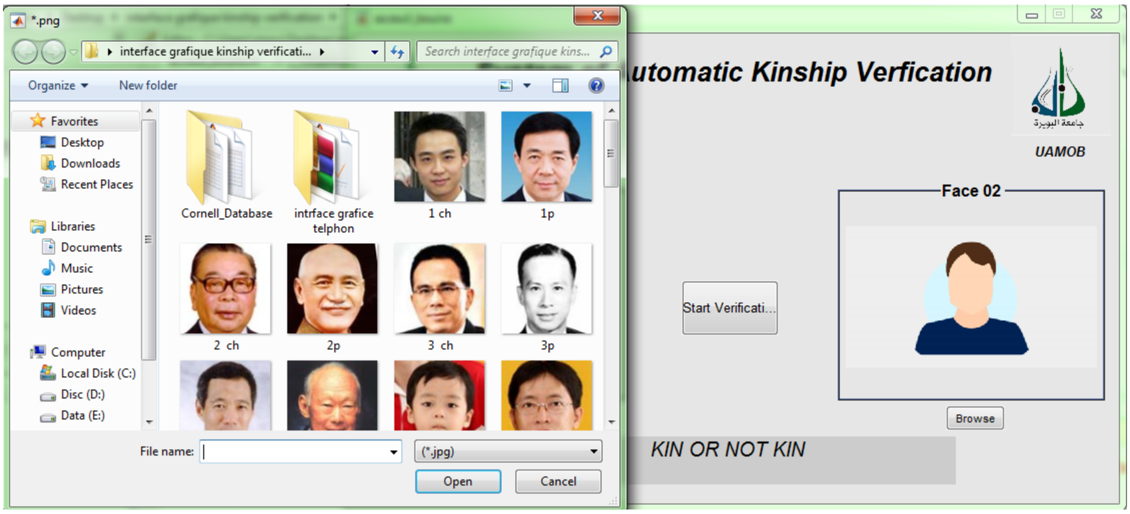

Figure 1.3: Load Face image 2

After the load of two images that we want to verify if they're related window will become like illustrated in **Fig A.4.**

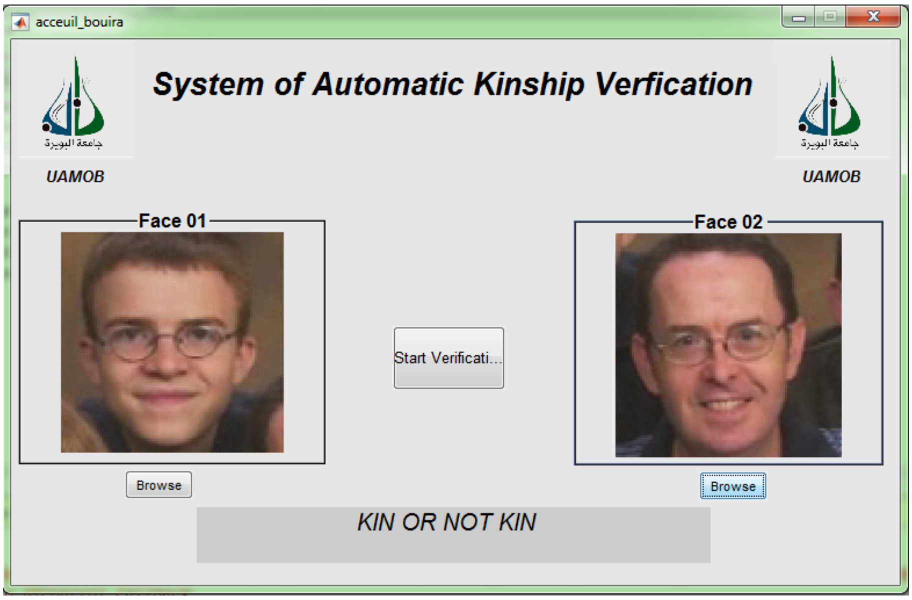

Figure 1.4: 2 face images loaded.

(3)- Start Verification button : by pushing it we start running the Verfication system and the result will pop up in **(4)** display as correct verification rate [%] (KIN Not KIN) as illustrated in the following figures.

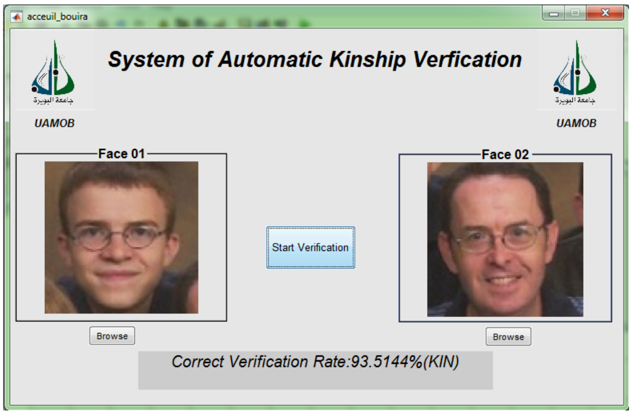

Figure 1.5: Kinship Verification test (KIN).

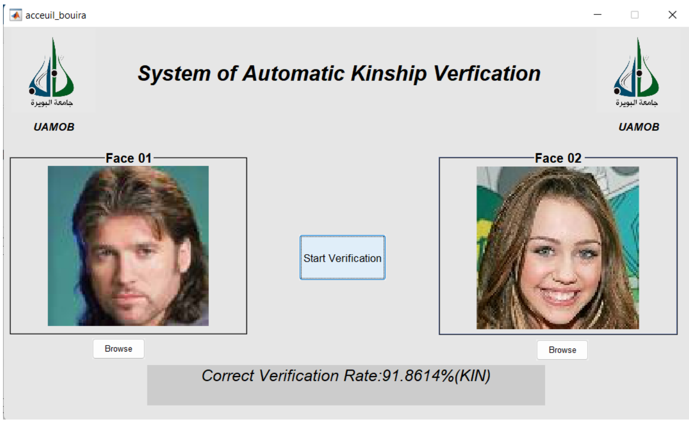

Figure 1.6: Kinship Verification test (KIN).

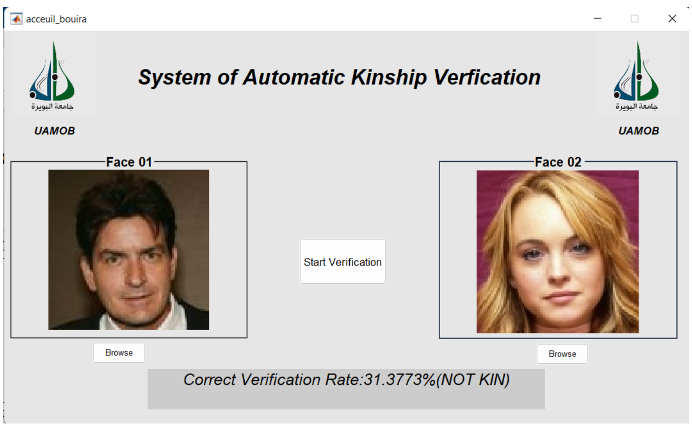

Figure 1.7: Kinship Verification test (NOT KIN).

الملخص

يجذب التحقق من القرابة من خلال صور الوجه المزيد و المزيد من الاهتمام من مجتمع البحث، و هو موضوع بحتي مستجد في رؤية الكومبيوتر. يمكن التحقق تلقائيا مما إدا كان شخصان من نفس العائلة أم لا من خلال صور الوجه. العديد من التطبيقات المحتملة: مثل إنشاء أشجار العائلة و تنظيم ألبومات العائلة، و وضع تعليقات توضيحية على الصور، البحث عن الأطفال المفقودين و الطب الشرعي، كلها مستهدفة بواسطة التحقق من القرابة.

يقدم هذا البحث نظاماًا ناجحاًا للتحقق من القرابة، والذي يستخدم طريقتين متتاليتين ((MSR + NDM

في مرحلة المعالجة المسبقة للصور لتحسين جودة الصورة والتغلب على المشكلات المتعلقة بالتباين والإضاءة والضوضاء. بالإضافة إلى ذلك، نقترح واصفاًا جديداً استناداًا إلى الرسوم البيانية للتحويل المويجي المنفصل ثنائي الأبعاد (Hist-2D DWT) نحن نحقق كذلك في تكامل الميزات الحرفية (Hist-2D DWT , LPQ) والميزات العميقة (VGG16 , RestNet50) من خلال دمجها على مستوى النتيجة باستخدام طريقة الانحدار اللوجستي.
تم إجراء تجارب مكثفة على مجموعتي بيانات قرابة، وتم الوصول إلى دقة تحقق تبلغ %91.81 te %95.18 ضمن مجموعتي بيانات و Coernell KiFace , TS KinFace

الكلمات المفتاحية

التحقق من القرابة، واصف Hist - 2D DWT ميزات عميقة،ميزات حرفية، MSR + NDM ، اندماج LR.

Résumé

La vérification de la parenté à partir d'images faciales attire de plus en plus l'attention de la communauté des chercheurs, est un sujet de recherche émergent en vision par ordinateur. Vérifier si deux personnes sont de la même famille ou non peut être automatiquement vérifié par des images faciales. De nombreuses applications potentielles : telles que la création d'arbres généalogiques, l'organisation d'albums de famille, l'annotation d'images; la recherche d'enfants disparus et la médecine légale, sont visées par la vérification de la parenté.

Cet article présente un système de vérification de la parenté réussi, qui utilise deux méthodes consécutives (MSR+NDM) dans la phase de prétraitement de l'image afin d'améliorer la qualité de l'image et de surmonter les problèmes liés au contraste, à l'éclairage et au bruit. En outre, nous proposons un nouveau descripteur basé sur les histogrammes d'une transformée en ondelettes discrète bidimensionnelle (Hist-2D DWT). Nous étudions de plus la complémentarité des caractéristiques artisanales (LPQ, Hist-2D DWT) et des caractéristiques profondes (VGG16, ResNet50) en les fusionnant au niveau du score à l'aide de la méthode de régression logistique.

Des expériences approfondies menées sur deux ensembles de données de parenté, des précisions de vérification de 95,18% et 91,81% ont été atteintes sous les ensembles de données Cornell KinFace et TS KinFace.

Mots Clés Vérification de la parenté, descripteur Hist-2D DWT, caractéristiques profondes, caractéristiques artisanales, MSR+NDM, fusion LR.

Abstract

Verification of kinship from facial images is attracting more and more attention from the research community, is an emerging research topic in computer vision. Checking whether two people are from the same family or not can be automatically checked by facial images. Many potential applications: such as creating family trees, organizing family albums, annotating images; the search for missing children and forensic medicine, are targeted by the verification of kinship.

This paper presents a successful kinship verification system, which utilizes two consecutive methods (MSR+NDM) in the image preprocessing stage to enhance image quality and overcome issues relating to contrast, lighting, and noise. Additionally, we propose a new descriptor based on the histograms of a Two dimensional Discrete Wavelet Transform (Hist-2D DWT). We further investigate the complementarity of handcrafted (LPQ, Hist-2D DWT) and deep features (VGG16, ResNet50) by fusing them at the score level using the Logistic Regression method.

Extensive experiments conducted on two kinship datasets, verification accuracies of **95.18%** and **91.81%** have been reached under Cornell KinFace and TS KinFace datasets.

Key words: Kinship Verification, Hist-2D DWT descriptor,Deep Features,ShallowFeatures, MSR+NDM, LR Fusion. . . .