Automatic
Kinship Verification

Amina Tidjani
University Kasdi Merbah Ouargla

A thesis submitted for the degree of
Doctor

Ouargla 2018-2019

DEMOCRATIC AND POPULAR REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY OF KASDI MERBAH OUARGLA

FACULTY OF NEW INFORMATION AND COMMUNICATION TECHNOLOGIES

DEPARTMENT OF ELECTRONIC AND TELECOMMUNICATION

Thesis submitted for the degree of Doctor LMD 3rd cycle

Communications and Signal Processing

By: Amina Tidjani

Face based Automatic Kinship Verification

Mohammed Lamine KHERFI	*MCA*	*Ouargla University*	*President*
Kamal Eddine AIADI	*Professor*	*Ouargla University*	*Supervisor*
Abdelmalik TALEB AHMED	*Professor*	*Valencienne University*	*Co - Supervisor*
Abdallah MERAOUMIA	*MCA*	*Tebessa University*	*Examiner*
Mohammed BOULASBAA	*MCA*	*Ouargla University*	*Examiner*
Fella CHARIF	*MCA*	*Ouargla University*	*Examiner*
Djamel SAMAI	*MCA*	*Ouargla University*	*Guest*

2018/2019

Dedication

This thesis is dedicated to

Our spiritual guide: ***Dr. Tidjani Sidi M. Laid***

My parents: who gave me life, the symbol of tenderness, who sacrificed themselves for my happiness and success, who gave meaning to my life, and who was the shadow during all the years of study

To my beloved husband: ***Youmbai mohamed zoheir***

To my dear sisters: ***Asma*** and ***Meriem***

To my beloved: ***Sidi Abd El jabar (Babi)***

To the memory of my uncle ***Mohamed Sghir Tidjani***

May god guard them and protect them.

To all my friends

To all those who are dear to me.

To all who love me.

To everyone I love.

Acknowledgements

Praise to our Lord "ALLAH" who has endowed us with wonderful faculty of reasoning. Praise to our Creator which prompted us to acquire knowledge.
I want to thank my dissertation supervisors Prof. Kamel eddine Aidi, Prof. Abdelmalik TALEB-AHMED, Dr. Djamel Samai, Prof. Abdenour Hadid, and Dr. Elhocine Boutella for his precious advice and help throughout the work period.
I am deeply grateful to all members of the jury for agreeing to read the manuscript and judge this thesis.
I would like to thank my parents Ahmed and Bahia for their love and constant support. I owe you everything.
My thanks also extend to all my teachers Kasdi Merbah University - Ouargla during the years of study.
And all those who are dear to me and who helped me to realize this work and gave me the will and the desire to do everything to succeed.
I will not forget to thank my colleagues Khaled Ben Sid, Maarouf Koraichi and Nariman.

Abstract

Face based automatic kinship verification is a novel challenging research problem in computer vision. It performs the automatic examining of the facial attributes and expecting whether two persons have a biological kin relation or not.

The focus is on providing novel solutions for challenges of family verification from faces with an efficient system with the aim of providing enhancement to the accuracy of kinship verification.

In our work, we analyzed the facial kinship verification systems in two modes the unimodal and multi-modal system. The feature extraction is a crucial step in the kinship recognition system. For this reason, we proposed two efficient feature learning extraction algorithms called discrete cosine transform network (DCTNet) and Context-Aware Local Binary Feature Learning (CA-LBFL).

Various databases are used and extensive experiments are carried out in order to validate our proposed methods and developed methods.

Besides, the experimental results demonstrated that the proposed methods achieved competitive results compared with other state-of-the-art.

Index Terms Biometrics, Kinship, Verification, DCTNet, CA-LBFL.

Résumé

En vision par ordinateur, la vérification automatique de la parenté est un nouveau systéme de recherche. Il effectue l'examen automatique des attributs faciaux et montre si deux personnes ont une relation biologique.

L'objectif est de fournir de nouvelles solutions aux problémes de vérification de la parenté avec un systéme efficace dans le but d'améliorer la précision du systéme de vérification de la parenté.

Dans notre travail, nous avons analysé le système de vérification de la parenté faciale selon deux modes, le système unimodal et le système multimodal. L'extraction des caractéristiques est une étape essentiel dans le systéme de reconnaissance de la parenté. Pour cette raison, nous proposons deux algorithmes d'extraction des caractéristiques efficaces appelés réseaux de transformation en cosinus discret et l'apprentissage des motifs binaires locaux contextuelles.

Diverses bases de données sont utilisées et différentes expériences sont menées afin de valider les approches proposées et les méthodes développées.

En outre, les résultats expérimentaux démontrent que les méthodes proposées a permis d'obtenir des résultats compétitifs par rapport á l'état de l'art .

Mots-clés - Biométrie, parenté, vérification, DCTNet, CA-LBFL.

Contents

List of Figures

List of Tables

Abbreviations

BNRML : Block-based Neighborhood Repulsed Metric Learning

BSIF : Binarized Statistical Image Features

CA-LBFL : Context-Aware Local Binary Feature Learning

CNN : Convolutional Neural Networks

COALBP : Co-Occurrence of Adjacent LBP

DCT : Discrete Cosine Transform

DCTNet : Discrete Cosine Transform Network

DMML : Discriminative Multi Metric Learning

DFT : Discrete Fourier Transform

FPLBP : Four Patch Local Binary Patterns

GSML : Generalized Sparse Metric Learning

HOG : Histograms of Oriented Gradient

HSV : Hue Saturation Value

KML : Kinship Metric Learning

LBP : Local Binary Patterns

LDA : Linear Discriminant Analysis

LMNN :	Large Margin Nearest Neighbor
LPP :	Locality Preserving Projections
LPQ :	Local Phase Quantization
MKSM :	Multiple Kernel Similarity Metric
NN :	Nearest Neighbor
NRML :	Neighborhood Repulsed Metric Learning
PCA :	Principal Component Analysis
RBF :	Radial Basis Function
ROC :	Receiver Operating Characteristic
SIFT :	Scale Invariant Feature Transform
SILD	Side-Information based Linear Discriminant
SML:	Single Metric Learning
SMCNN	Similarity Metric based Convolutional Neural Network
SRC :	Sparse Representation Classifier
SSML :	Sparse Similarity Metric Learning
SVM :	Support vector machines
TOP :	Three Orthogonal Planes
TPLPB :	Three Patch Local Binary Patterns
WLD :	Weber Local Descriptor

Chapter 1

Introduction

This Chapter contains:

The ability to automatically verify whether two persons are from the same family is referred to as family (kinship) verification. The general objective of this thesis is to research, develop and evaluate novel computational models to determine whether two persons are from the same family or not based on their faces. For example, the inputs could be two faces (Face A and Face B) and the expected output could be a decision whether Person A is the father, mother, sister, brother of Person B or not.

The focus is on providing novel solutions for challenges of family verification from faces with an efficient system with the aim of providing enhancement to the accuracy of kinship verification.

This chapter introduce the main topic of this thesis that is facial kinship verification. The challenge is to automatically learn and extract the most significant features of human faces. This chapter begins by outlining the kinship verification topic in Section 1.1. Section 1.2 provides the contribution of this thesis. Section 1.3 describes the thesis organization.

1.1 Kinship Verification

Facial appearance is a primary source of information regarding the persons identity, gender, ethnicity, affective state, head pose, age and kinship relations. Hence, the perception of facial attributes governs person perception, interpersonal attraction, and consequently pro-social and social behavior [1], [2] and [3].

Automatic detection of kinship from facial appearance is a challenging problem, and it has been thoroughly studied across several disciplines such as psychology e.g., [4], [5], [6] sociology e.g., [7], [8], medicine e.g., [9] , anthropometry e.g., [10], Human Behavior and Evolution Society e.g., [11], [12] and recently in computer vision and machine learning e.g., [13], [14], [15], [16], [17], [18], [19].

In this thesis, we focus on using computer vision techniques to automatically detect the relationships between two persons from their faces.

Facial kinship verification is a module aiming to analyze pedigree construction and calculate similarity indices for different facial traits. This problem has several interesting applications, such as family album organization, finding missing children and social media it still challenging for real applications because there are usually large variations in pose, expression, illumination, age, ethnicity, especially, when face images are captured in unconstrained environments.

The studies of kinship verification in psychology and sociology [7], [8] have found that:

1. Humans can recognize kin relationships among unknown individuals.

2. Human faces convey essential cues helping to determine the kin relationships between individuals.

3. Since facial appearance is found to be a useful cue for genetic similarity, previous studies attempted to identify which facial features are providing clues for humans to recognize the relationships between two people.

Inspired by the observations and results of the psychological studies and by capitalizing on recent advances in computer vision and machine learning, researchers investigated the kinship verification problem from facial images, aiming to develop computational models and algorithms to automatically verify kin relationships between people. Usually, four different types of kinship relations are considered: father-son (F-S), father-daughter (F-D), mother-son (M-S), and mother-daughter (M-D).

During the past years, the challenge of kinship verification has attracted different levels of research. In 2010, Fang et al. [20] the first researchers who tackled the kinship verification challenge, they extracted low-level features from facial parts. In 2011, researchers started to investigate this problem with several mid-level features [21]. Recently, some researchers used the high-level feature representations.

Fundamentally, two important challenges in kinship verification are addressed. The first one is related to the environment of the database such as age, illumination, facial expression, pose variations, resolution and rotations. The second one is related to kinship itself this latter is based on checking the existence of kinship by verifying feature resemblance between kin. Moreover, features extraction is important and most sensitive because special feature have an observable impact on the efficiency of the recognition system.

This thesis focus on kinship verification based feature extraction task. We present an original investigation to detect the most significant characteristic features of human faces and it should help from the improved performance of kinship verification.

1.2 Research Contributions

The main contributions of this thesis are highlighted as follows:

1. Firstly, the biggest challenge is how to describe a face, how to recognize the different facial features for children, and then learn how to relate them to their corresponding parents without being affected by different challenges. We addressed many challenges: i) diverse ages, expressions, gender, skins, lighting changes even dramatically illumination; ii) several features extraction methods are explored, such as: the Hand-Crafted Features and the learned features (No Hand-Crafted Features).

2. An experiment study is provided. We investigate the global and local features and matching approaching to gain insights into the problem. The global appearance-based methods try to find a suitable representation of the whole image. The local texture descriptors represent certain region properties.

3. We explore the feature combination to perform multiple feature fusion to extract complementary information to improve the kinship verification performance.

4. We have proposed an efficient system based on Context-Aware Local Binary Feature Learning (CA-LBFL) for kinship verification. The CA-LBFL is a method has applied to learn contextual features from raw pixels directly and to eliminate the dependence on hand-crafted features.

5. We have proposed an efficient system with the aim of providing enhancement to the accuracy of kinship verification. The system based on a simple deep learning method called Discrete Cosine Transform Network (DCTNet), where

2D-DCT is adopting as a filter bank to extract the most significant inherited facial features. To the best of our knowledge, the DCTNet is being used for the first time in our work for the kinship verification. We introduced the multimodal system based on some discriminative biological information.

1.3 Thesis Organization

The rest of thesis is divided as following:

Chapter 2 details the facial kinship verification task. We give an overview of automatic kinship verification from faces by presented some notions, definitions and terminology required for understanding this topic. Next, we describe the system design of facial kinship verification. Also some applications and the challenges related to facial kinship verification. We summarize some existing databases.

The chapter presents also a review of the current state-of-the-art in automatic kinship verification is provided. We discussed the features extraction methods that divided into two approaches. The first one based on the Hand-Crafted Feature-based methods including (Local Binary Patterns (LBP), Histogram of Gradient (HOG), Gabor Wavelets... etc.). The second approach based on the Deep Learning Feature-based method. Also several proposed metric learning to tackle the kinship verification problem. However, three different kinship verification setting has been proposed to evaluate the performance of the system.

In chapter 3, the features extraction and classification method used for the kinship verification are discussed in detail in this chapter. First, we introduce the hand-crafted features extraction techniques including global and local features. Next, we present two methods based on learned (No Hand-Crafted) features learning, the first one is the CA-LBFL and the second approach is based on DCTNet. The learned features are proposed in order to eliminates the dependence on hand-crafted features (extract and represent features automatically instead of selected manually) and to extract the most significant inherited facial features. The results obtain are discussed and shows that the proposed approach provided a valuable solution to the facial kinship verification problem.

Two matching methods are proposed to deal with kinship verification. Firstly, we discuss a supervised methods that called Support vector machines (SVM). SVM performs classification by finding the hyperplane that gives the largest minimum distance to the training examples. The second method present a simple approach that requires no training data by computing the distance between pairs of features with different functions. The rest of the chapter, provide a brief discussion of the tools and techniques used. Functionality of feature selection algorithm is explained. Also, we explains the different fusion techniques and its related strategies. We present the details of performance evaluation measurement throughout the thesis.

In chapter 4, experimental results of facial kinship verification analysis are presented and discussed. We have investigated the kinship verification task using three experiments. In experiment 1, we have investigated several feature extraction methods and similarity measures. In experiment 2, we tackled the problem using a novel solution using a DCTNet via 2D-DCT filters bank. To improve the kinship verification performance, a discriminative biological information are used. From a biological opinion, the chromaticity of the face is tied to genetically expressed characteristics, such as eye color or skin tone, In experiment 3, we proposed an effective method called CA-LBFL. The CA-LBFL is a method has applied to learn contextual features from raw pixels directly and to eliminate the dependence on hand-crafted features.

Finally, Summary and future work are drawn in **Chapter 5**.

Chapter 2

Facial Kinship Verification

This Chapter contains:

2.1 Introduction

Kinship verification is operation can verify the relationship of a person based on facial characteristics. This skill is well demonstrated in recognizing people in images. In this chapter, we present the background information and the notions required for understanding the kinship verification topic. Also, we present an overview of the challenges related to kinship verification through human facial images and we introduce the main remedies proposed by state-of-the-art works.

Firstly, we discuss the kinship terminology and then we present the concept of kinship verification in computer vision. We highlighted the difference between facial verification and kinship verification system. Also, we devoted to system design and application. The chapter also summarizes some existing databases. Finally, we present the approaches and protocols proposed for solving this problem.

2.2 Kinship terminology

We need to be familiar with various kinship terminologies before deepening into the problem and the solutions.

As a special type of social relationship, kinship plays an important role in the field of social network analysis [22]. Kinship is a term that has several meanings. In general, kinship (similarity, familiarity) is a word used to recognize relationships between individuals in a family. In biology, kinship is the degree of genetic relatedness between two family members. In psychology, Kinship is the state of being related by birth, common ancestry, marriage or adoption [23]. In the anthropological context, kinship refers to the network of social relationships between people that form an important part of the lives of most humans in most societies [24]. In Physiognomy, is to estimate the relationships of children and their parents, based on signs and to the likeness of them, and to identify the proportions of the child in view of the members of his body and the members of his father. Moreover, different societies classify kin relationships differently and therefore use different systems of kinship terminology [25].

2.3 Kinship in computer vision

A biometric system is essentially a pattern-recognition system that recognizes a person based on a feature vector derived from a specific physiological or behavioral characteristic (also known as modalities) that the person possesses [26].

These modalities provide a very high level of protection against fraud and it has several key advantages such as not transferable, non-repudiation, not guessable. The technology has been successfully implemented in different real-life applications such as forensics, government agencies, banking, and financial institutions, enterprise identity management.

Indeed, one of the most used biometric modalities is the face, according to its potentiality and capability to distinguish between people; it has several advantages over other biometric modalities e.g., fingerprint [27], [28], Palm-prints [29], [30], voice [31], [32] or iris [33], [34]. It is more accepted due to its major advantage: it is the only physiological biometric that can be reliable measured remotely and, moreover, the authentication of the users can happen without their explicit interaction with the sensor or their knowledge. Several areas are emerging, such as: age estimation [35], face anti spoofing [36], facial cosmetics [37], analyzing attractiveness, for surgical/ orthodontics planning [38].

At recent days, one of the new areas of application of biometric that exploit the face modalities is the kinship verification.

Automatic kinship verification system tries to recognize the relationships between people based only on photographs of their faces. Kin recognition is the task of training the machine to recognize the genetic kin and non-kin based on features extracted from digital images and try to determine whether or not kinship exists between a pair of faces, but they do not aim to recognize the exact type of kinship [39].

In computer vision, kinship verification is a very challenging problem it encounters many variations as in face recognition problems such as low-resolution images, illumination changes, pose variations, effects of aging, mixed ethnicities, and multiple age groups.

In addition, the kinship system is the classification of persons related through kinship based on patterns such as their faces (see figure 2.1). For example, the inputs could be two faces (Face A and Face B) and the expected output could be a decision whether Person A is the father/sister/mother/brother of Person B or not.

2.4 Face verification and kinship verification

At first look, the kinship verification process may seem similar to face verification process. In effect, these two processes are different things but they have common factors. We might say that kinship verification is the highest level of face verification [39].

The basic structure is the common factor between kinship verification and face verification system and the differences between the facial and kinship verification is shown in Table 2.1.

Table 2.1: Difference between kinship and face verification system.

Facial Verification	Kinship Verification
Extract features from same person	Extract features from different person
Verify or identity	Verify the relationships
Same trait of query image1 and query image 2	Different trait of query image1 and query image 2
Height level system	Highest level system
The most common application is the security, monitoring and extensively by government a agencies	Studying the biological relationship and utilizes by people and less by government agencie
In decision stage Matched or not matched	In decision stage Kin or not kin
Performance of the machine is very roughly as accurate as human	Accuracy is not satisfactory

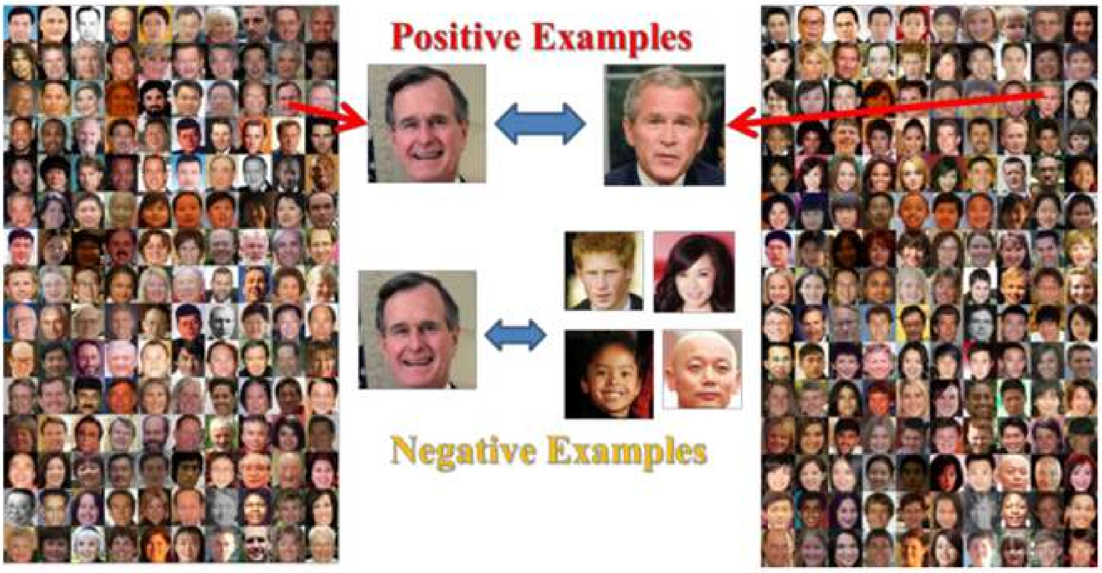

Figure 2.1: An example of kinship verification via face images. The images on the left correspond to parents and those on the right to children. The challenge is to automatically find the correspondences parent-child.

2.5 System design

The automatic detection of kinship from facial images is a difficult problem that recently received attention from the computer vision and pattern recognition research community. Kinship is a genetic relationship between two family members. Specifically, there are four different types of kinship relations: Mother-Daughter (M-D), Father-Daughter (F-D), Mother-Son (M-S) and Father-Son (F-S) kinship relations see Figure 2.2.

The kinship verification has been studied in psychology and sociology [20], [40] and [41]. The findings from the literature include that:

- Humans are able to recognize kinship relationships between unknown individuals.
- The mechanism of kinship perception is probably different from identity recognition.
- Human faces can convey some important cues to identify the kin relations of persons.
- Even attempted to identify the facial features providing kinship clues.

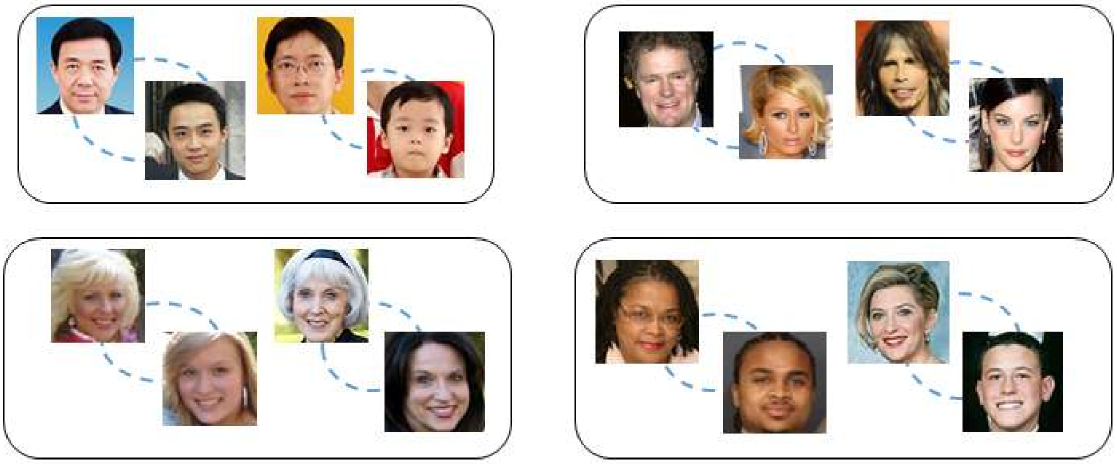

Figure 2.2: KinFaceW-II database. Four type relationships: (F-S), (F-D), (M-D), and (M-S).

- Facial appearance is a useful cue for genetic similarity.

Inspired by this observation, researchers started to investigate this problem from facial images, where the objectives are to develop computational models and algorithms to verify human kin relations.

The aim of kinship verification is to recognize the genetic kin or not-kin relationship based on image feature. In recent years, several methods have been proposed to investigate this problem using computer vision and machine learning. In 2010, Fang et al. [20] was tackle the first attempt of kinship verification; they proposed a method to automatically verify parent-child image pair relationships through the analysis of facial features. His method works according to the following steps:

- **Step 1:** Parent-child database collection. Over on-line search, they collected the databases that contain facial pairs image (parent-child) from celebrities and public figures.

- **Step 2:** Inherited facial feature extraction. They used a set of low-level image features extractions including, facial parts, color and geometry distances between parts and gradient of the face.

- **Step 3:** Classifier training and testing. Objective of this experiment is to classify the child - parent pairs into two false and true categories. Same number of negative pairs (no kin relation) as positive pairs (with a kin relation) is created by associating faces from persons which do not have a kinship relation.

The positive examples are the true pairs of parents and children and negative examples are each parent with a randomly selected child from the children images who is not his/her true child. Using the extracted feature vectors, they calculate the differences between features vector of the corresponding parents and children, and apply different machine for classification.

Figure 2.3 presents a general schema of kinship verification based on facial images. This kinship system has the following essential components:

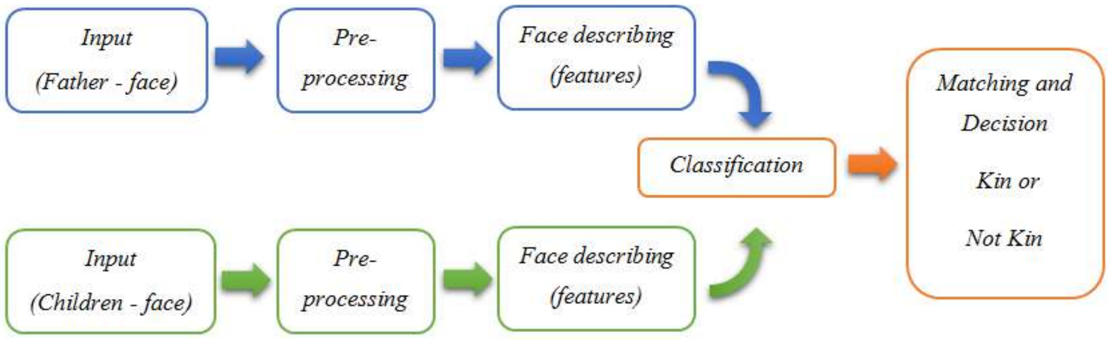

Figure 2.3: General schema of kinship verification system.

1. **Input (face datasets):** includes public figure face images of family members. Therefore, the face images are captured under uncontrolled environments with no constraints.

2. **Pre-processing:** is an important component. In order to localize the face of the image, this latter is then cropped, such that the non-facial regions such as the background and hairs were removed and only facial region was used for kinship verification. If those are color images, they converted into gray-scale images. For each cropped image, histogram equalization was applied to mitigate the illumination.

3. **Face describing:** the best way to learn kinship analysis is first to learn how to recognize the different facial features for children, and then learn how to relate them to their corresponding parents. Facial features must be described in a way that enables them to be efficiently descriptors.

4. **Classification and Decision:** kinship Verification is bi-classification problem where the pair of input faces is classified as positive (the true pairs of parents and children) or negative (the false pairs of parents and children).

2.6 Application

Determining if someone is a father, mother, son, or daughter is a complicated task. It is possible to use the most advanced of methods to determine if two persons are likely to have the suspected relation or not. Researchers started to investigate the problem of kinship verification, where the objectives are to estimate relatedness between closely related individuals based on face features. There are some potential applications for kinship verification such as family album organization, social media, missing child search and entertainment.

2.6.1 Social media and family album organization

Over the past few years, with the development of technology in social media, most smart phone and social media such as Google+ and Facebook use facial recognition system to automatically organize billions of images. There are two principal questions to be responded: (1) who these people are, and (2) what their relations are. The first question can be addressed with Face recognition technique and kinship verification is a useful system to approach the second question. When the relationships between people are recognized, it is possible to automatically construct the family tree from this network society.

2.6.2 Finding missing children

Another important application of kinship verification is missing children search. Although a DNA test is the most accurate test for family kinship verification, but, it is restricted in some applications. It unfortunately cannot be used in many scenarios such as in video surveillance, and the DNA privacy testing and the coast is very high.

However, facial kinship verification can solve these limitations because verifying kinship relationships from facial images is very convenient and its cost is very low. For example, if we want to find a missing child from thousands of children, it is difficult to use the DNA testing to verify their kin relation due to privacy concerns. However, if our kinship verification method is used, we can quickly first identify some possible candidates which have high similarity from facial images. Then, the DNA testing is applied to get the exact search result.

2.6.3 Entertainment

The film making industry used the visual effects that can age or rejuvenate the actors. These effects are not limited to movies but are also widely applied to photo editing. The imminent integration of such tools into popular design software will make for more realistic retouch of photos. Make-up artists that specialize in transforming the face can leverage the construction of person, age and kin specific morph-able models. Guided by those models, the artists will transform the face of the actor for roles that demand sibling-like similarity actors.

2.7 Kinship challenges

In this thesis, the challenge is to improve the accuracy of facial kinship verification so that the several complications do not cause a reduction of the performance of the system. Fundamentally, we focus on kinship verification based feature extraction task. We considered the facial kinship verification problem as features extraction problem. In particular, facial kinship verification is very challenging. There are at least two major challenges. The first one is related to the environment of the database. The second one is related to kinship itself.

- The face of pairs (parent- child) may look different due to variance in age, gender, and mixed ethnicity environments and also to deal with other factors such as: pose variations, illumination, and expression, especially, when face images are captured in unconstrained environments.

- Secondly, extracting features can be considered as a main problem to determining the relationships. We must initially determine which facial features are most relevant for determine parentchild relationships. It is necessary to describe and extract the most inherited feature in order to perform robust kinship verification system.

2.8 Databases

The appearance of a face is affected by a large number of factors including identity, face pose, illumination, facial expression, age, occlusion, and facial hair. The development of algorithms robust to these variations requires databases of sufficient size that include carefully controlled variations of these factors. Furthermore, common databases are necessary to comparatively evaluate algorithms. Kinship verification continues to be one of the most popular research areas of computer vision and machine learning. Along with the development of kinship verification algorithms, a comparatively diverse databases have been collected. Examples of different data sets are shown in Figure 2.4.

Table 2.2 provides a list of datasets that can be used for kinship verification.

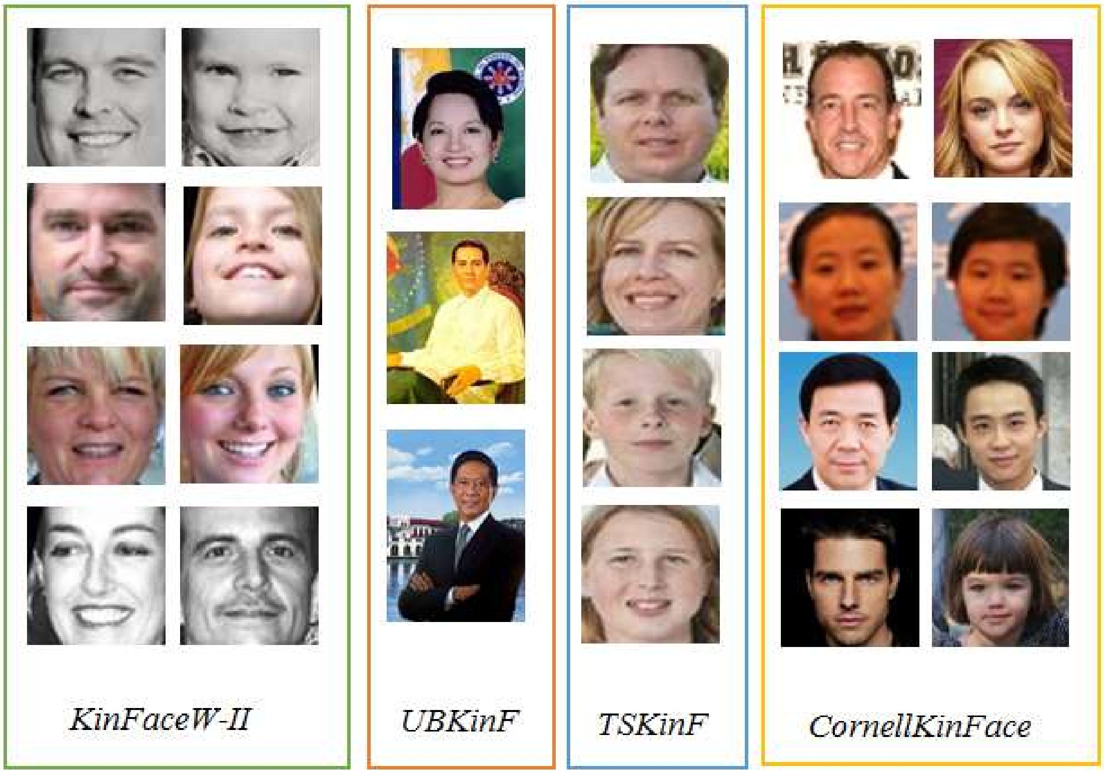

Figure 2.4: Samples of images from a different database for different kin relations. KinFacew-II. From top to bottom are the (F-S), (F-D), (M-S) and (M-D) kinship relations. UB inFace (Daughter-Yong father and Old Father), TSKinFace (F-M-S-D), Cornell KinFace (F-D), (M-D), (F-S) and (F-D).

In this section we review different databases:

- KinFaceW-I , KinFaceW-II [42] [1]: includes public figure face images of family members collected from the Internet. Therefore, the face images are captured under uncontrolled environments with no constraints. The difference between KinFaceW-I & KinFaceW-II is that face images with a kin relation were acquired from different photos in KinFaceW-I while, in most cases, faces are cropped from the same photo in KinFaceW-II. KinFaceW-I dataset contains 156 F-S, 134 F-D,

[1] http://www.kinfacew.com/

Table 2.2: Kinship verification databases.

Input	Database	Number of Pair image	Type of relation	kinship Pairs	Controlled condition	Pairs from same photo	Publicly available
Image	KinFaceW-I	533	F-S	156	No	Partially	Yes
			F-D	134			
			M-S	116			
			M-D	127			
	KinFaceW-II	1000	F-S	250	No	Yes	Yes
			F-D	250			
			M-S	250			
			M-D	250			
	UBKinFace	400	Old P-C	200	No	No	Yes
			Yong P-C	200			
	TSKinFace	1015	F-M-S	285	No	Yes	Yes
			F-M-D	274			
			F-M-S-D	228			
	Family 101	206	FamilY	206	No	No	Yes
	IIITD Kin	272	S-S	52	No	No	No
			F-D	33			
			F-S	52			
			M-D	26			
			M-S	15			
			B-S	49			
			B-B	42			
	Cornell Kin	143	P-C	143	No	Partially	Yes
Video	UvA-NEMO Smile	95	F-S	94	Yes	No	No
			F-D	58			
			M-S	82			
			M-D	133			
			B-B	28			
			S-S	52			
			B-S	66			
	KFVW	418	F-S	107	No	Partially	No
			F-D	101			
			M-S	100			
			M-D	110			

116 M-S, and 127 M-D pairs of kinship images. For the KinFaceW-II dataset, each relation contains 250 pairs of kinship images.

- UB KinFace [41] [2]: comprises 600 images of 400 people which can be separated into 200 groups. Each group is composed of child, young parent and old parent images. Most of images in the database are real-world collections of public figures (celebrities and politicians) from Internet. The face images were collected without restriction in terms of pose, expression, illumination, background, age, ethnicity, and occlusion.

- Cornell Kin Face [20] [3]: contains 143 of parents-children pairs. The database is collected through a controlled on-line search for images of public figures and celebrities and their parents or children. The database includes faces with variations in age, gender, race, career, etc.

- TSKinFace (Tri-Subject Kinship Face) [43] [4]: The "TSKinFace" dataset is the first large-scale dataset of families for one-versus-two kin relation. It contains 1015 different family with distinct family names, including 2,589 individuals, with 787 images. All images in the dataset are harvested from the internet based on knowledge of public figures family and photo-sharing social network such as flickr.com. Each family contains one child and two parents. The final dataset includes 274, 285 and 228 family photos for Father-Mother-Daughter (FM-D), Father-Mother-Son (FM-S) and Father-Mother-Son- Daughter (FM-SD), respectively. Two kinds of family-based kinship relations are constructed in the TSKinFace database: Father-Mother-Son (FM-S) and Father-Mother-Daughter (FM-D). The FM-S and the FM-D contain 513 and 502 groups of tri-subject kinship relations, respectively. Hence we have 1015 tri-subject groups in our database totally. The families included in our database are diverse in terms of races as well. For FM-S relation, there are 343 and 170 groups of tri-subject kinship relations for Asian and non-Asian, respectively. And for FM-D relation, the numbers for Asian and non-Asian groups are respectively 331 and 171. For pair-wise relationships, there are 513 father-son relations, 502 father-daughter relations, 513 mother-son relations, and 502 mother-daughter relations.

[2] http://www.ece.neu.edu/ yunfu/research/Kinface/Kinface.htm
[3] http://chenlab.ece.cornell.edu/projects/KinshipVerification
[4] http://parnec.nuaa.edu.cn/xtan/data/TSKinFace.html

- The "Family101" dataset [40] [5]: it is a large-scale dataset of families across several generations and race. The "Family101" includes around 72% Caucasians, 23% Asians, and 5% African Americans. For relationships, there are 213 father-son relations, 147 father-daughter relations, 184 mother-son relations, and 148 mother-daughter relations.

- IIITD Kinship [44]: includes celebrities face images collected from the Internet. Therefore, the downloaded face images are classified into four ethnicities: Asian, Afro-American, Indian and American. The kinship relation has been categorized into the following seven relations: Brother-Brother, Brother-Sister, Father-Daughter, Father-Son, Mother-Daughter, Mother-Son, and Sister-Sister.

- UvA-NEMO Smile Database [45] [6]: is a large-scale smile database which has 1240 smile videos (597 spontaneous and 643 posed) from 400 subjects. Ages of subjects vary from 8 to 76 years. Videos are in RGB color and recorded with a resolution of 19201080 pixels at a rate of 50 frames per second under controlled illumination conditions. For further illumination and color normalization, a color chart is present on the background of the videos. Many families participated in the database collection, allowing its use for evaluation of automatic kinship from videos. A total of 95 kin relations were identified between 152 subjects in the database. There are seven different kin relations between pairs of videos: Sister-Sister (S-S), Brother-Brother (B-B), Sister-Brother (S-B), Mother-Daughter (M-D), Mother-Son (M-S), Father-Daughter (F-D), and Father-Son (F-S). The association of the videos of persons having kinship relations gives 228 pairs of spontaneous and 287 pairs of posed smile videos.

- Kinship Face Videos in the Wild (KFVW) [46]: the KFVW dataset included 418 pairs of face videos was collected from TV shows on the Web under uncontrolled condition such as lighting, pose, occlusion, age, expression, makeup, background, etc. Each video contains about 100 500 frames The average size of a video frame is about 900 500 pixels. There are four kinship relation types in the KFVW dataset: Father-Son (F- S), Father-Daughter (F-D), Mother-Son (M-S), and Mother-Daughter (M-D), and there are 107, 101, 100, and 110 pairs of kinship face videos for kin relationships F-S, F-D, M-S, and M-D respectively.

[5] http://chenlab.ece.cornell.edu/projects/KinshipClassificatio

[6] https://www.uva-nemo.org/

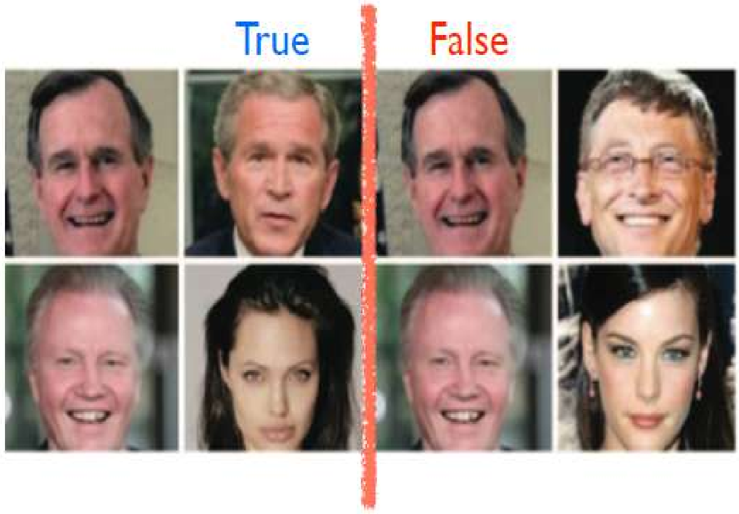

Figure 2.5: Examples of collected data base (parent-child) pairs from the work of Fang et al.[20].

2.9 State of the art in Kinship verification

Face recognition is defined as the automated identification or verification of individuals based on facial physiological characteristics. Face recognition has become a popular topic of research in image processing, computer vision and biometrics over the last years and one of the most successful applications of image analysis. The realm of face recognition systems growing both in the business and security screening sectors, thereby making them a prevalent solution for old techniques. Face recognition tools were developed to verify the relationship between two family members. This technology is called family (kinship) verification. This concept based on analyzing family relationships.

Automatic kinship verification aims to recognize the degree of kinship of two individuals (e.g. parents and children) from their facial images.

A remarkable interest has been given to the problem of kinship verification by researchers these last years. This interest is motivated by the potential applications of the topic, especially in social data mining.

Fang et al. [20] were among the first researchers who tackled the kinship verification. Database used containing 150 pairs image, collected from the Internet, as illustrated in Figure 2.5. They used the Pictorial Structures Model to locate facial parts (eyes, nose, mouth, etc.) by representing an object by a collection of parts arranged in a deformable configuration. Then, they extracted a set of 22 low-level features from these parts including:

- Parts color: the central position of the facial part is found, and the color at this point is used such as eye color and hair color. For skin color, the center of the nose was used. For hair color, a sub-window of the top of the image was taken, and a mode filter is applied to this sub-window to obtain the most commonly occurring color in this region.

- Facial parts (geometry): they detected the central position and the boundaries of each part. With these image coordinates, they extracted the sub-window for each face.

- Facial Distances: distances between parts are calculated using the Euclidean distance.

- Histogram of Gradients feature (HoG).

The K-Nearest-Neighbor (KNN) and Support Vector Machine (SVM) are used to classify the pairs of images as having kin relation or not. They performed the K-Nearest-Neighbors with K = 11 and Euclidean distance, and Support Vector Machine with a radial basis function (RBF) kernel and the LibSVM package [47]. The best classification achieved 70.69%, outperforming the 67.19% obtained by a panel of human raters on the same data.

Motivated by the first results, several scientists have investigated the kinship verification via facial images, different approaches have been proposed. These methods can be roughly divided into two approaches. The first one based on Hand- Crafted Feature-based methods. The second one based on the Deep Learning Feature-based methods.

On the other hand, various metric learning methods have been investigated for tackling the facial kinship verification problem. Furthermore, different protocols are proposed. Table 2.3 and Table 2.4 collects some of the main features used in facial automatic kinship verification.

2.9.1 Approaches based on hand-crafted features

The most proposed to tackle the facial kinship verification used traditional hand-crafted features (shallow structure). Typical the hand-crafted features descriptor include Histogram of Oriented Gradients (HOG) [67], Scale-Invariant Feature Transform (SIFT) [68], Gabor Filter [69], Local Binary Pattern (LBP) [70], Local Phase Quantization (LPQ) [71] etc. However, these features are based on the learning of the

Table 2.3: Summary of the Hand-Crafted features.

Paper	Year	Database	Features
[20]	2010	CornellKinFace	Face appearance & geometry
[21]	2011	UBKin Face	Gabor
[41]	2012	UBKin Face	Gabor
[44]	2012	IITD Kin	DoG salient points
[48]	2012	UBKin Face	Binary, Relative attributes
[40]	2013	Family 101	SIFT
[13]	2013	UvA-NEMO Smile	Dynamic features, CLBP-TOP, CLBP
[49]	2013	CornellKinFace	Geometric, WLD
[42]	2014	KinFaceW-I & II	LBP, LE, SIFT, TPLBP
[50]	2014	KinFaceW-I & II CornellKinFace UBKin Face	LBP, SIFT, SPLE
[51]	2014	KinFaceW-I & II	LE, LBP, TPLBP, SIFT
[52]	2014	Family101	LBP, Grassman Manifold
[53]	2015	KinFaceW-I & II	LBP, SIFT, SPLE
[54]	2015	KinFaceW-I & II	LPQ, TPLBP, FPLBP, WLD
[43]	2015	TSKinFace	SIFT
[55]	2015	KinFaceW-I & II CornellKinFace UBKin Face	LBP, SIFT
[56]	2016	KinFaceW-I & II	LBP, HOG
[57]	2016	KinFaceW-I & II	LBP, HOG
[58]	2016	KinFaceW-I & II	HOG, SIFT
[59]	2016	KinFaceW-I & II	SIFT
[60]	2017	KinFaceW-I & II	LBP, DSIFT, HOG, LPQ
[61]	2017	UvA-NEMO	BSIFTOP, LBPTOP, LPQTOP
[62]	2017	KinFaceW-I & II	LE
[63]	2017	KinFaceW-I & II	LTP
[46]	2018	KFVW	LBP, HOG
[64]	2018	KinFaceW-I & II CornellKinFace TSKinFace	BSIF, LPQ, CoALBP
[65]	2018	KinFaceW-I & II	LBP, HOG
[66]	2018	KinFaceW-I & II	LBP, HOG, SIFT, LPQ, WLD

surface properties and appearance of an object given by the shape, size, arrangement, density, a proportion of its elementary parts.

For instance, two papers on kinship verification [41] and [72] are proposed by Xia et al. The first paper [41] involved three different types of data sets, old parents, young parents and children set. The first two are images of the same persons at different ages (i.e. young and old). The authors employed transfer subspace learning (TSL) to find the similarity between children-young and children-old parents pairs. The Gabor wavelets are used for feature extraction.

In the second paper [72], their basic idea is to partition the face into five layers and use Gabor filters as face descriptor. The contextual feature used to determination the relationships with Transfer Subspace Learning.

Kohl et al. [44] tackled this new technology by proposing a new algorithm works according following steps: first, they normalize the face detected using the weber normalization. Next, The key points are extracted by taking the local extrema of the above difference of Gaussian (DOG) named (self-similarity descriptors). Finally, they show the result of classification accuracy of SVM.

Fang et al. [8] proposed a method to address the kinship challenge as reconstructing the query face from a mixture of parts from a set of families. They reconstructed the query face from a sparse set of samples among the candidate families. The family classification is determined based on the reconstruction error for each family. Twelve facial parts are selected to build the part-based dictionaries for sparse representation based classification (SRC).

Lu et al. [42] proposed a new largest kinship data sets called KinFace in the Wild I (KinFaceW-I) and KinFace in the Wild II (KinFaceW-II). There are four different relationships in both the KinFaceW-I and KinFaceW-II data sets: Father-Daughter (F-D), Father-Son (F-S), Mother-Daughter (M-D) and Mother-Son (M-S).

They addressed the facial kinship verification problem with a new neighborhood repulsed metric learning (NRML) method. They used a metric learning to seek an effective distance between pairs images. They aim to seek an effective distance between pairs images. They projected the distance as close as possible for facial images with kinship relations and those without kinship relations in the neighborhoods are pushed away as far as possible. They experimented with four features descriptors (LBP, HOG, LE and SIFT).

Kou et al. proposed method in [53] for explicitly learning a genetic similarity measure. The method is based on quadruple input, using LBP, SIFT, LE features, and issues a similarity matrix as output.

Qin et al. [43] exploiting the informations from both parents (Father-Mother) to detect the kinship relationship. Novel Relative Symmetric Bilinear Model (RSBM) was introduced to estimate the similarity between the parents and the child. They use the LBP, LE, SIFT and TPLBP for face representation.

Recently, many researchers have been using a combined of facial descriptors to verifying the kinship relation. For example, in the last kinship competition [73], all the proposed methods used three or more descriptors. The best performing method in this competition employed different local features (LBP, HOG, and OCLBP).

Xiaoting Wu et al. [71] investigated the usefulness of color information in the kinship verification from facial images. To encode both the chrominance and the luminance information in the color images, they extracted joint color-texture features. The performance of kinship verification using joint color-texture (LPQ, BSIF, and NRML) in the three color spaces (RGB, HSV, and YCbCr) is then compared against approaches using only Grey-scale information.

On the other hand, statistical techniques are applied to classify the kinship relationships. They are usually used to learn an effective classifier such as metric learning [42], transfer learning [41], and subspace learning [41].

Lu et al. [42] learned that the distances between pair image in same classes are as small as possible, while the distances are as large as possible between pair image different class based on textural features.

Yuan et al. [74] learned similarity matrix for kinship verification, by proposing a Sparse Similarity Metric Learning (SSML) method which enforces both the PSD constraints and the group sparsity.

Hu et al. [75] learned multiple global and local distance metrics, by maximising the correlations of features and the distance between each positive pair is less than a low threshold, and that for each negative pair is greater than a high threshold.

P. Ajit et al [59] presented a novel SIFT flow based genetic Fisher vector feature (SF-GFVF) which enhances the facial genetic features for kinship verification. The proposed SF-GFVF feature is derived by applying a novel similarity enhancement method based on SIFT flow and learning an inheritable transformation on the Fisher vector feature so as to enhance and encode the genetic features of parent and child image in kinship relations.

B. Patel et al.[63] explored the effectiveness of periocular region in verifying facial kinship captured in the wild. They proposed a block-based Neighborhood Repulsed Metric Learning (BNRML) to learns multiple local distance metrics from different blocks of the images represented by local ternary patterns. Moreover, to contemplate diversity in discrimination power of different blocks, weighted score-level fusion scheme is used to obtain a similarity score of image pair. Extensive experiments on KinFaceW-I and KinFaceW-II datasets demonstrated the potential of periocular features for kinship verification.

To evaluate the performance of different kinship verification algorithms three setting has been proposed in the FG kinship competition (the 2015 IEEE International Conference on Automatic Face and Gesture Recognition, Ljubljana, Slovenia.) [73]. The unsupervised setting, image-restricted setting and image-unrestricted setting. For each face image, they extracted two different feature descriptors: the LBP and HOG.

- Unsupervised setting: No labeled kin relation information is used. Given a face pair, the cosine similarity of their features is used to compute their similarity directly.

- Image-restricted setting: Only the given kin relation information is used in the training splits. For each face image, we first apply PCA to project feature into a low-dimensional feature vector and then Side-Information based Linear Discriminant analysis (SILD) [76] is employed to learn a distance metric. Specifically, the positive pairs and negative pairs in the training set were used to estimate the within-class and between-class variations of LDA. Finally, the cosine similarity of each test pair in the learned LDA space is computed.

- Image-unrestricted setting: The identity information of the person is available to potentially form additional negative pairs in the training splits. For each face image, PCA is first used to project feature into a low dimensional feature vector and then neighborhood repulsed metric learning (NRML) is employed to learn a discriminative distance metric. Specifically, the label of training sample is used to seek the most similar intra-class neighbors to learn the distance metric. Finally, the cosine similarity of each test pair in the learned NRML space is computed.

In this evaluation, the Image-restricted setting shows better kinship verification performance than the baseline methods. The most challenging Protocol is the unsupervised setting and the image-unrestricted setting is the easiest one.

Haibin.Y et al. [46] investigated the problem of video-based kinship verification. They present a new video face dataset called Kinship Face Videos in the Wild

(KFVW) which were captured in wild conditions for the video-based kinship verification. On the other hand, they compared the performance of several state-of-the-art metric learning based kinship verification methods.

Zhao, Yan-Guo, et al. [66] investigate the facial kinship verification based on the similarity computation witch is essentially an implicit nonlinear feature transformation. They proposed a novel Multiple Kernel Similarity Metric (MKSM) which cover a flexible family of linear and nonlinear metrics. The proposed MKSM method provides a simple framework for metric learning and feature fusion/selection in the Kinship Verification task.

In general, local descriptor constitutes power visual cues for feature representation. They provide discriminative information about small appearance details in local neighborhoods. So, they are robust to local changes databases such as illumination, identity, and expression.

However, these features are not learned, can be lost, difficult to design, and the performance using these features degrades dramatically in variations and unconstrained environments [39], therefore, it is required to use the deep learning technique for solving this problem.

2.9.2 Approaches based on deep learning features

Deep learning approaches are proposed to automatically learn and understand features from a huge amount of data like images, rather than designing features manually. Accordingly, we can gain useful features to perform tasks without much effort and request expertise. The deep features can provide highly informative representations for a classification task and thus lead to improved accuracy. In kinship verification, a few work has been proposed using deep learning.

In 2014, Dehghan, Afshin, et al. [77] introduced a new method for learning discriminative, genetic features for describing the parent-offspring relationship. They uncover three key insights that bridge the gap between anthropological studies and computer vision. They proposed an algorithm that fuses the features and metrics discovered via gated auto-encoders with a discriminative neural network layer that learns the genetic features to delineate parent-offspring relationships. Further, they analyzed the correlation between features detected and those found in anthropological studies.

Recent research [78] pointed out the misleading of kinship problem by automatically learning and extracting important descriptors using manually designed features. It used Similarity Metric based Convolutional Neural Network (SMCNN). The general

Table 2.4: Summary of the proposed deep learning methods.

Paper	Year	Database	Approach
[77]	2014	KinFaceW-I & II	Discriminative neural network layer
[79]	2015	KinFaceW-I & II	Deep Convolutional Neural Networks (CNN)
[78]	2016	KinFaceW-I & II	Similarity Metric Based Convolutional Neural Network (CNN)
[61]	2017	UvA-NEMO Smile	Deep Convolutional Neural (CNN)
[80]	2018	KinFaceW-I & IIe CornellKinFace TSKinFace	Deep Convolutional Neural (CNN)
[81]	2019	KinFaceW-I & II CornellKinFace UBKin Face	Deep Neural Network (DNN)

idea is to extract highly discriminative multi-scale features by using two convolution layers for eliminating the redundant information generated by the symmetrical structure in SMCNN. A cost function of the Siamese architecture is used to calculate the distance between the image pair.

Zhang et al. [79] proposed deep Convolutional Neural Networks (CNN) to extract the high-level features where produced from the neuron activations of the last hidden layer, and then fed the extracted feature into soft-max classifier to verify the kinship relationships. They also extracted the key-points-based features.

E. Boutellaa et al. [61] proposed a method to address the kinship verification challenge from spatio-temporal descriptors and explore the use of texture features (BSIFTOP, LPQTOP, and LBPTOP) and deep Convolutional Neural Networks (CNN) for characterizing faces, his experimental based on videos over still images. They have performed various experiments on UvA-NEMO Smile database. Support Vector Machine (SVM) was used to classify a pair of facial features as a positive or negative sample.

M. Dawson et al. [80] investigated the problem of facial kinship verification by training the deep Convolutional Neural Networks (CNN) classifier to determine if two faces are from the same photograph or not. Since faces from the same photograph are more likely to be from the same family.

Newly, Zhou, X et al. [81] proposed a method to address the facial kinship verification from kinship metric learning (KML) with a coupled Deep Neural Network (DNN) model. The proposed method aimed to learn a deep compact cross-generation similarity metric. Moreover, the proposed KML implicitly learns to fuse a pair of deep embeddings for robust similarity measure of the parent-child pairs.

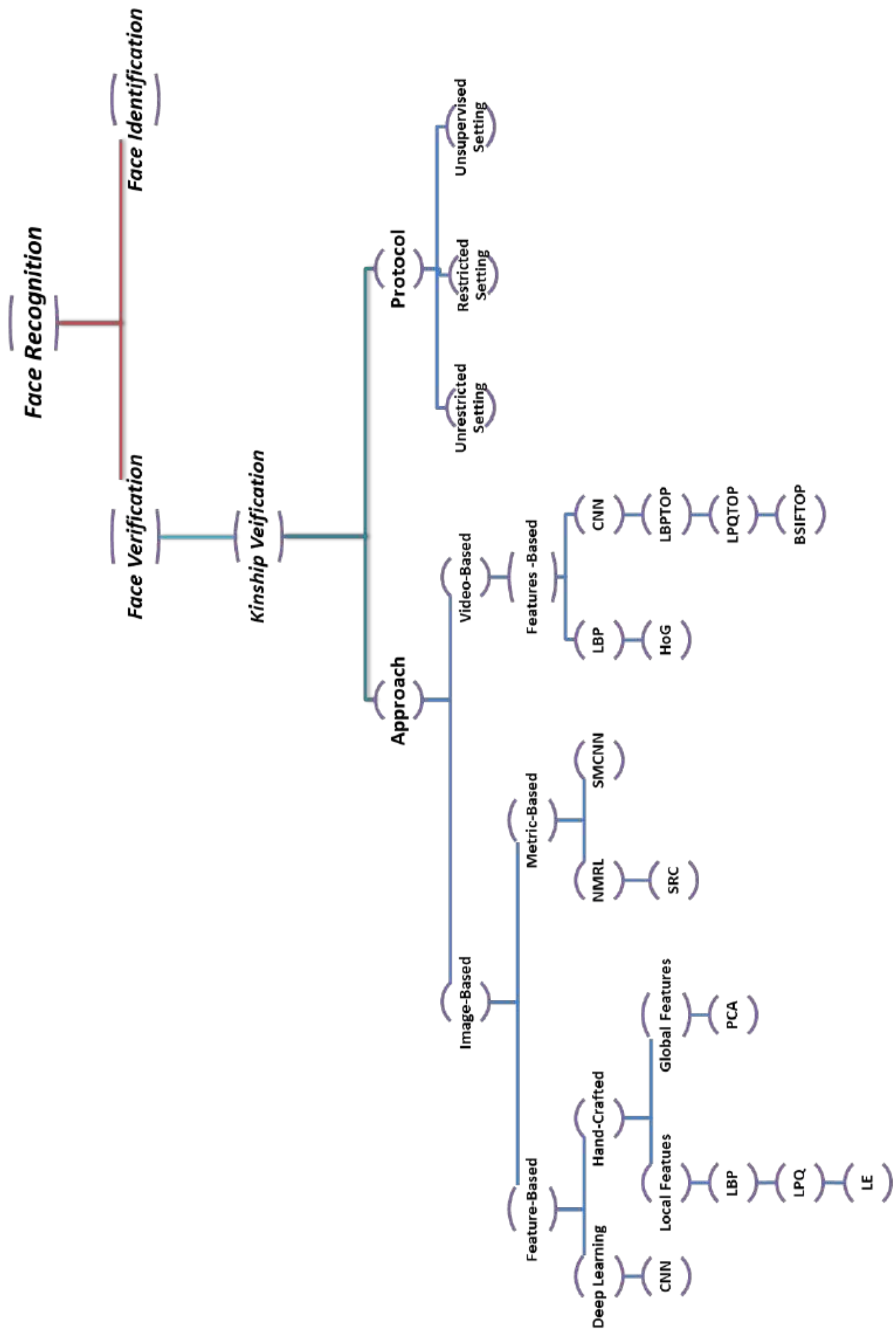

Figure 2.6: An overview of some issues related to facial kinship verification.

Figure 2.6 provide an overview of some issues related to facial kinship verification.

2.10 Conclusion

In this chapter, we gave an overview of automatic kinship verification from faces. We presented some notions and several definitions. The system design of facial kinship verification is described.

We discussed some exciting applications for the continuous research in this topic. We have also enumerated a number of practical challenges that restrain the facial kinship verification problems and we summarized the existing databases. Also, we have presented the main state of the art and an overview of automatic kinship verification approaches. We discussed some used methods that divided into Hand-Crafted Feature-based methods and Deep Learning Feature-based method. Several metric learning are proposed to describe the basic kinship verification like SSRW, NRML, SRC and transfer subspace learning.

In kinship verification problem, deep learning approaches are proposed were numerous layers of information processing stages are exploited, and deep learning was implemented by convolutional neural network (CNN) to feature extraction based and classification based.

Also, different kinship verification setting has been proposed to evaluate the performance of the system. In the next chapters of the thesis, we will be presented with the technical details of features extraction methods, protocols, classifiers used in this work.

Chapter 3

Kinship verification Methods

This Chapter contains:

3.1 Introduction

In this chapter, features extraction and classification methods are discussed. A facial image is a crucial clue that contains many useful human characteristics, such as identity, race, gender, age, expression, ethnicity, etc. What type of features help to identify the relationships and how the features are represented to determine the relationship, therefore, it becomes a major problem and an important challenge in estimating kinship. It is necessary to focus on the features extraction stage as it has a noticeable impact on the performance of the recognition system. This chapter begins by introducing the features extractions based on Hand-Crafted features. some important cues are proposed such as PCA, LDA, LBP, LPQ ...etc. Next, two approaches based on learned (No Hand-Crafted) features are proposed. Secondly, we focus the research on kinship classification and techniques used to improve the accuracy of system performance. Kinship Verification is bi-classification problem where the pair of input faces are classified as positive (the true pairs of parents and children) or negative (the false pairs of parents and children). Two different classifiers are used the support vector machine (SVM) and several metric learning (distances) such as L_1 and L_2.

Next, an optional phase called feature selection is conducted. The aim of the feature selection algorithm is to select the best subset of extracted features that gives the smallest classification error. The rest of this chapter presented the details of performance evaluation and fusion process.

3.2 Features extractions based on Hand-Crafted features

3.2.1 Global Appearance-based

Global appearance-based methods try to find a suitable representation of the whole image, all pixels are regarded, by approximating (reduce the dimension) the original data and keeping as much information as possible. Therefore, represented by the vector containing the weights of a linear combination of the basis vectors.

The holistic (global) approach takes the input face images globally and extracts important facial features based on the high-dimensional intensity values of face images automatically. The main feature of the global appearance-based is that they capture both facial texture and geometry information.

In this work, we extract three global descriptors: Principal Component Analysis (PCA)[82], [83], Linear Discriminant Analysis (LDA)[84], [83] and Locality Preserving Projections (LPP)[85], [83].

- The Principal Component Analysis (PCA), first introduced by Sirovic and Kirby [82], have been lately used for face analysis. PCA is one of the most successful techniques that have been used in face recognition. PCA is a mathematical method which transforms a set of correlated variables into uncorrelated variables by using an orthogonal components known as eigenvectors and eigenvalues.

  In mathematical terms, we wish to capture the variation in a set of face images and utilize this information to encode and compare images of individual faces in a holistic way. Specifically, a set of eigenfaces can be created by implementing a mathematical process named principal component analysis (PCA) on a large set of human faces images, or equivalently, the eigenvectors of the covariance matrix of the set of face images.

  The PCA was implemented for kinship verification on each set of images, to obtain the eigenvectors (or eigenfaces) of face corresponding to the higher variance among the images being analyzed. Aim to indicate only the most significant facial feature.

  Method:

  - Define a data matrix, $X = [v_1, v_2, ..., v_N]$
    Where N is the total number of learning images.
  - Calculate the mean.
    $$\Psi = \frac{1}{N} \sum_{i=1}^{N} v_i \tag{3.1}$$
  - Subtract the mean of the distribution from the data set.
    $$\phi_i = v_i - \Psi, i = 1...N \tag{3.2}$$
  - Calculate the covariance matrix.
    $$C = \sum_{i=1}^{N} \phi_i \phi_i^T = AA^T, A = [\phi_1, \phi_2, ..., \phi_N] \tag{3.3}$$
  - Calculate the eigenvalues and the eigenvectors of the covariance matrix. The weights are then obtained by the normalized components of the eigenvector (v_1) which corresponds to the highest eigenvalue:
    $$\beta_i = \frac{v_{1i}}{\sum_{i=1}^{N} v_{1i}} \tag{3.4}$$

- Linear Discriminant Analysis (LDA) or Fisherfaces method is one of the most popular supervised feature extraction techniques. LDA aims to find a linear combination of features characterizes or separates two or more classes. LDA seeks an optimal set of discriminant projection vectors to map the original data space onto a lower dimensional feature space, by maximizing the Fisher criterion. It takes into account the scatter between-classes and the scatter within-classes by maximize the between-class and minimizing the within-class. It is also adapted to distinguish the image variation due to identity from variation due to other sources such as expression and illumination.

  Method:

  - Let S_B be the between-class scatter and S_W the within-class scatter.

    $$S_B = \sum_{i=1}^{c} N_i(m_i - m)(m_i - m)^T \tag{3.5}$$

    $$S_w = \sum_{i=1}^{c}\sum_{j=1}^{N_i}(x_{ij} - m_i)(x_{ij} - m_i)^T \tag{3.6}$$

    Where x_{ij} denotes the jth training sample of the ith class, m_i is the mean of the training sample of the ith class and m is the mean of all the training samples.

  - The objective function of LDA is defined as:

    $$\max_{w} \frac{w^T S_B w}{w^T S_W w} \tag{3.7}$$

  - Solving the generalized eigenvalues problem.

    $$S_B w = \lambda S_W w \tag{3.8}$$

  - Let $w_1, w_2, ..., w_k$ be the eigenvectors corresponding to the k largest eigenvalues $\lambda_i \mid i = 1, 2, ..., k$ decreasingly ordered $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$, then $W = [w_1, w_2, ..., w_k]$ is the learned mapping of LDA. Since the rank of S_B is bounded by $c - 1$, k is at most equal to $c - 1$.

  The idea behind the choice of LDA is to separate the relation into two class, first class is with kinship relation and the second class is without kinship relation.

- The Locality Preserving Projections (LPP) was introduced by He and Niyogi [42], it is unsupervised technique and performs a linear transformation. Therefore, the Laplacian notion was used by LPP to builds a graph incorporating neighborhood information of the data set. The transformation matrix which maps the data points to a subspace was built. This linear transformation optimally conserves the local neighborhood information. In kinship verification, LPP aims to preserve the intrinsic geometry structure of the face and to maintain the locality relationship after projection. Before using those descriptors, the features are extracted by converting each image into gray-scale and averaging the whole image.

  Method

  1. Given a dataset of N samples, $X = x_1; x_2, ..., x_N; where x_i \in \Re^D$
  2. Find the transformation matrix W of size $D \times d$.
  3. map $y_i = WTx_i, where y_i \in \Re^d$
  4. Construct the adjacency graph either by
     - Neighborhood.
     - K-nearst neighbors.
     - Find the similarity matrix S (using Heat kernel).

     $$S_i^j = \begin{cases} \dfrac{\exp \parallel x_i - x_j \parallel^2}{t} & i, j connected \\ 1 & otherwise \end{cases}$$

  5. Compute generalized eigenvalues problem

     $$XLX^TW = \lambda XDX^TW \tag{3.9}$$

     Where $D : Diagonal matrix D_{ij} = \sum_j S_{ij} and L = D - S$

  The main difference between PCA, LDA, and LPP is that PCA and LDA focus on the global structure of the Euclidean space, while LPP focuses on the local structure of the manifold, but they are all considered as linear subspace learning algorithms.

3.2.2 Local Texture-based:

Local texture descriptors represent certain region properties by multi-dimensional histograms. Very often geometric properties (e.g., location, distance) of interest points in the region (corners, edges) and local orientation information (gradients) are used.

- Local Binary Patterns (LBP) were first introduced by Ojala et al. [86] to classify texture patterns. Latterly, Ahnon.t et al. [70] used the LBP for face description. The LBP is an operator assigns a label to every pixel of an image by thresholding its value with neighborhood pixels with the center pixel value and considering the result as a binary number. The basic LBP operator is illustrated in Figure 3.1. The histogram of the labels can be used as a texture descriptor. Mathematically, the LBP operator is defined as in Equation 3.10:

$$LBP(N,R) = \sum_{n=0}^{N-1} s(I_n - I_c)2^n \tag{3.10}$$

  Where N is the number of pixels in the neighborhood, R is the radius, and the threshold function s(x) = 1 if $x \geq 0$, otherwise s(x) = 0. The I_c and I_n values are the gray levels of the center pixel and thexth surrounding pixel, respectively.

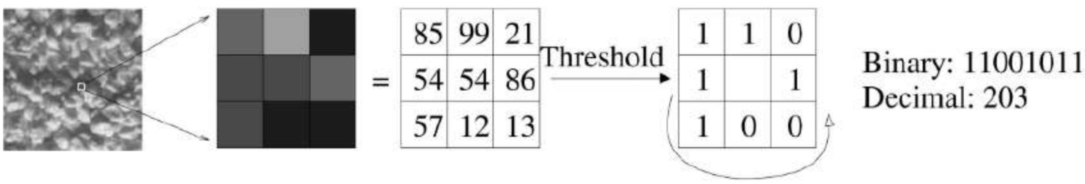

Figure 3.1: The basic LBP operator.

- Three Patch LBP (TPLBP) [87] was proposed to produce a single bit value in the code assigned to each pixel by comparing the values of three patches. For each pixel in the image, we consider a $w \times w$ patch centered on the pixel, and S additional patches distributed uniformly in a ring of radius r around it (Figure 3.2). For a parameter α, we take pairs of patches,$\alpha-$ patches apart along the circle, and compare their values with those of the central patch. The value of a single bit is set according to which of the two patches is more similar to the central patch. The resulting code has S bits per pixel. Specifically, the Three-Patch LBP is produced by applying the following formula to each pixel:

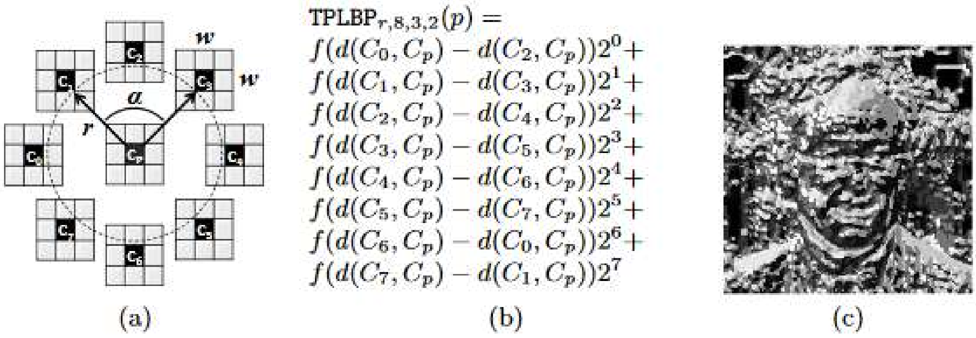

Figure 3.2: (a) The Three-Patch LBP code with $\alpha = 2$ and S=8. (b) The TPLBP code computed with parameters S= 8, w= 3, and $\alpha = 2$. (c) Present an example of LBP encoding (different intensities representing different codes) [87].

$$TPLBP_{r,S,w,\alpha}(p) = \sum_{i}^{S} f(d(C_i, C_p) - d(C_{i+\alpha mod S}, C_p))2^i \tag{3.11}$$

- Four Patch LBP (FPLBP) [87] is an operator assigns to compare two center symmetric patches in the inner ring with two center symmetric patches in the outer ring positioned patches away along the circle (say, clockwise). One bit in each pixels code is set according to which of the two pairs being compared is more similar. Thus, for S patches along each circle we have $S/2$ center symmetric pairs which is the length of the binary codes produced. The formal definition of the FPLBP code is as follows:

$$\begin{aligned} FPLBP_{r1,r2,S,\omega,\alpha}(p) = \sum_{i}^{S/2} & f(d(C_{1i}, C2, i + \alpha mod S) \\ & - d(C_{1i+S/2}, C2, i + S/2\alpha mod S))2^i \end{aligned} \tag{3.12}$$

  For every pixel in the image, we look at two rings of radii r_1 and r_2 centered on the pixel, and S patches of size $\omega \times \omega$ spread out evenly on each ring (Figure 3.3).

- Local Phase Quantization (LPQ) [88] was proposed by Ojansivu et al. [89] as a descriptor for texture which is robust to image blurring. The LPQ encodes the local phase information of four frequencies of the short-term Fourier transform

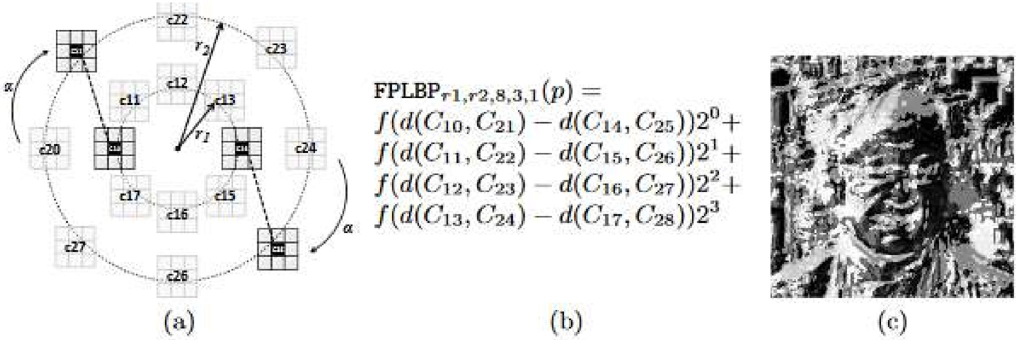

Figure 3.3: (a) The Four-Patch LBP code. Four patches involved in computing a single $\alpha = 1$ bit value with parameter are highlighted. (b) The FPLBP code computed with parameters S= 8, w= 3, and $\alpha = 1$. (c) Present an example of LBP encoding (different intensities representing different codes) [87].

(STFT) over a local window. The corresponding LPQ function is defined as in Equation 3.13:

$$F(u, x) = \sum_{y} f(x - y) \exp(-j2\Pi)u^T y) =^T {}_u W x_f \tag{3.13}$$

Where W_u is the basis vector of the 2-D Discrete Fourier Transforms (DFT) at frequency u and f_x is another vector containing all M^2 image samples from N_x. Figure 3.4 illustrated the basic LPQ operator.

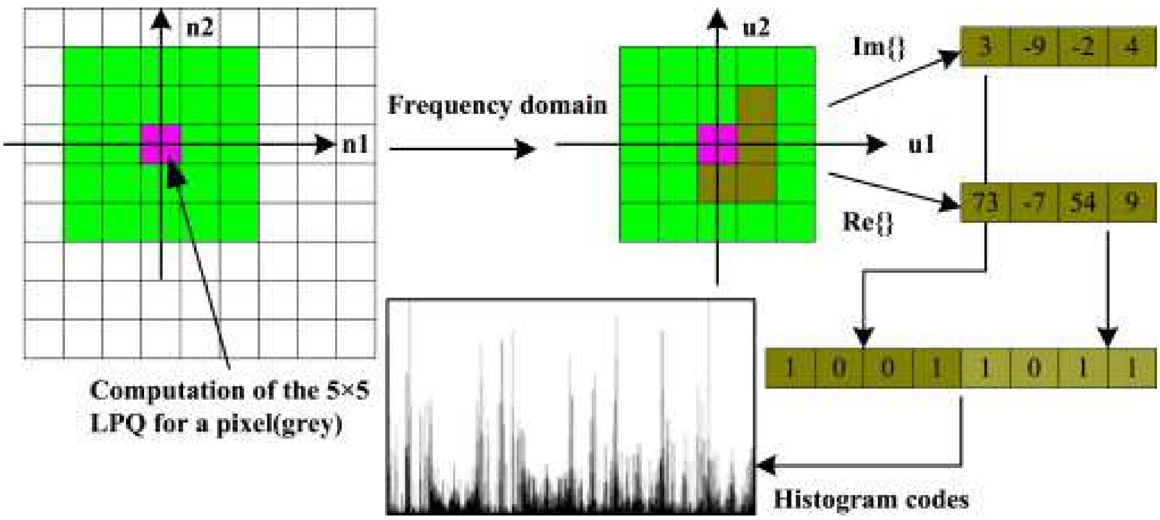

Figure 3.4: The basic LPQ operator [88].

- Binarized Statistical Image Features (BSIF) [90] is a method based on efficient scalar quantization scheme and independent component analysis to construct a local texture descriptors. Given an image patch X of size $(l \times l)$ pixels and a linear filter Wi of the same size, the filter response si is obtained as in equation 3.14

$$s_i = \sum_{u,v} W_i(u,v)X(u,v) = W_i^T x \tag{3.14}$$

Where vector notation is introduced in the latter stage. Given n linear filters W_i, we stack them into a matrix W and compute all responses at once $S = W_x$ Next, given a random sample of natural image patches, we compute the filters W_i so that the elements s_i of s are as independent as possible when considered random variables.

- Histograms of Oriented Gradient (HOG) have been proposed for face recognition purposes [67] [92]. The essential thought behind the HOG descriptors is that local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. The implementation of these descriptors can be achieved by dividing the image into small connected regions, called cells, and for each cell compiling a histogram of gradient directions or edge orientations for the pixels within the cell. The combination of these histograms then represents the descriptor. The gradient magnitude, GM, and the gradient angle, GA, are computed by

$$G_M = \sqrt{G_X^2 + G_Y^2} and G_A = atan(G_Y, G_X) \tag{3.15}$$

Where G_X is the image gradients in the horizontal directions and G_Y is the image gradients in the vertical directions. The gradient orientations, or angles, for each pixel are used to select the histogram bin. Figure 3.5 shown a simplified procedure for HOG features.

- Gabor features [93] Gabor filter is one of the best-known tunable filters, which is appropriate for capturing orientation information from the image. Gabor filters (Gabor wavelets) can be used to extract components corresponding to different scales and orientations from images. In [94], the Gabor filter-based features are directly extracted from the gray-level images. In the spatial domain, a two-dimensional Gabor filter is a Gaussian kernel function modulated by a complex sinusoidal plane wave, defined as:

$$G(x,y) = \frac{f^2}{\pi\gamma\eta} \exp(\frac{-(x'^2 + \gamma^2 + y'^2)}{2\sigma^2}) \exp(j2\pi f x' + \phi) \tag{3.16}$$

$$x' = xcos\theta + ysin\theta \tag{3.17}$$

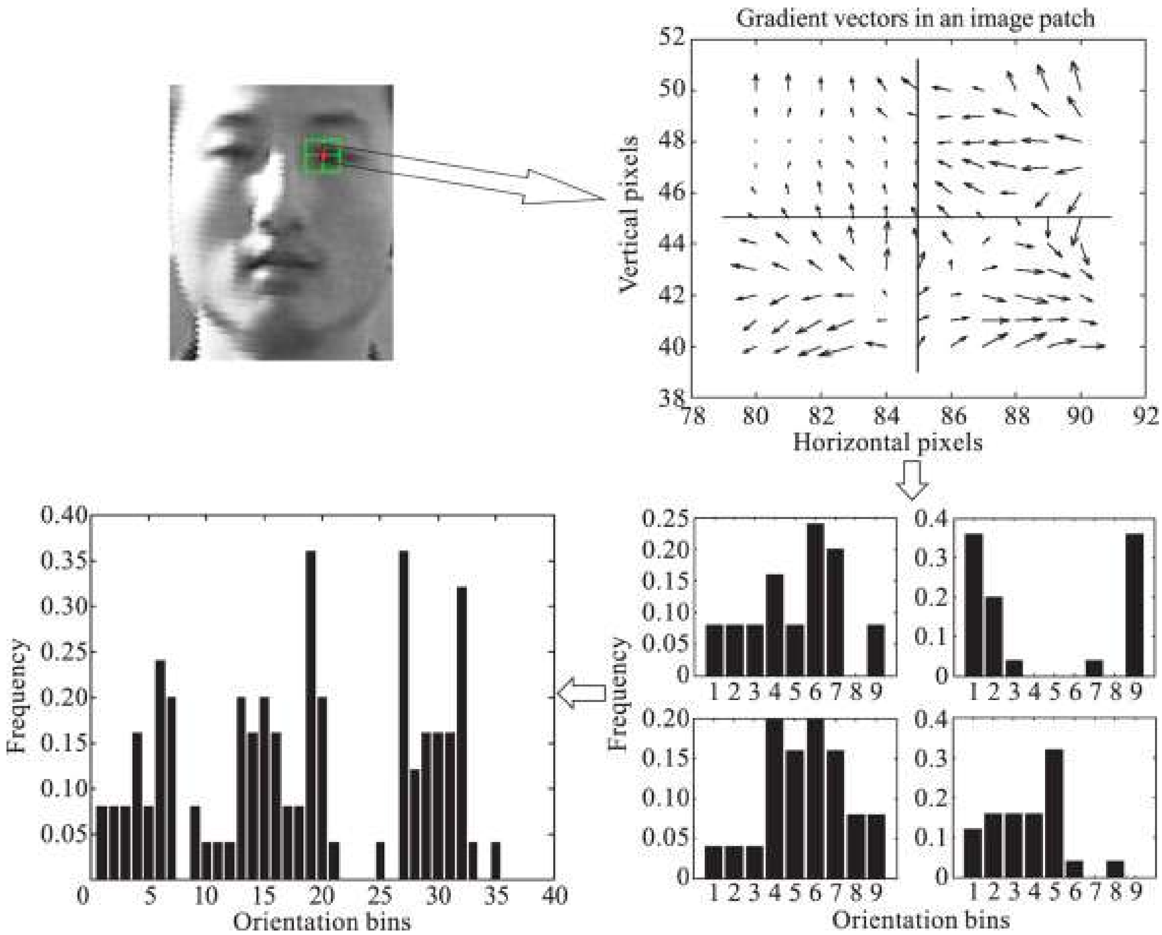

Figure 3.5: A simplified procedure for HOG features [92].

$$y' = -xsin\theta + ycos\theta \tag{3.18}$$

Where f represents the frequency of the sinusoid, h is the orientation of the normal to the parallel stripes of a Gabor function, ϕ is the phase offset, σ is the standard deviation and γ is the spatial aspect ratio.

3.3 Face Representation

- Multi-Block (MB) [91] is a variant that replaces intensity values in the computation of descriptor with the mean intensity value of image blocks. The MB is a technique that divides the face into $(n \times n)$ blocks. On each block, we apply a texture descriptor to get more features of the face.

- Multi-Level (ML) representation is a technique that combines the features extracted from consecutive different MBs. In other terms, we extract features of the whole image, and then we divide it into different blocks of different sizes and extract features of each block. The entire features are then concatenated into one vector.

Figure 3.6 explains the Multi-Level (ML) approach.

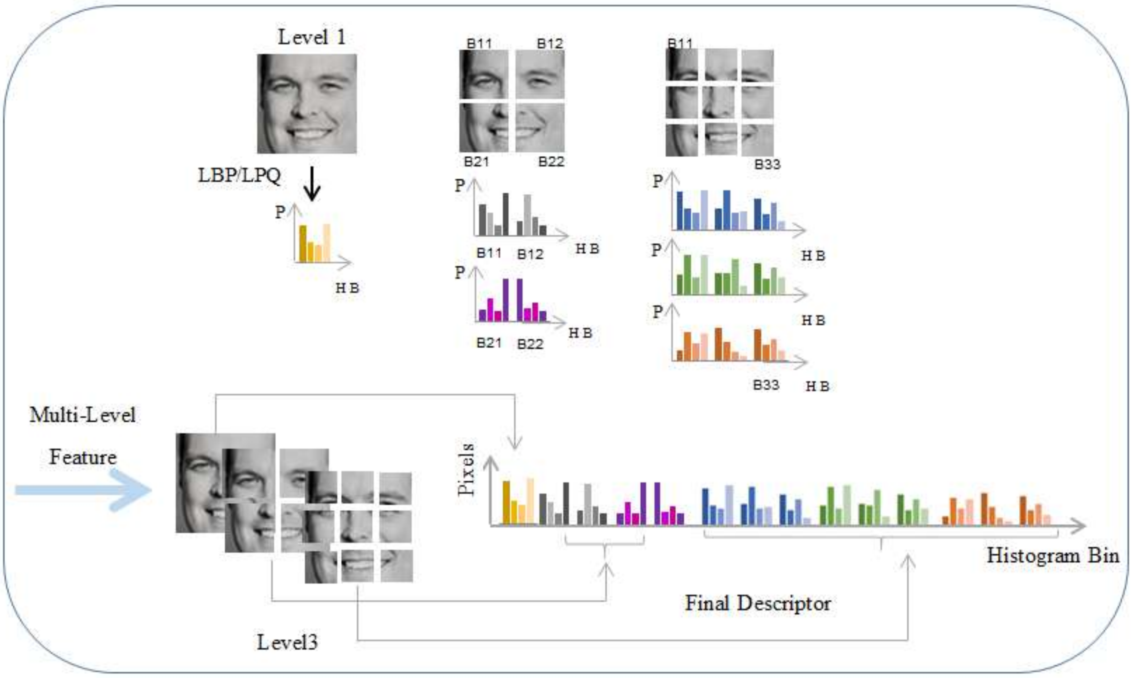

Figure 3.6: Example of multi-level features extraction with ($n = 3 \times 3$ sub-blocks).

3.4 Features extractions based on learned (No Hand-Crafted) features

3.4.1 Context-Aware Local Binary Feature Learning

In this thesis, we propose a new local binary feature learning method for kinship verification. Local binary descriptor constitutes power visual cues for feature description and classification. It provides discriminating information about small appearance details in local neighborhoods. So, it is robust to local changes databases such as illumination, identity, geometric distortions and transformations, expression, age, and occlusion. Unlike existing local descriptors are not discriminatory enough to estimate the relationship of face images because are hand-crafted features, which previous knowledge is required. To better utilize feature descriptors for facial kinship verification, we proposed the Context-Aware Local Binary Feature Learning (CA-LBFL)

[95] method to learn contextual features from raw pixels directly. It is applied to eliminates the dependence on hand-crafted features (extract and represent features automatically instead of selected manually).

After the success of local binary learning in face image verification, the CA-LBFL is the new variant of a local binary learning method. Thus, the block diagram of CA-LBFL algorithm presented in Figure 3.7 can be described as follows:

Pixel Difference Vectors (PDV): Given any pixel in the image, first, we compute the differences between the central pixel and its $(2R+1)\times(2R+1)$ neighboring pixels, where R is the size of neighborhood. Then, these differences are aligned as a vector which becomes the PDV feature of the pixel. We select 8 neighboring pixels, so, The PDV is a 8-dimensional vector.

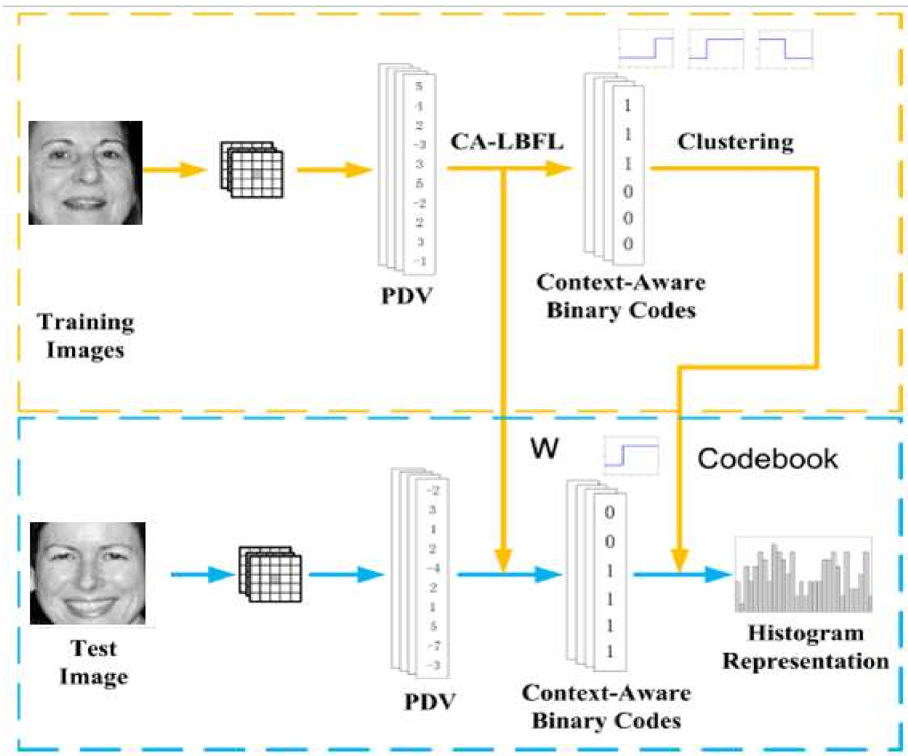

Figure 3.7: The block diagram of CA-LBFL [95].

Projection Matrix W: We learn K hash functions to obtain context-aware binary codes, we map and quantize each x_n into a binary vector $b_n = [b_{1n}, ..., b_{Kn}]^T \epsilon 0, 1^{K\times1}$. Let $w_k \epsilon R^d$ be the projection vector for the k_{th} function, and the k_{th} binary code b_{kn} of x_n can be computed as:

$$b_{kn} = 0.5 \times (sgn(w_k^T x_n) + 1) \tag{3.19}$$

where sgn(v) equals to 1 if $v \geq 0$ and -1 otherwise.

To obtain the binary codes for feature representation, we formulate the following optimization objective function:

$$\begin{aligned} min_{wk} J = J_1 + \lambda_1 J_2 + \lambda_2 J_3 + \lambda_3 J_4 \\ = \sum_{n=1}^{N} || \sum_{k=1}^{K-1} ||b_{kn} - b_{(k+1)n}||_2 - 1||^2 \\ + \lambda_1 \sum_{n=1}^{N} \sum_{k=1}^{K} ||(b_{kn} - 0.5) - w_k^T X_n||^2 \\ + \lambda_2 \sum_{k=1}^{K} \sum_{n=1}^{N} ||(b_{kn} - 0.5||^2 \\ - \lambda_3 \sum_{n=1}^{N} \sum_{k=1}^{K} ||(b_{kn} - \mu_k||^2 \end{aligned} \tag{3.20}$$

Where N is the PDVs number which is obtained from the original images, the mean of the k_{th} bit of all N PDVs is μ_k. The three parameters λ_1, λ_2 and λ_3 are used to equilibrate the weight of variant terms. In equation 3.20, the physical meaning of $\sum_{k=1}^{K-1} ||b_{kn} - b_{(k+1)n}||^2$ is the sum of bitwise 0/1 changes in each binary vector.

The minimization of J_1 executes the contextual information and makes the codes more robust to the noises. J_2 aims to minimizes the loss of energy in the process of projection which reduce the loss of quantization between the binary codes and the original features. J_3 is to evenly distribute the feature in the binary codes. J_4 project the vector as independent as possible by maximize the variance of the binary codes. We can mapped the projection matrix W and each sample x_n into a binary vector as follows:

$$0.5 \times (sgn(W^T x_n) + 1) \tag{3.21}$$

Then, (2) can be re-written into the matrix form as:

$$\begin{aligned} min_w J = J_1 + \lambda_1 J_2 + \lambda_2 J_3 + \lambda_3 J_4 \\ = tr(((AB)^T (AB) - I_N)^2) \\ + \lambda_1 ||(B - 0.5) - W^T X||_F^2 \\ + \lambda_2 ||(B - 0.5) \times 1^{N \times 1}||_F^2 \\ - \lambda_3 tr((B - U)^T (B - U)) \end{aligned} \tag{3.22}$$

Where $W = [w_1, w_2, ...w_K] \in R^{d\times K}$, the matrix of all binary codes B defined as $B = 0.5(sgn(W^TX + 1) \in \{0, 1\}^{K\times N}$, $U \in R^{K\times N}$ is the mean matrix repeating the row vector of the mean of all binary bits, I_N is the identity matrix, and they minimize the difference between adjacent bits in binary codes using matrix $A\epsilon 0, 1, -1^{K-1} \times K$ as follows:

$$\begin{cases} 1, & \text{i=j;} \\ -1, & \text{i=j-1;} \\ 0, & \text{otherwise.} \end{cases} \tag{3.23}$$

The element of the matrix A is a_{ij}, and the indices are i and j. The differences between all the adjacent bits in learned binary codes are represented in matrix AB. and the diagonal of $(AB)^T(AB)$ is the sum of bitwise changes. Thus, J_1 can be rewritten as follows:

$$\begin{aligned} J_1(W) = tr(((AW^TX)^T(AW^TX) - I_N)^2) \\ = tr(X^TWA^TAW^TXX^TWA^TAW^TX) \\ - 2 \times tr(X^TWA^TAW^TX) + N \end{aligned} \tag{3.24}$$

Similarly, we rewrite $J_3(W)$ and $J_4(W)$ as:

$$\begin{aligned} J_3(W) = ||(W^TX - 0.5) \times 1^{N\times 1}||_2^2 \\ = tr(W^TX_1^{N\times 1}1^{1\times N}X^TW) \\ - N \times tr(1^{1\times K}W^TX1_{N\times 1}) \\ + 0.25 \times 1^{1\times N}1^{N\times 1} \end{aligned} \tag{3.25}$$

$$\begin{aligned} J_4(W) = tr(W^TXX^TW) - 2 \times tr(W^TXM^TW) \\ + tr(W^TMM^TW) \end{aligned} \tag{3.26}$$

where $M\epsilon Rd \times N$ is the mean matrix which are repeated row vector of the mean of all PDVs. Therefore, they optimize W and B using the following iterative approach. Obtaining B with a fixed W: when W is fixed, equation 3.27 can be written as follows:

$$min_B J(B) = ||(B - 0.5) - W^TX||_F^2 \tag{3.27}$$

As B is a binary matrix, the solution can be directly obtained as:

$$B = 0.5 \times (sng(W^T X) + 1) \tag{3.28}$$

Learning W with a fixed B: when B is fixed, the equation 3.29 can be written as follows:

$$\begin{aligned} min_W J(W) = tr(&X^T W A^T A W^T X X^T W A^T A W^T X \\ &- 2 \times (tr(X^T W A^T A W^T X) + tr(W^T Q W) \\ &- 2 \times \lambda_1 tr((B - 0.5) \times X^T W) \\ &- \lambda_2 \times N \times tr(1^{1\times K} W^T X_1^{N\times 1}) \end{aligned} \tag{3.29}$$

subject to$W^T W = I$. Where

$$\begin{aligned} Q = \lambda_1 X X^T + \lambda_2 X_1^{N\times 1} 1^{1\times N} X^T \\ - \lambda_3 (X X^T - 2 X M^T + M M^T) \end{aligned} \tag{3.30}$$

Unsupervised Clustering. The conventional K-means (unsupervised clustering) is used to learn the codebook to represent all binary codes as a histogram features.

3.4.2 Discrete Cosine Transform Network (*DCTNet*)

Deep learning is proposed to eliminate the dependence on handcrafted features and learn features from pixels directly. Deep learning has been recently outperforming state-of-the-art in various applications in general and faces recognition in particular, [96], [97], [98]. However, deep learning recognition methods automatically learn, understand and design features from an image manually. It requires recognizing visual patterns from pixels directly with minimal pre-processing [98], it provides robustness to geometric distortions and transformations, and other 2-D shape variations, such as illumination, pose, and scale. Our proposed approach using Discrete Cosine Transform Network (*DCTNet*) [99] works according to the following steps:

1. Filter Banks: Discrete Cosine Transform (DCT) adopting 2D-DCT basis as filter banks. This is motivated by the fact that 2D DCT basis is indeed a good

approximation for high ranked eigenvectors. The transformation of 2D-DCT can be extended from 1D-DCT as following:

$$f(x,y,u,v) = C(u)C(v)cos(\frac{(2x+1)u\pi}{2N})cos(\frac{(2y+1)v\pi}{2N}) \tag{3.31}$$

Where

$$C(u) = \begin{cases} \sqrt{\frac{1}{N}}, for & \text{u=0;} \\ \sqrt{\frac{2}{N}}, for & \text{u=1,2,....N-1.} \end{cases} \tag{3.32}$$

Note that C(v) is defined the same way as in Equation 3.32.

2. Convolution layers: The network is composed of 2 convolution layers. Since DCT basis of each layer can be combined to form a flat single layer network. The last convolution layer of DCTNet forms the real-valued outputs. The output of input image I_d of size nm with D channels given by:

$$O_d^p = \{I_d * W_p^l\}_{p=1}^{p_l} \tag{3.33}$$

where $*$ denotes the discrete convolution and the size of output O_d^p is same as I_d and $W_p^l \in R^{k \times k}, p = 1, 2, ..., P_l$ is 2D-DCT bases from P_l filters at layer l. Of each layer, we have d outputs.

3. Binarization and Block-wise Histogramming: Binarization is performed on each set separately by first binarising the responses with a threshold at zero [99]. In this stage, the real value of the last convolution layer is turned into a binary format by thresholding to zero (value 1 for a positive response, zero otherwise) denoted by BIN:

$$BIN(O_d^p) = \begin{cases} 1 & \text{if } O_d^p \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.34}$$

Where $BIN(O_d^p)$ is a binary image. Then, each of these binarized images is partitioned into non-overlapping blocks. The characteristics of these images are obtained by concatenating all the histograms of each block B such as:

$$H = \{H_b^d\}_{b=1,d=1}^{B,D} \tag{3.35}$$

Where $b = 1, 2, ..., B; d = 1, 2, ..., D$

The combination of binarization and block-wise histogramming is expected to be able to extract discriminative features.

4. Histogram Tied Rank Normalization (*TR* normalization): is applied to eliminate the disparity of histogram vector. Based on the tied rank principle used by robust statistic and intra-normalization. The input of *TR* normalization is $\boldsymbol{H}$ the extracted block-wise histogram of an image and the output is $\boldsymbol{v}$ *TR* normalized histogram feature vector. Given H as the extracted histogram where $H = H_{b=1,d=1}^{B,D}$.

   The histogram *TR* normalization is presented in Algorithm 1.

Data: Extracted block-wise histogram of an image: H
Result: *TR* normalized histogram feature vector: v
1 Compute $\overline{H_b^d}$.;
2 Calculate v_b^d.;
3 Normalize v_b^d.;
4 Concatenated all v_b^d. $v = [\hat{v}_1^1, \hat{v}_2^1, ..., \hat{v}_B^1, \hat{v}_1^2, ..., \hat{v}_B^D] \epsilon R^{(2^P L)BD}$.;

Algorithm 1: Histogram *TR* Normalization

In order to give a complete description of the histogram *TR* normalization methodology, we detail each step of the previous algorithm:

(a) For each H_b^d we compute tied rank without bin with zero occurrence yields $\overline{H_b^d}$. This is because bin with zero occurrences is not a sample in histogram. It should be ignored in the ranking process.

(b) Calculate v_b^d where $v_b^d = \sqrt{\overline{H_b^d}}$.

(c) Normalize v_b^d with L_2 norm to obtain $\hat{v_b^d}$.

(d) The final TR normalized histogram feature vector is obtained by Concatenate all $\hat{v_b^d}$. Where $v = [\hat{v}_1^1, \hat{v}_2^1, ..., \hat{v}_B^1, \hat{v}_1^2, ..., \hat{v}_B^D] \epsilon R^{(2^P L)BD}$.

The block diagram of DCTNet algorithm presented in Figure 3.8.

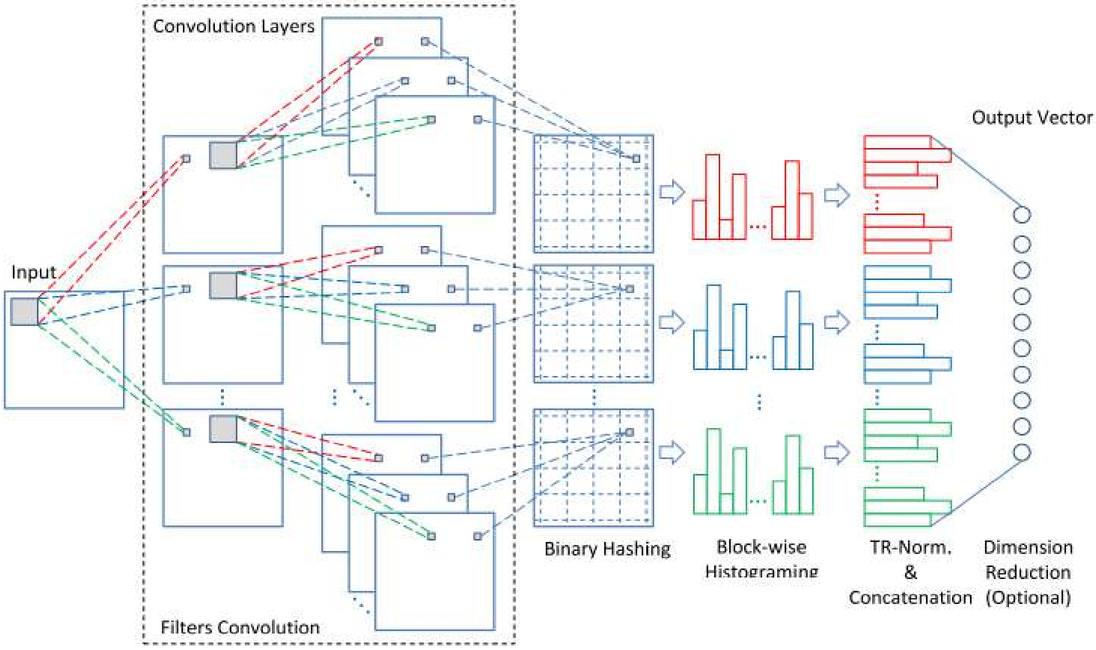

Figure 3.8: The block diagram of DCTNet [99].

3.5 Classification

After the feature representations are adopted, the next step is image classification. The choice of the classifier has a great impact on the performance or recognition accuracy of the kinship verification task.

In terms of biometrics, classification means finding a proper identity for the query [72]. Generally, image classification is performed by calculating the similarity between a target discriminating vector of feature and a query discriminating vector of feature [100]. In kinship verification, the facial images are classified into two classes. The first class is the true pairs of parents and children and the second one is the false pairs of parents and children.

In this work, two classifiers are used, the Support vector machines (SVM) and metric learning.

3.5.1 Support vector machines

In machine learning, Support vector machines (SVM) is a supervised classification, regression and outliers detection algorithm. They were originally proposed by Vapnik [101] and improved by Vapnik and Corinna Cortes [102].

SVM performs classification by finding the optimal separating hyperplane which maximizes the margin (the distance of closest data, regardless its class, to the hyperplane) of the high dimensional training data. The training feature vectors along with their labels are input to SVM, which outputs a model able to predict the labels of new unseen data.

In other words, the operation of the SVM algorithm is based on finding the hyperplane that gives the largest minimum distance to the training examples. Twice, this distance receives the important name of margin within SVMs theory. Therefore, the optimal separating hyperplane maximizes the margin of the training data. Figure 3.9 illustrate the SVM idea.

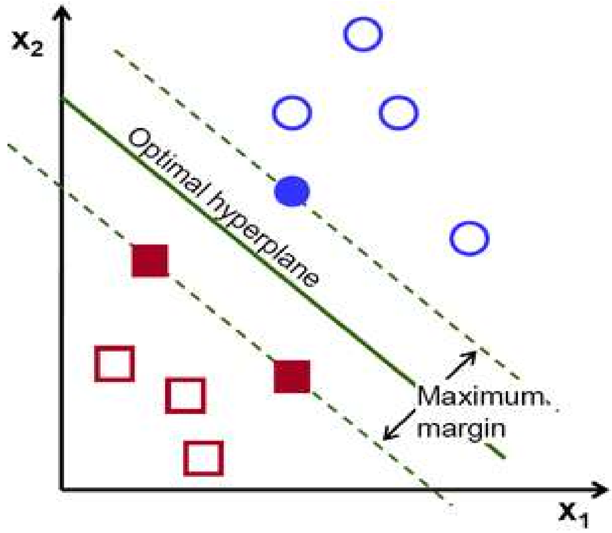

Figure 3.9: SVM example.

In this work, a Library for Large Linear Classification (LIBLINEAR) [103] is used where the pair of input faces is classified as positive (the true pairs of parents and children) or negative (the false pairs of parents and children). LIBLINEAR is an open source library for large-scale linear classification. It supports linear support vector machines and implements the one-against-one approach.

In our experiments, a linear support vector machine with 0/1 labels based on the kinship relationship is used for bi-classification. The five fold cross-validation strategy is applied to the training set to evaluate the performance of the system. However, difference SVMs method is very sensitive to chosen parameters. The C parameter of the SVM are set by sampling on a grid.

In SVM, we look for two things i) setting a larger margin, ii) lowering misclassification rate. For large value of parameter C will cause a small margin hyperplane and small values of C will cause a large margin hyperplane.

Before training the SVM machine, we convert a pair of features into a single feature vector as required by the classifier. We have tried different techniques of combining a pair of features, such as concatenation and vector distances. We have empirically found that utilizing the normalized absolute difference shows the best performance. The corresponding normalized absolute difference function is defined as in Equation 3.36

$$f_i = \sum_i \frac{| x_i - y_i |}{\sum_i (x_i + y_i)} \tag{3.36}$$

Where $X = x_1, ..., x_n$ and $Y = y_1, ..., y_n$ are pair of feature vectors and represented by the vector $F = f_1, ...f_d$. Figure 3.10 depicts the basic schema for facial kinship verification based on SVM classifier.

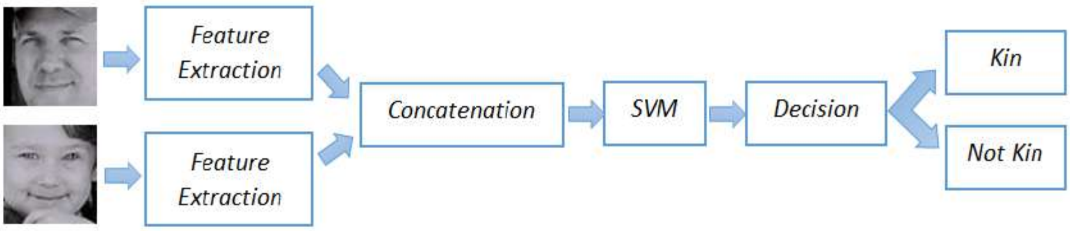

Figure 3.10: Kinship verification flowchart based on SVM classifier.

3.5.2 Metric Learning

In computer vision and machine learning, metric learning has received a lot of attention in last years. Number of metric learning algorithms has been proposed in the literature, they concerned to learn a distance function. The metric learning methods aim to seek an appropriate distance metric for classification tasks and have achieved a good performance in many facial image analysis applications.

In this work, we present a simple method that requires no training data, we compute the distance between the pairs of features with different functions and use the result as a score for deciding whether the pair is a kin or not. Several following distances are used in our case :

City block distance (L_1) [104]: City-Block distance or Manhattan distance, also called, L_1 distance and L_1 norm. It represents the distance between points in a city road grid. It examines the absolute differences between the coordinates of a pair of objects as follows:

$$d_{L_1} = \sum_i |x_i - y_i| \tag{3.37}$$

Euclidean distance (L_2) [105]: The Euclidean distance is the most used distance in literature. It based on calculating the distance between two points in Euclidean space by examining the root of squares difference between a coordinate of a pair of the image.

$$d_{L_2} = \sqrt{\sum_i (x_i - y_i)^2} \tag{3.38}$$

Chi-Square distance (χ^2) [106]: The chi-squared distance is a nonlinear metric and is widely used to compare histograms. Chi-Square distance can be defined as the distance between two histograms.

$$d_{\chi^2} = \sum_i \frac{(x_i - y_i)^2}{x_i + y_i} \tag{3.39}$$

Cosine distance (cos) [107]: Cosine similarity is a measure of similarity between two vectors and can be computed as:

$$d_{cos} = \sum_i \frac{x_i y_i}{\sqrt{\sum_i x_i^2}\sqrt{\sum_i y_i^2}} \tag{3.40}$$

Where x and y are the two feature vectors of the pair of images being compared.

Figure 3.11 depicts the basic schema for facial kinship verification based on Metric Learning classifier.

Figure 3.11: Kinship verification flowchart based on Metric classifier.

3.6 Feature selection

In machine learning, Feature Selection (FS) is a data preprocessing step that is applied after feature extraction step. Feature selection method aims to select the most relevant subset of extracted features to provide a useful classification with the smallest error and to discard non-relevant features. In this thesis, we used Fisher/Correlation feature selection using Wrapper for Spider toolbox [108], [109] and [110].

Spider Wrapper is a library of objects in Matlab. It is meant to handle large unsupervised, supervised or semi-supervised machine learning problems.

The fundamental idea of choosing Fisher/Correlation (F/C) [111] score for kinship verification is to find a subset of characteristics, the distances between data points in same classes (positive pairs) are as small as possible, while the distances are as large as possible between data points in different classes (negative pairs).

Let I = $(x_i, y_i)|i = 1...N$ be the training data of pairs of parent and child images, where x_i is the vector of the ith parent and y_i of the ith child, and L the class of this data. Each x_i and y_i are ranked with Fisher/Correlation algorithm, next the size of the optimal (F/C) set is heuristically found.

The proposed feature selection method is presented in Algorithm 2.

Data: Input images, labels and the number of features to be selected.
Result: The selected features.
1 Trasformation the pair features into a single feature vector using U.;
2 Calculate S.;
3 Output the features selected.;

Algorithm 2: Feature Selection Algorithm.

In order to give a complete description of Feature Selection methodology, we detail each step of the previous algorithm: The input of the algorithm are: data $I = (x_i, y_i)|i = 1...N$, labels of data and the number of features to be selected.

First step. It based on normalized absolute difference U to transform the pair features into a single feature vector.

$$U = \sum_i \frac{|x_i - y_i|}{\sum_i (x_i + y_i)}.$$

Where x are the parents images and y are the children images.

Second step. Select the most relevant descriptors of the input features based on Fisher score S.

$$S_i = \frac{\sum_{k=1}^{K} n_j(\mu_{ij} - \mu_i^2)}{\sum_{k=1}^{K} n_j \eta_j \rho_{ij}^2}.$$

Where μ^j is the mean, σ^j denote the standard deviation of the whole data set and K= number of selected features.

Fisher score is one of the most widely used supervised feature selection methods. However, it selects each feature independently according to their scores under the Fisher criterion, which leads to a suboptimal subset of features.

We calculates the Fisher score for each feature by returning the fisher object initialized with parameter (K: numSelect; the number of descriptors that are selected for each feature). In our experiments we change the K in order to identify the best performance for the algorithm.

Result. The output of the algorithm is features selected.

3.7 Performance evaluation

In machine learning research, Receiver Operating Characteristic (ROC) [112] is an accepted method and widely used to show the effectiveness of a recognition system. ROC curve, can defined as a function of the decision threshold, it created by plotting the rate of true positive (ie. image classified correctly as positive pairs or negative pairs), against the rate of false positive (ie. image classified incorrectly) at various threshold settings. An illustration of a ROC curve is presented in Figure 3.12. Each point on the ROC curve represents a positive/ negative pair corresponding to a particular decision threshold. therefore, our evaluation performance is based on ROC curve and accuracy. More specifically:

$$Accuracy = \frac{TP + TN}{P + N} \tag{3.41}$$

Where TP is the number of image that classified correctly as positive pairs, TN is the number of image that correctly classified as negative pairs, P is the number of all positive pairs, and N is the number of all negative pairs.

3.8 Fusion Process

Multimodal systems overcome many practical problems that occur in single modality biometric systems, such as noisy sensor data, non-universality and/or lack of distinctiveness of a biometric trait, unacceptable error rates, and spoof attacks, by consolidating multiple biometric information pertain to the same identity [113]. Biometric fusion can be implemented at various levels, such as raw data level, image level, feature level, rank level, score level, decision level and color may be seen as a fusion process too.

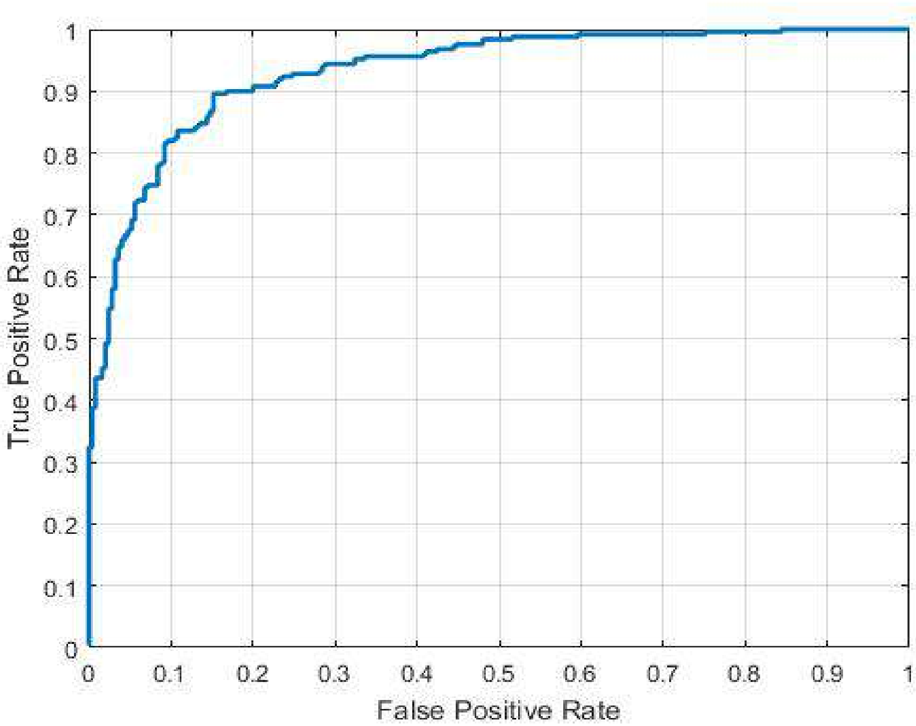

Figure 3.12: Example of ROC curve: true positive rate against false positive rate.

In this work, we proposed two types of fusion strategies, namely, score fusion [114] and color fusion [71].

3.8.1 Score Fusion

The fusion score level is to combine scores obtained from different features. In our work, we have combined two features which produce the best score, in order to combine the decisions of several features into one final decision and to achieve higher kinship verification accuracy. Therefore, a simple sum rule at the score-level is opted to perform the fusion.

3.8.2 Feature Fusion

The color-based method is proposed to explore the role of color features and to encode both the luminance and the chrominance. In this work, we used the psycho-visual feature (HSV). However, The Hue-Saturation-Value (HSV) color space is motivated by human vision system, in which the light intensity (V channel) is separated from the color tonality (H channel) and the saturation of the color (S channel).

The following steps covers the color fusion experiments:

1. A pair of face images is given as an input. These two images were first splitting into different color spaces HSV.

2. Then, from each channel of the considered color space H, S and V we extracted the features (e.g., LBP).

3. The features were then concatenated to form one enhanced feature vector.

4. Finally, we applied classifiers between the feature vectors of the pair of the two face images.

3.9 Conclusion

In this chapter, different features extraction and matching methods are proposed to deal with kinship verification problem. Firstly, we discussed the methods based on Hand-Crafted features and approaches based on learned (No Hand-Crafted) features.

The main advantage of the first method is constituted of power visual cues for feature description. They provide discriminating information about small appearance details in local neighborhoods. So, they are robust to local changes databases such as illumination, identity, geometric distortions and transformations, expression.

The second method introducing the *CA-LBFL* and the *DCTNet*. The *CA-LBFL* is applied to learn contextual features from raw pixels directly. Furthermore, the (*DCTNet*) is applied to extract and represent features automatically where numerous layers of information are exploited.

Secondly, we discussed a supervised methods that called (*SVM*). SVM performs classification by finding the hyperplane that gives the largest minimum distance to the training examples.

The second method present a simple approach that requires no training data by computing the distance between pairs of features with different functions and use the result as a score for deciding whether the pair is a kin or not. Several distances are used (the City block distance, the Euclidean distance, the Chi-Square distance and the Cosine distance.)

The rest of the chapter, provided a brief discussion of the tools and techniques used. Functionality of feature selection algorithm is explained. Also, we explained the different fusion techniques and its related strategies. We presented the details of performance evaluation measurement throughout the thesis.

Chapter 4

Results and Discussions

This Chapter contains:

4.1 Introduction

In this chapter, we develop different studies to employ the ideas introduced throughout this thesis on the facial kinship verification.

We represent the experimental results obtained by the tests we have performed of facial kinship verification. We investigate several features characteristics: texture-based methods, the global-based methods, and learning methods. The first methods includes LBP, TPLBP, FPLBP, HOG, Gabor, BSIF, and LPQ. The second methods includes PCA, LDA, and LPP. The learning methods include DCTNet and the CA-LBFL. We conduct intensive experiments on different existing kinship databases. The results are analyzed and compared and interesting conclusions are deduced.

4.2 Experimental 1

We will start by investigation several local feature extraction and similarity computation approaches and we will address many challenges: diverse ages, expressions, gender, skins, lighting changes even dramatically illumination. Also, we will explore the feature combination to perform multiple features fusion to extract complementary information to improve the kinship verification performance.

4.2.1 Experimental Evaluation and Databases

In this experimental, we investigated various feature extraction and similarity computation approaches for the problem of kinship verification from facial images. We conduct intensive experiments on different existing kinship databases. The results are analysed, compared and interesting conclusions are deduced.

We analysed the performance of five texture descriptors (LBP, TPLBP, FPLBP, HOG and Gabor wavelets) on four kinship datasets that are publicly available: *KinFaceW-I*, *KinFaceW-II*, *Cornell Kin Face* and *UB KinFace V2*.

LBP and HOG are selected for their popularity in computer vision and showed very promising results in different problems. TPLBP and FPLBP are chosen to change the scale of LBP and provide other categories of local information.

Gabor is one of the best known tunable filters. Moreover, we address the problem with a simple setting: For each face image, *PCA* is first used to project feature and then the cosine similarity of each test pair is computed. Finally, error normalization is employed.

For the experiments, each face is cropped using Viola Jones detector and then resized to 64×64 pixels and the gray histogram is equalized. Examples of cropped image are shown in Figure 4.1.

Then same number of negative pairs (no kin relation) as positive pairs (with a kin relation) is created by associating faces from persons which do not have a kinship relation. We apply the Principal component analysis (PCA) for the dimensionality reduction of feature vectors. For each of the four kinship subsets (i.e. F-D, F-S, M-D and M-S), we recommend taking the feature normalization technique to improve

Figure 4.1: Examples of cropped face image of KinFaceW-I & II database.

recognition rate. The result of normalization is that the features will be rescaled so that they all have properties of a standard normalization distribution.

We perform 5-fold cross validation and compute the mean accuracy of the five folds. We report the mean verification accuracy of the four kinship subsets as the measure of system performance.

4.2.2 Feature Extraction

LBP, TPLBP, FPLBP: The face image is divided into 4×4 non overlapping blocks and the size of each block is 16×16 and we applied the LBP, TPLBP, FPLBP for each block. The face is represented by concatenating the vectors from all the blocks.

HOG: HOG features are quantized with 9 histogram bins and the number of HOG windows are (3×3).

Gabor: The parameters of Gabor filters were empirically determined for the face images. These were set as orientation v = 8 and scales u = 5. The feature vectors are normalized to zero mean and unit variance.

4.2.3 Results and Analysis

In this work, we compared the performance of five texture description methods on kinship recognition, using the Support Vector Machines (SVM) and Nearest Neighbor (NN). Our system classification based on the determination of the similarity from two feature descriptors by measuring the score. We investigated the system in two modes: the unimodal and the multimodal system.

a) Unimodal system

In this context, we have performed two experiments to assess the performance of the proposed approach. In the following, we present the reported results.

Table 4.1: Mean kinship verification accuracy in % on different database.

database	Features	SVM	NN			
			L1	L2	χ^2	cos
KinFaceW II	LBP	63.20	53.65	53.25	53.35	75.85
	TPLBP	61.20	54.35	54.05	53.05	70.00
	FPLBP	58.59	53.25	53.05	53.15	69.80
	HOG	66.00	54.05	53.35	53.85	75.10
	GABOR	63.85	53.35	53.75	54.35	72.95
KinFaceW I	LBP	66.97	54.68	54.65	53.09	68.92
	HOG	67.94	53.90	53.92	54.81	67.38
	GABOR	62.65	58.36	53.49	54.95	65.04
UBKF Yong P	LBP	57.80	54.77	53.83	55.09	64.70
	HOG	58.00	52.88	55.14	52.26	63.25
	GABOR	55.00	54.80	52.87	53.55	59.90
UBKF Old P	LBP	58.75	55.08	54.14	54.79	59.75
	HOG	58.25	51.94	50.97	55.79	60.00
	GABOR	58.25	54.80	52.87	53.55	60.50
CornelKinFace	LBP	61.99	54.09	55.98	52.59	68.92
	HOG	59.19	52.96	55.23	55.23	64.90
	GABOR	60.85	55.21	54.12	52.25	62.65

Nearest Neighbor. We presented a simple method that requires no training data, we compute the distance between (parent-children) vectors, including L_1, L_2, χ^2, and cos.

Comparing distances: Table 4.1 shows the mean verification rates of different metric learning on different kinship datasets. Our results show that the average accuracy of human performance on the task of kinship verification is 50.97 - 75.85 % on the four data bases. With suitable descriptors the recognition rate of the Cosine distance exceeds the recognition rates of the other distance, while L_2 shows the worst performance.

Comparing features: The results for different features are reported in Table 4.1. The performances of the LBP show competitive results on different kinship database. The LBP descriptor can achieve a higher accuracy equal to 75.85 % at the *KinFaceW-II*. On the other hand, Gabor descriptor report the best performance on UBkinFace (Old Parent).

Comparing databases: the classification accuracy obtained by cosine distance method on *UB Kin Face* (young parents) reaches around 62.48%, this is comparable to *UB Kin Face* (old parents) who's reaches 60.08%. These results are maybe due to the different age of the pairs. The best verification accuracy is obtained on KinFaceW-II while the lowest are on *UB Kin Face* (old parents). These results are maybe due to the different Database environment.

Support vector machines To classify the parent-child pairs into two true and false categories, we used the binary linear classifier without including kernel, we first create training data by concatenate the (parent-child) pairs, the parameter of training model is $C = 1$. From Table 4.1, the best results is the HOG feature that produce 67.94% on KinFaceW-I.

b) Multimodal system

We have fused the LBP feature and the HOG feature (two best features) to check their complementarity. Our experiments based on score-level fusion.

From Table 4.2, the higher accuracy equal to 80.80% with $NN - cosine$ on *KinFaceW II*. Overall, the fusion enhanced the verification accuracy of unimodal system.

The experiment result shows that the cosine similarity is appropriate to measure the relationships of facial image and the LBP is more informative in learning this metric learning.

Table 4.2: Verification accuracy (in %) for feature fusion on different databases.

	KFW-I	KFW-II	UBK-Y	UBK-O	Cornell
SVM	68.01	69.00	61.96	60.75	60.71
NN-cos	70.47	80.80	67.96	65.75	68.71

From this experimental, we can see that there are several labels affecting the accuracy of an automatic kinship.

Ethnicity: The accuracy of human performance on the task of kinship verification on *KinFaceW II* (where the faces are cropped from the same images) have a high potential for biased data sets, which includes a majority of Caucasian individuals, the other datasets used are more balanced.

Age: The labels such as age and gender are helpful to determine the relationship. UB kinface included, for each child, images of his/her young and old parent, which makes it impossible to evaluate the classification rate.

Data base environment: The appearance of a face is affected by a large number of factors including identity, face pose, illumination, facial expression and resolution. The development of algorithms robust to these variations requires databases of sufficient size that include carefully controlled variations of these factors. Furthermore, common databases are necessary fair comparison.

Finally, we can see that: i) The pertinent features that help us to recognize the kinship relationship are: the genetic relatedness, age variation and data base environment; ii) The best features are obtained by the LBP algorithm which estimates the resemblance based on local special information and those features extended to detect the relation in different poses and orientations; iii) The best classifier is the NN which attributes a sample to the class with the closest samples according to some similarity measure.

4.3 Experimental 2

Because of the large variance in the facial appearance of the parent and children, kinship verification is very challenging. In general, a suitable face representation must be provided to deal with this challenge. In this experimental, various global and local face representations were introduced. The Global appearance-based methods try to find a suitable representation of the whole image. The Local texture descriptors represent certain region properties. Many different types of visual features derived from these measures were over-viewed. Furthermore, some of the ideas suggested overcoming the problem were summarized.

Also, we proposed a novel learning method for kinship verification consists of four main stages: 1) A DCTNet applied to each face image in order to extract the most significant inherited facial features through convolutional layers based on 2D DCT filter bank. 2) The response of the last layer is binarized and partitioned into non-overlapping block-wise histograms. 3) A Tied Rank (TR) Normalization is used to eliminate the disparity of histogram vectors of DCTNet. 4) The last stage is to distinguish between the different pairs. The distances between data points in same classes (positive pairs) are as small as possible, while the distances are as large as possible between data points in different classes (negative pairs).

Experiments are conducted on three public databases (UBKinFace, KinFaceW-I and KinFaceW-II). They show significant performance improvements compared to state-of- the-art methods.

Finally, we introduced the multimodal system based on some discriminative biological information to improve the kinship verification performance.

This investigation are published in papers C (1) 1.

4.3.1 Experimental Evaluation and Databases

We report our experimental investigations on the UBKinFace (set 1 and set 2), Kin Face in the Wild II (KinFaceW-II) and Kin Face in the Wild I (KinFaceW-I) datasets showing auspicious performance compared to state-of-the-art.

For the experiments, each face is converted into gray-scale image, then we cropped and resized to 32×32 pixels and the gray histogram is equalized, then the same number of negative pairs (no kin relation) as positive pairs (with a kin relation) is created by associating faces from persons who do not have a kinship relation.

To evaluate the performance of kinship verification, three step is performed: feature selection, dimensionality reduction and classification. Firstly, we used the Fisher/Correlation algorithm to select the most informative features from the obtained histograms. Next, the obtained features were projected into a 399-dimensional using PCA. Finally, the cosine similarity of each test pair is computed. For each of the four kinships, we performed 5-fold cross-validation and computed the mean accuracy of the five folds. To evaluate the performance of system, the Receiver Operating Characteristic (ROC) is used. We reported the mean verification accuracy of the three kinship subsets as the measure of the system performance.

4.3.2 Feature Extraction

For the proposed method, first, we tested the mean area under ROC with the restricted setting on F-S relationship in KinFace in the Wild-II database with varying the number of different parameters, and then the number who is performed the best performance for all relationships including F-D, M-D, and M-S.

- Appearance: The feature dimension for appearance-based was empirically set to be 300.

- Texture: for LBP, we divide each image into (6×6) blocks where the size of each block is (10×10) with neighborhood P = 8, a radius R = 1. With ML representation using level 6.

  For LPQ, we divide each image into (4×4) blocks where the size of each block is (10×10) with a window of size (3×3), a Gaussian derivative quadrature filter for local frequency estimation was used. With ML representation using level 7.

  For BSIF, we divide each image into (4×4) blocks where the size of each block is (16×16), a linear filter of size (8×8) was used. With ML representation using level 4.

- DCTNet: The network is composed of 2 convolution layers, to achieve the best performance, we tested different values of filters number, filters size, and block wise histogram size. The results are shown in Table 4.3.

  The parameters selected are six filters for the first layer and five filters for the second layer, and the perfect result gave when the filter size equal (8×8). The outputs of convolution layer give real values. The output of the last layer is turned into binary format. Then, we divided binary images into six non-overlapping blocks. The characteristics of these images are obtained by concatenating all the histograms of each block. Next, we apply (*TR Normalization*) to eliminate the disparity of histogram vector. Finally, we minimize the data using *PCA*.

We exploit a deep learning machine, where numerous layers of information processing stages are used. To extract the discriminant features for kinship verification, the input of the network is a gray-scale image.

After, 2D DCT filter bank is applied to the image to extract the most significant inherited features. As shown in Figure 4.2, we can see that some part of the face

Table 4.3: Average area under ROC of KinfaceW-II versus varying the parameters.

Number of Filters	The filter Size	The block-Wise Histogram Size	Accuracy Father-Sun
[2 2]	[2 2]	[2 2]	73.40
[2 2]	[4 4]	[2 2]	82.00
[2 2]	[8 8]	[2 2]	81.40
[2 6]	[5 5]	[6 6]	83.60
[2 6]	[5 5]	[8 8]	84.40
[4 4]	[4 4]	[2 2]	77.60
[5 6]	[5 5]	[6 6]	82.60
[5 5]	[5 5]	[12 12]	80.20
[6 6]	[5 5]	[6 6]	84.80
[6 6]	[5 5]	[8 8]	84.60
[6 6]	[5 5]	[10 10]	84.20
[6 5]	**[8 8]**	**[6 6]**	**88.80**
[6 5]	[8 8]	[14 14]	82.20
[6 6]	[5 5]	[5 5]	82.20
[7 7]	[5 5]	[6 6]	84.20
[8 8]	[5 5]	[6 6]	81.60
[8 2]	[5 5]	[8 8]	80.60
[10 10]	[5 5]	[5 5]	75.60
[14 14]	[5 5]	[8 8]	73.60

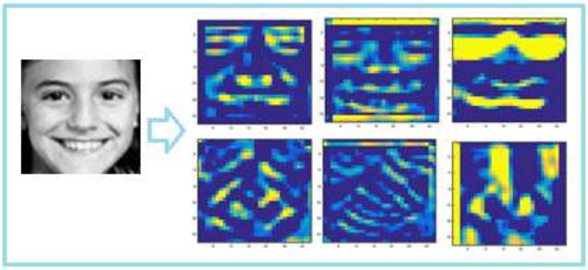

Figure 4.2: Example of the DCTNet filter bank on gray-scale image.

have a higher frequency like eye and nose, while other parts have a smaller frequency like frontal face. The results of this filters give only the most significant facial features. The light blue and yellow regions in the visualization highlight the inheritable genetic feature regions in the kinship images. One essential issue to address when adopting 2D DCT basis into the network as filter bank is the basis selection. The prior knowledge of human face characteristic is taken into account. Since human face distinct features are composed of more high-frequency horizontal components (eyes, eyebrows, and lips) than low-frequency vertical component, it is natural to rank the 2D DCT bases by horizontal-frequency major order. The DCTNet keeps the importance of horizontal frequency direction at each turn to extract inherited feature passed from parents to his children that are a result of genetic inheritance. On the other hand, block-wise histogramming, which is capable of implicitly encoding spatial information of image regions, is useful for the classification task. We propose an efficient method normalization (TR Normalization) to regulate the histogram of DCTNet for robustness. We adopt the ranking idea by quantifying the correlation between variables.

4.3.3 Results and analysis

This section demonstrates the performance of our proposed method on four kinship relations: father-son (F-S), father-daughter (F-D), mother-son (M-S), and mother-daughter (M-D) on KinFaceW-II database. In this experiment, we use different feature descriptors to determine to which extend the dominant facial features are crucial in verifying and establishing the relationship between individuals and how the features are represented to explain the relationship.

1. First, the results of our proposed appearance method are reported in Table 4.4. Considering the average accuracy and the accuracy of all the kinship relations,

Table 4.4: Kinship verification accuracy in % using appearance features on KinFaceW-II database.

Type	Features	FS	FD	MD	MS	Mean
Appearance based	PCA	74.40	70.00	73.00	70.60	72.00
	LDA	**77.00**	**75.60**	**75.80**	**74.20**	**75.70**
	LPP	70.90	63.30	65.40	65.00	67.00

Table 4.5: Kinship verification accuracy in % using texture features on KinFaceW-II database.

Type	Features	FS	FD	MD	MS	Mean
Texture based	MB-LBP	79.00	77.00	75.00	73.20	76.10
	MB-LPQ	81.20	74.00	74.00	73.20	75.00
	MB-BSIF	83.00	77.00	76.00	75.00	78.00
	ML-LBP	82.60	77.80	**76.80**	75.20	78.10
	ML-LPQ	80.20	75.00	75.40	74.60	76.30
	ML-BSIF	**83.20**	**79.00**	76.80	**75.60**	**79.20**

Table 4.6: Kinship verification accuracy in % on KinFaceW-II using DCTNet.

TR Normalization	FS	FD	MD	MS	Mean
Yes	**88.80**	**80.80**	**83.80**	**85.60**	**84.75**
No	84.60	78.80	78.80	78.60	80.20

LDA is the best performing method with accuracy equal to 75.7% , while LPP shows the worst performance. For global descriptors, we conducted that LDA is more informative in learning kin relationship.

2. Table 4.5 shows a comparison between representations (MLs and MBs), texture descriptors (LBP, LPQ, and BSIF). We observed from the table that the performance of MLs is better than MBs when using the same descriptor. This is due that to the fact MLs representation gives more detailed information of the image than the MBs. When comparing the different baseline methods, we can see that BSIF yields in the best overall performance.

3. We tested our proposed method DCTNet with TR normalization against the DCTNet without TR normalization. The results are reported in Table 4.6. The performance of DCTNet feature with normalization technique report the best performance on the four relationships. Overall, the TR Normalization enhanced the verification accuracy by a significant margin, F-S (improved by 4.2%), F-D (improved by 2.0%), M-D (improved by 5.0%), and M-S (improved by 7.0%).

4. On the other hand, DCTNet features report the best performance on all kinship relations significantly improving the verification accuracy. The gain in verification performance of the deep features varies between 1.8% and 10%, compared with the best texture accuracy. These results highlight the high ability of DCTNet in learning kinship verification.

5. Deep features against texture and appearance features: Local-based features perform better than global based. These results are due to the local features which have a highly discriminative, invariance to monotonic gray-level changes and computationally efficient, but the global features have no discriminating ability for constituent parts of the image. One can conclude that checking the kinship relation is easier when using facial parts rather than the whole

Table 4.7: Kinship verification accuracy in % on KinFaceW-II database of proposed multimodal system using texture features.

Type	Features	FS	FD	MD	MS	Mean
Color based	ML-LBP	86.4	79.0	82.0	83.8	82.80
	ML-LPQ	85.8	77.0	79.2	82.4	81.00
	ML-BSIF	**86.7**	**80.0**	**82.4**	**86.2**	**83.80**

Table 4.8: Mean accuracy in % of proposed multimodal system using DCTNet on KinFace-W.

Data base	FS	FD	MS	MD	Mean
KinFace in the Wild-I	82.33	75.15	72.36	68.60	74.61
KinFace in the Wild-II	91.40	88.20	87.40	90.00	89.25

face. The local descriptors achieved a good performance, but the features are designing manually, incomprehensible regarding of gender, age variation and need to search over space and scale. So no relation between blocks exists.

We have proposed a simple deep Learning (DCTNet) technique to address this problem. Deep learning recognition methods automatically learn high-level features from multilayer and parameters, rather than designing features manually and a spatial relation between blocks are encoded.

6. For the multimodal system, we extracted features based on color cue, we have used ML method and DCTNet which they achieve the best accuracy. The results are shown in Table 4.7 and Table 4.8. From Table 4.8, we observed that DCTNet descriptor achieves a higher accuracy equal to 89.25%. The performance is improved unlike the use on the gray-scale image directly. The proposed color-texture features outperform the gray-texture features methods in verification accuracy 4.4%, 5.5% and 4.6% and 4.5 % on LBP, LPQ, BSIF and DCTNet features respectively. DCTNet reported the best performance on all kinship

relations and the higher accuracy of human performance on the task of kinship verification is 91.4% on Father-Son relationship.

We can conclude that color does provide some discriminative information that can help in boosting the kinship verification performance. From a biological opinion, the chromaticity of the face is tied to genetically expressed characteristics, such as eye color or skin tone.

Our method showed good generalization capabilities and outperformed the recognition capabilities of human raters and previous approaches in the literature.

7. Comparison with Databases: In order to validate our proposed DCTNet, we conducted experimental based on KinFaceW-I, KinFaceW-II, and UBKinFace databases. The results for different relationships (F-S, F-D, M-S, and M-D) on KinFaceW databases is reported in Table 4.8 and UBKinFace is reported in Table 4.9. Our results show that the average accuracy of human performance on the task of kinship verification is 67.00 - 89.25 % on the three databases. The KinFaceW-II database is the best performing. These results are maybe due to the nature KinFaceW-II images (Image cropped from same photo) data sets. However, we showed that the images in these data sets have a high potential for biased results.

   The classification on UB Kin Face (young parents) reaches around 70%, and this is comparable to UB Kin Face (old parents) who reaches 67%. Moreover, the variation in age of the persons has an effect on the kinship verification accuracy because facial features continuous changes over time. These results are maybe due to the different age of the pairs. One can conclude that checking the kinship relation is easier between persons of close age.

8. Comparing relations. On the three databases, the best verification accuracy is obtained for F-S relationship. The ROC curves for separate relations are depicted in Figure 4.3, 4.4, 4.5 and 4.6. These results are maybe due to the different sex of the pairs. One can conclude that the labels such as sex are helpful to determine the relationship. Although this finding is not an absolute rule. It is remarkable that the performance of kinship verification between M-S is better than M-D in different approaches. These results motivate us to try to identify the characteristic that distinguishes the differences between sexes.

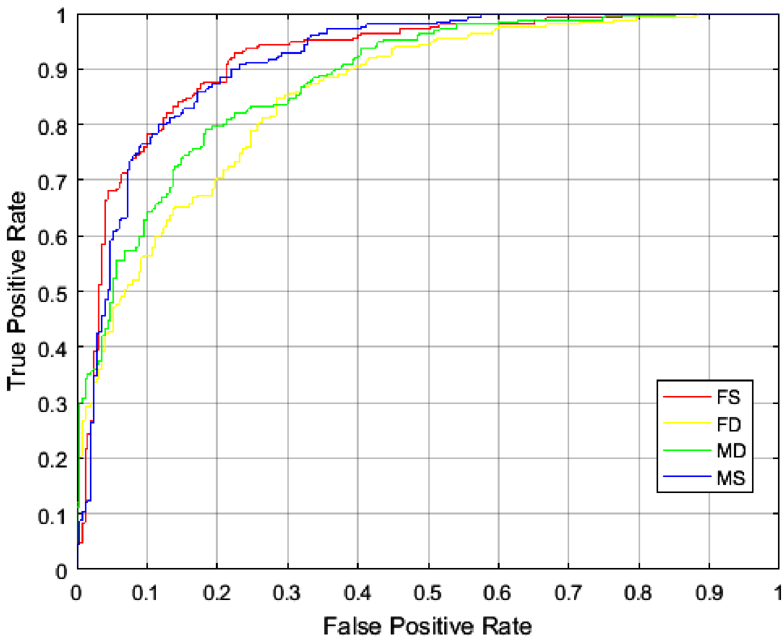

Figure 4.3: The ROC curves of BSIF method obtained on KinFaceW-II data sets.

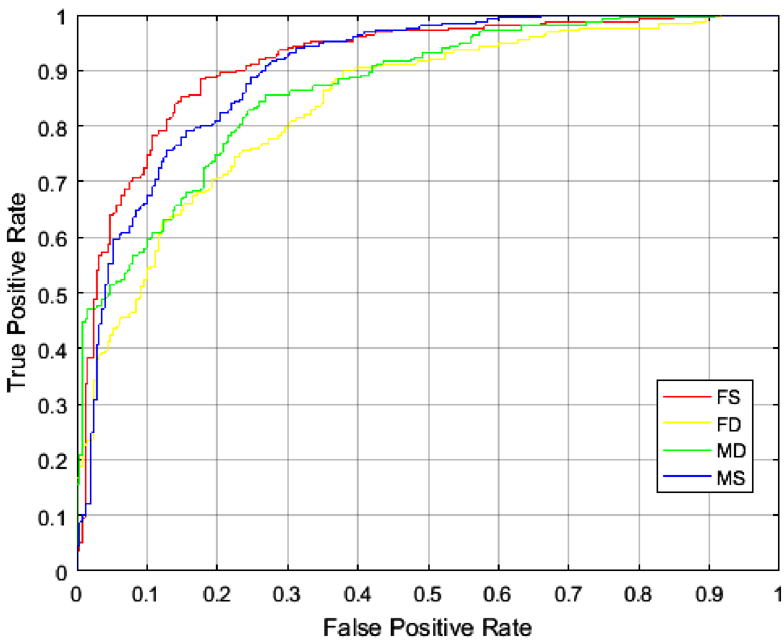

Figure 4.4: The ROC curves of LBP method obtained on KinFaceW-II data sets.

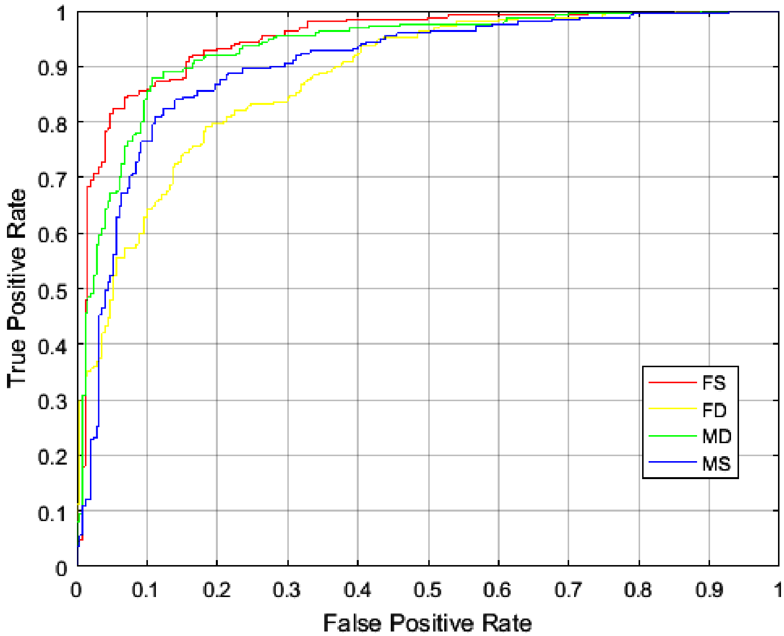

Figure 4.5: The ROC curves of DCTNet method obtained on KinFaceW-II data sets.

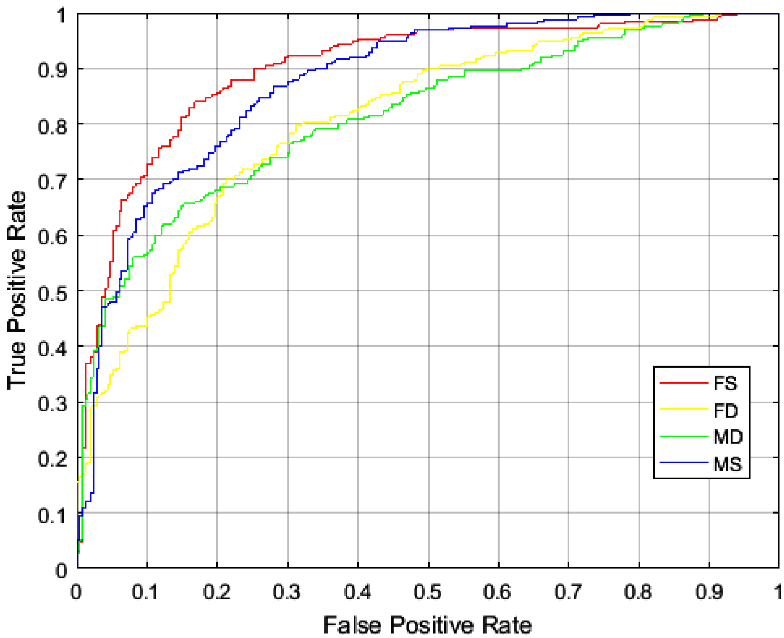

Figure 4.6: The ROC curves of LPQ method obtained on KinFaceW-II data sets.

Table 4.9: Mean kinship verification accuracy in % on *UB Kin Face* using DCTNet.

Database	Mean Accuracy
UB Kin Face (Young Parent)	70.00
UB Kin Face (Old Parent)	67.00

Interestingly, specific facial parts like eyes, nose, and mouth can be considered as strong similarities which are considered as kinship clues in daily life. But, this idea is sensitive because the family may have weak or strong similarities in specific facial parts. For instance, the eyes of the son may be very similar to his father's eyes but very different from his mother's eyes, thereby complicating the analysis.

Comparisons with other methods. We presented the comparison between our proposed *DCTNet* method and other state-of-the-art deep learning, feature learning under different metric learning methods. The comparison results are reported in Table 4.10. Table 4.11 shows a comparison of the multimodal results with the previous works. Firstly, we can see that deep learning method perform better than other traditional machine learning methods.

Unimodal system: Comparing our results against the state-of-the-art ones demonstrates considerable improvements in all the kinship subsets. The improvement in verification accuracy of our approach compared with the best performing method ranges from 3.4% to 7.2% respectively depending on the relationship type.

Multimodal system: We can remark that our advanced approach outperforms another state-of-the-art approach concerning the mean verification rate. Compared with the best performing methods, our method makes improvement in two relationships F-S and F-D because our proposed fused method provides some discriminative biological information that can help in boosting the kinship verification performance.

We can explain the reason why our DCTNet achieves better verification rate when compared to the state-of-the-art approaches. By using the 2D DCT as a filter bank, proposed filters enjoys a good approximation for high ranked eigenvectors, the prior knowledge of human face characteristic is taken into account, and the obtained data contain a higher-level information. By creating a spatial relation between blocks using

Table 4.10: Comparison of proposed DCTNet against state of the art on KinFaceW-II (unimodal system).

Method	Years	Features	FS	FD	MS	MD	Mean
NRML [42]	2014	LBP	79.20	71.60	72.20	68.40	72.85
		HOG	80.80	72.80	74.80	70.40	74.70
SM [53]	2015	SPLE	76.90	74.30	77.40	77.60	76.60
		LBP	69.20	68.30	68.90	69.3	68.90
LGM [53]	2015	LBP	69.20	68.30	68.90	69.30	68.90
		SPLE	74.90	71.00	76.90	76.40	74.80
SSML [74]	2016	LBP	82.40	78.60	79.80	77.93	79.68
		HOG	85.00	77.00	80.40	78.40	80.15
LMNN [74]	2016	LBP	77.80	73.20	70.60	70.40	73.00
		HOG	83.20	75.60	77.60	77.40	78.45
GSML [74]	2016	LBP	75.60	72.00	69.20	71.80	72.15
		HOG	83.38	75.20	75.80	76.40	77.70
SMCNN [78]	2016	CNN	75.00	79.00	85.00	78.00	79.20
LML [75]	2017	LE	76.80	74.20	76.60	73.80	75.40
		LBP	66.00	64.80	67.80	66.80	66.40
		TPLB	68.60	66.20	65.40	70.80	67.80
		SIFT	72.20	66.00	68.20	66.20	68.20
SML [75]	2017	LE	76.20	70.10	72.40	71.80	72.60
		LBP	66.90	65.50	63.10	68.30	66.00
		TPLBP	71.80	63.30	63.00	67.60	66.40
		SIFT	68.10	63.80	67.00	63.90	65.70
Our method	2018	DCTNet	**88.80**	**80.80**	**83.80**	**85.60**	**84.75**

Table 4.11: Comparison of proposed DCTNet against state of the art on KinFaceW-II (Multimodal system).

Method	Reference	FS	FD	MS	MD	Mean
Polito	[73]	85.3	85.8	87.5	83.7	86.30
$MNRML$	[42]	76.9	74.3	77.4	77.6	76.60
LM^3L	[75]	82.4	74.2	76.6	78.7	78.70
$DMML$	[50]	78.5	76.5	78.5	79.5	78.20
IML	[50]	79.4	71.5	76.3	77.3	76.10
$DDMML$	[60]	87.4	83.8	83.2	83.0	84.30
CNN-Basic	[79]	84.9	79.6	88.3	88.5	85.30
CNN-Points	[79]	89.4	81.9	**89.9**	**92.4**	88.40
Our	-	**91.4**	**88.2**	87.4	90.0	**89.25**

block-wise histogramming. By applying Fisher/Correlation functions, the method can discard the duplicate features and provide a useful classification with the smallest error.

Compared with Convolutional Neural Networks (*CNN*): Discrete Cosine Transform Network (*DCTNet*) is a lightweight design compared with a convolutional neural network (*CNN*).

In *CNN*, Different filters were applied to get a varied information and no prior knowledge of human face characteristic is taken into account. We need thousands of facial images to train a deep convolutional network. In kinship verification problem, it is very hard to build a basic model and to estimate the CNN parameters because the kinship databases are not big enough. Instead, a simple *DCTNet* was proposed to learn *2D DCT* filters bank by binarization and a spatial relation between blocks are encoded by block-wise histogramming where the prior knowledge of human face characteristic is taken into account. The *DCTNet* network was built with an independent set of data.

4.4 Experimental 3

Local binary descriptor constitutes power visual cues for feature representation. They provide discriminative information about small appearance details in local neighborhoods. So, they are robust to local changes databases such as illumination, identity, and expression. Unlike existing local descriptors is not discriminatory enough to estimate the relationship between two people. This is mainly due to the learning feature code individually and the hand-crafted features which previous knowledge is required. In this experimental, we proposed an effective Context-Aware Local Binary Feature Learning (CA-LBFL) for kinship verification in order to solve the proposed problem. CA-LBFL a method has applied to learn contextual features from raw pixels directly and to eliminates the dependence on hand-crafted features. Experimental results demonstrate that the proposed method achieves competitive results compared with other states-of-the-art.

To evaluate the effectiveness of proposed CA-LBFL methods, we performed kinship verification experiments on the challenging KinFace W-I, KinFace W-II and UB KinFace V2. A brief presentation of the used databases is provided bellow.

4.4.1 Experimental Evaluation

For the experiments, same number of negative pairs (no kin relation) as positive pairs (with a kin relation) is created by associating faces from persons which do not have a kinship relation. We apply the features selection technique to select the most relevant features. For each of the four kinship subsets (i.e. F-D, F-S, M-D and M-S), we recommend taking the feature normalization technique to improve recognition rate. The result of normalization is that the features will be rescaled so that they all have properties of a standard normalization distribution. We perform 5-fold cross validation and compute the mean accuracy of the five folds. Finally, the cosine similarity of each test pair is computed. We report the mean verification accuracy of the four kinship subsets as the measure of system performance.

Table 4.12: Mean kinship verification accuracy in % on *KinFaceW-II* versus varying λ_1, λ_2, and λ_3.

Parameters			Accuracy
λ_1	λ_2	λ_3	F-S
10^2	10^2	10^5	79.0
10^2	10^3	10^6	81.5
10^3	10^4	10^7	84.2
10^3	10^3	10^8	86.8
10^3	10^2	10^8	**88.8**
10^4	10^2	10^9	83.0
10^4	10^3	10^9	82.3
10^2	10^3	10^8	81.0

Table 4.13: Mean kinship verification accuracy in % on *KinFaceW-II* versus varying local region.

Local	Region	Accuracy
4	4	80.0
6	6	83.0
8	**8**	**88.8**
10	10	82.2

This investigation are published in conference Papers C (2) 1.

4.4.2 Parameter Adaptation

For the proposed CA-LBFL method, first we tested the mean area under ROC on KinFace in the Wild-II database with different parameters, and then applied these parameters for the rest of experiments. We fixed dictionary size as 500, the binary code length K as 30 and set R as 3. To achieve the best performance, we tested different values of λ_1, λ_2 and λ_3. The results are shown in Table 4.12.

The three parameters λ_1, λ_2 and λ_3 were selected as 10^3, 10^2 and 10^8.

We also tested different values of local region, finally, was fixed to 8×8, each face image was represented as a 38400-dimensional feature vector after using CA-LBFL ($38400 = 600 \times 8 \times 8$). The results are shown in Table 4.13.

4.4.3 Results and Analysis

To achieve the best performance, we tested different size of features to be selected on KinFaceW-II (F-S relationship) and then applied for the rest of relationships. The results are shown in Table 4.14. The final vector has size 662.

Table 4.14: Kinship verification accuracy in % of proposed method on KinFaceW-II databases (F-S) versus varying features number.

num Features	FS relationships
2592	81.80
2016	84.00
1440	85.80
1152	85.20
720	87.40
662	**88.80**
605	88.60
576	88.20
432	87.40
259	88.00
202	86.60
115	84.40

Table 4.15: Kinship verification accuracy in % of proposed method on KinFace W databases.

Data Base	FS	FD	MS	MD	Mean
KinFaceW-II	88.80	79.00	83.40	80.80	83.00
KinFaceW-I	72.84	69.17	65.57	68.63	69.05

Table 4.16: Kinship verification accuracy in % on *UB Kin Face*.

Data Base	Mean Accuracy
UB Kin Face Young Parent	**68.00**
UB Kin Face Old Parent	61.00

The results for different relationships (F-S, F-D, M-S and M-D) on different databases are reported in Table 4.15 and Table 4.16. Our results show that the average accuracy of human performance on the task of kinship verification is 61.00 - 83.00 % on the three data bases.

The best verification accuracy is obtained for F-S with 88.80% on The KinFace W-II database while the lowest is F-D. The classification on UB Kin Face (young parents) reaches around 68.0%, this is comparable to UB Kin Face (old parents) who's reaches 61.0%.

Also, we designed our evaluation with two different settings: image-unrestricted (for each face image, we first apply PCA to project feature into a low-dimensional feature vector and then side-information based linear discriminant analysis (SILD) is employed to learn a distance metric) and image-restricted (for each face image, PCA

Table 4.17: Kinship verification accuracy in % of proposed method on KinFaceW-II databases.

Setting	FS	FD	MS	MD	Mean
image-unrestricted	79.40	75.00	72.00	72.60	74.75
image-restricted	81.40	75.00	78.60	73.60	77.15

Table 4.18: Kinship verification accuracy in % of proposed method on KinFaceW-II databases under restricted setting.

Features selection	FS	FD	MS	MD	Mean
No	80.00	70.80	74.60	73.50	74.72
Yes	81.40	75.00	78.60	73.60	77.15

is first used to project feature into a low dimensional feature vector and then neighborhood repulsed metric learning (NRML) [42] is employed to learn a discriminative distance metric). Table 4.17 list the mean accuracy of the two different settings. From the experimental results, we can see that the evaluation under image-restricted report 77.15% better than image-unrestricted with 74.75%, while our proposed setting report the best performance with 83.00%.

Under restricted setting, we tested our proposed method CA-LBFL with features selection against the CA-LBFL without features selection method. The results reported in Table 4.18 . The performance of CA-LBFL feature with features selection technique report the best performance on the four relationships. Overall, the features selection enhanced the verification accuracy by a significant margin, F-S (improved by 1.4%), F-D (improved by 4.2%), M-S (improved by 4.0%), and M-D (improved by 0.1%).

Comparison with Baseline Methods: A comparison against other reported methods on Kin Face W-II is reported in Table 4.19. We compare the proposed method with various features under different strategies including Neighborhood Repulsed Metric Learning (NRML) [115], Large Margin Nearest Neighbor (LMNN) [51], Generalized Sparse Metric Learning (GSML) [74], single metric learning (SML) [75], Individual Metric Learning (IML) [50], Local Large-Margin Multi-Metric Learning (LM^3L)[75] and Sparse Similarity Metric Learning (SSML) [74].

These results clearly show that our proposed approach outperform the state-of-the-art methods in all configurations (i.e. for different protocols and parameters). The F-S, M-D, M-D and F-D relationship, our method outperforms all the others and provides state-of-the-art performance. The success of the proposed method is mainly due to the robust information. It constrains the number of shifts from different binary

Table 4.19: Comparison of our approach for kinship verification against state of the art on *KinFaceW-II* database .

Method	Feature	FS	FD	MS	MD	Mean
NRML [42]	LBP	79.20	71.60	72.20	68.40	72.85
	HOG	80.80	72.80	74.80	70.40	74.70
LMNN[74]	LBP	77.80	73.20	70.60	70.40	73.00
	HOG	83.20	75.60	77.60	77.40	78.45
GSML[74]	LBP	75.60	72.00	69.20	71.80	72.15
	HOG	83.38	75.20	75.80	76.40	77.70
SML[75]	LE	76.20	70.10	72.40	71.80	72.60
	LBP	66.90	65.50	63.10	68.30	66.00
	TPLBP	71.80	63.30	63.00	67.60	66.40
	SIFT	68.10	63.80	67.00	63.90	65.70
IML [50]	ALL	79.4	71.50	76.3	77.30	76.10
LML[74]	LE	76.80	74.20	76.60	73.80	75.40
	LBP	66.00	64.80	67.80	66.80	66.40
	TPLB	68.60	66.20	65.40	70.80	67.80
	SIFT	72.20	66.00	68.20	66.20	68.20
LM^3L[75]	ALL	82.40	74.20	76.60	78.70	78.70
SSML[74]	LBP	82.40	78.60	79.80	77.93	79.68
Our method	CA-LBFL	**88.80**	**79.00**	**83.40**	**80.80**	**83.00**

bits to exploits the contextual information of adjacent bits and obtains more robust local binary features.

Finally, Figure 4.7 and Figure 4.8 provide the performance of our best approaches.

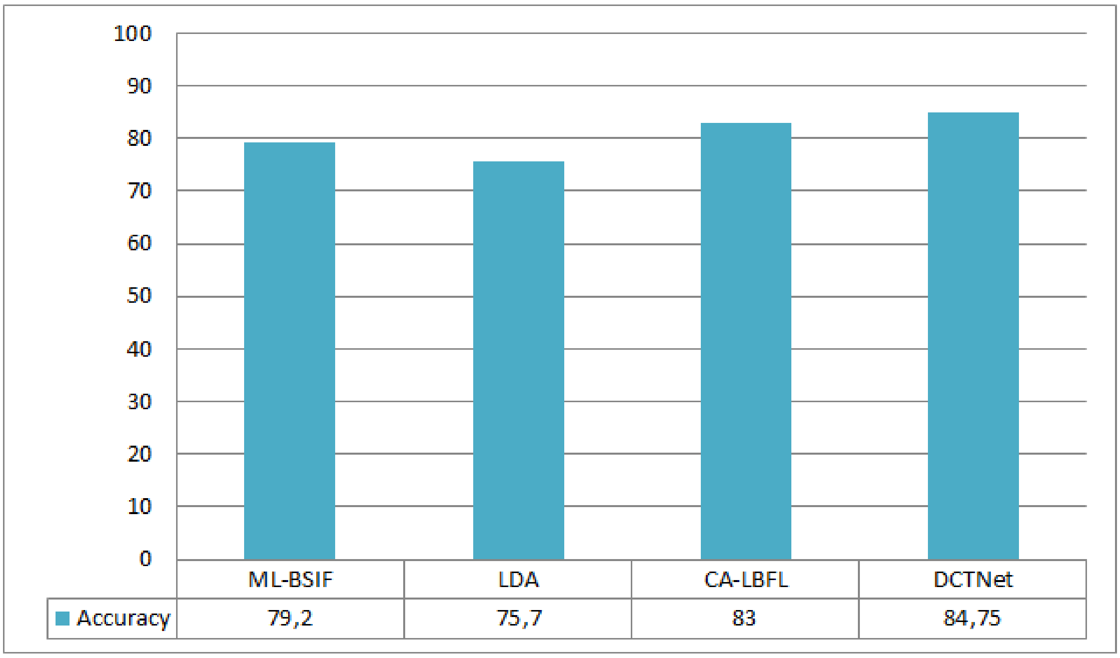

Figure 4.7: The performance of our best approaches on unimodal system.

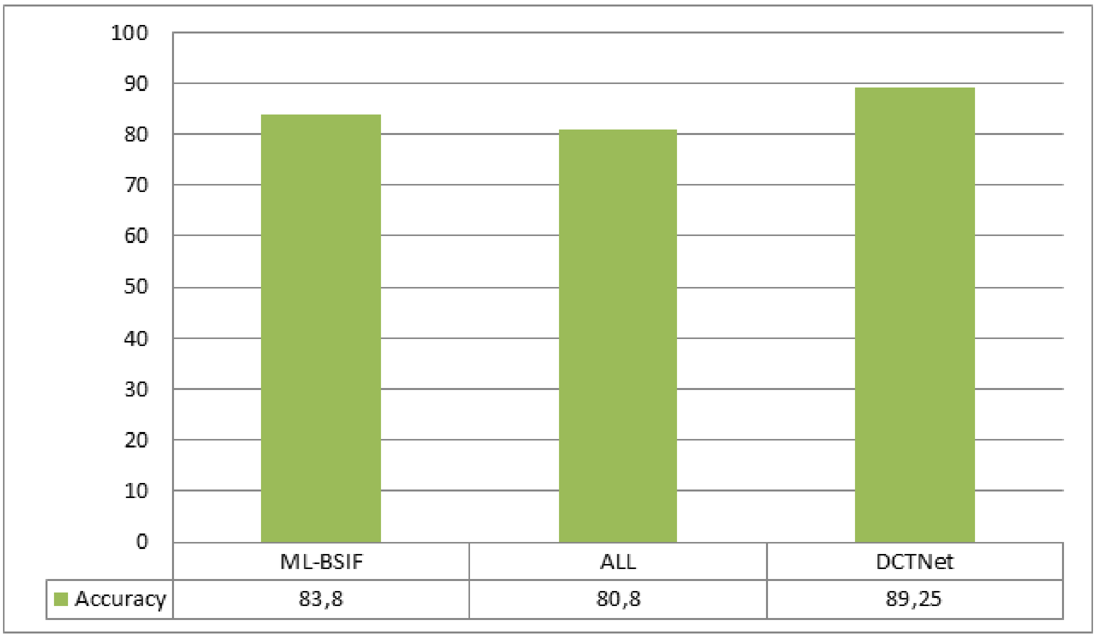

Figure 4.8: The performance of our best approaches on multimodal system.

4.5 Conclusion

In this chapter, we have investigated the kinship verification problem from faces using three experiments.

In experimental 1, we have investigated several feature extraction methods and similarity measures. We performed extensive evaluation on four publicly available databases (KinFaceW-I & II, UBKinFace and Cornell Kin Faces). Faces are described using LBP, TPLBP, FPLBP, HOG and Gabor features. Support vector machines and Nearest neighbor (City block distance, Euclidean distance, Chi-Square distance and Cosine Distance) are used for classification. Our study demonstrates that the performance of kinship verification affected by several labels such as age and database environments. Furthermore, The best features are obtained by the Local Binary Pattern and The best classifier is the cosine similarity.

In experimental 2, we tackled the problem of automatic kinship verification from facial images considering four relationships: F-S, F-D, M-S and M-D. We proposed a novel solution using a Discrete Cosine Transform Network (DCTNet) via 2D-DCT filters bank. To improve the kinship verification performance discriminative biological information are used. Our proposed method demonstrates a high efficiency of deep features in describing faces for inferring kinship relations on the used KinFaceW-I, KinFaceW-II and UBKinFace. Furthermore, comparison of our approach to the previous work indicates significant improvements in verification accuracy.

In experimental 3, we proposed an effective Context-Aware Local Binary Feature Learning (CA-LBFL) for facial kinship verification. The CA-LBFL is a method has applied to learn contextual features from raw pixels directly and to eliminate the dependence on hand-crafted features. Our study demonstrates the high efficiency of Context-Aware Local Binary Feature Learning in describing faces for kinship verification and shown its effectiveness compared with several state-of-the-art methods.

Chapter 5

Summary and future work

The purpose of the work presented in this thesis was the kinship verification from face images. Several features extractions approaches are proposed in the literature, Hand-Crafted approaches and Deep learning approaches. On the other hand, various metric learning methods have been investigated for tackling this problem.

This thesis focus on kinship verification based feature extraction task. We presented an original investigation to detect the most significant characteristic features of human faces and it should help from the improved performance of kinship verification.

In the first part of this work, we focused on the approaches based Hand-Crafted features (Global and Local features), which is very used in face recognition. These features are based on the learning of the surface properties and appearance of an object given by the shape, size, arrangement, density, a proportion of its elementary parts.

The Global appearance-based methods try to find a suitable representation of the whole image, all pixels are regarded, by approximating the original data and keeping as much information as possible. On the other hand, Local texture descriptors represent certain region properties by multi-dimensional histograms. Very often geometric properties (e.g., location, distance) of interest points in the region (corners, edges) and local orientation information (gradients) are used.

These methods have been validated on different databases, and have been implemented under Matlab. Also, we explored the feature combination to perform multiple feature fusion to extract complementary information to improve the kinship verification performance.

The choice of these attributes to characterize a face is a critical problem requiring increased experience in the field of kinship recognition. The existing local and global descriptors are designing manually, which previous knowledge is required.

We have proposed an efficient system with the aim of providing enhancement to the accuracy of kinship verification. The system based on a simple deep learning method called Discrete Cosine Transform Network (DCTNet), where 2D-DCT is adopting as a filter bank to extract the most significant inherited facial features. Deep learning recognition methods automatically learn high-level features from multilayer and parameters, rather than designed features manually and a spatial relation between blocks are encoded. To the best of our knowledge, the DCTNet is being used for the first time in our work for the kinship verification. Moreover, we introduced the multimodal system based on some discriminative biological information. Our results highlight the high ability of DCTNet in learning kinship verification.

Finally, we proposed an effective Context-Aware Local Binary Feature Learning (CA-LBFL) for kinship verification. The CA-LBFL is a method has applied to learn contextual features from raw pixels directly and to eliminate the dependence on hand-crafted features. Experimental results demonstrate that the proposed method achieves competitive results compared with other states-of-the-art.

In the future, more work should be carried out for developing of advanced methods for kinship verification from videos and possibly also 3D face images. Also, development of novel methods for Kinship verification from gait, voice (and possibly also recording new databases if needed) and development of multimodal methods for Kinship verification from gait and voice.

Appendix A

Side-Information Linear Discriminant Analysis (SILD)

Side-Information based Linear Discriminant Analysis (SILD) [83] is proposed in which the within-class and between-class scatter matrices are directly calculated by using the side-information of pairs of data points.

In kinship verification, SILD is employed to learn a distance metric learning by estimate the within-class scatter matrix C_p by employing positive pairs and the between-class scatter matrix C_n by using negative pairs in training set:

$$C_p = \sum l_{ij} = 1(x_i - x_j)(x_i - x_j)^T$$

$$C_n = \sum l_{ij} = -1(x_i - x_j)(x_i - x_j)^T$$

Then, SILD learns a discriminative linear projection $W \in \Re d \times m, m \leq d$ by solving the optimization problem:

$$max_w \frac{det(W^T C_n W)}{det(W^T C_p W)}$$

By diagonalizing C_p and C_n as:

$$C_p = U D_p U^T, (U D_p^{-1/2}) C_p (U D_p^{-1/2}) = I$$

$$(U D_p^{-1/2}) C_n (U D_p^{-1/2}) = V D_n V^T$$

the projection matrix W can be computed as:

$$W = U D_p^{-1/2} V,$$

in which U and V are an orthogonal matrices, and D_p and D_n are a diagonal matrices. In the transformed subspace, the squared Euclidean distance of a pair of data points x_i and x_j is calculated by:

$$d_w^2(x_i, x_j) = \|W^T x_i - W^T x_j\|_2^2 = (x_i - x_j)^T W W^T (x_i - x_j) = (x_i - x_j)^T M (x_i - x_j)$$

This distance is equivalent to computing the squared Mahalanobis distance in the original space, and we have $M = WW^T$.

Appendix B

Cross Validation

Cross-validation, sometimes called rotation estimation or out-of-sample testing is any of various similar model validation techniques for assessing how the results of a statistical analysis will generalize to an independent data set. It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice [1] [116].

It is commonly used in applied machine learning to compare and select a model for a given predictive modeling problem because it is easy to understand, easy to implement, and results in skill estimates that generally have a lower bias than other methods.

B.1 K-fold Cross Validation

In k-fold cross-validation, the data is first partitioned into k equally (or nearly equally) sized segments or folds. Subsequently k iterations of training and validation are performed such that within each iteration a different fold of the data is held-out for validation while the remaining k-1 folds are used for learning. Data is commonly stratified prior to being split into k folds. Stratification is the process of rearranging the data as to ensure each fold is a good representative of the whole [2].

In other words, the K-fold Cross-validation is a computer intensive technique, using all available examples as training and test examples [117]. Figure B.1 showed an example of 4-cross validation.

[1] https://en.wikipedia.org/wiki/Cross-validation(statistics)
[2] http://leitang.net/papers/ency-cross-validation.pdf

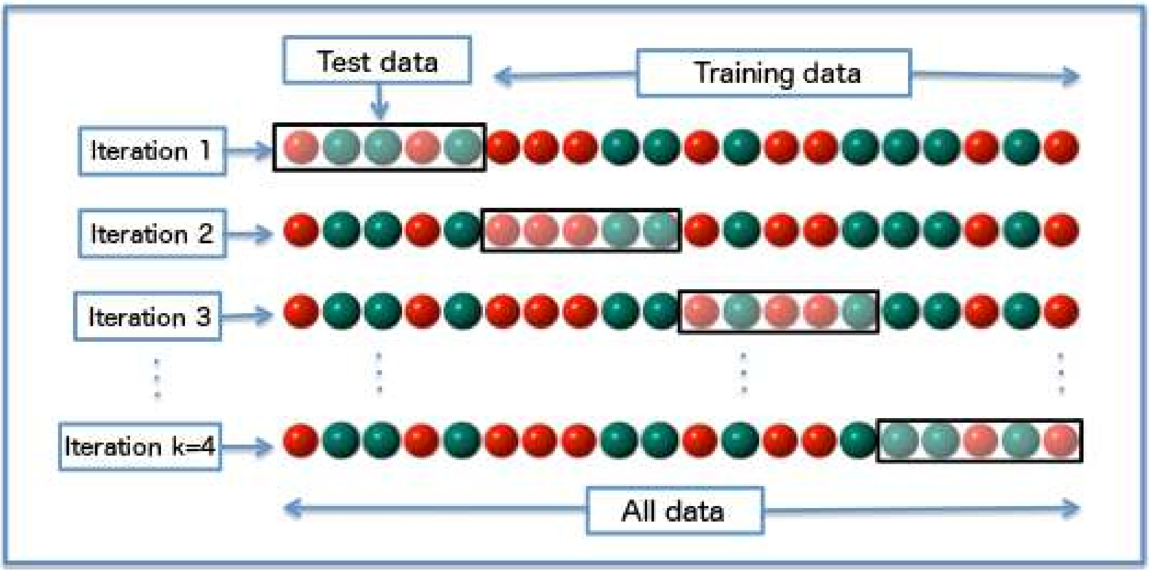

Figure B.1: Example of 4-cross validation.

Mathematically, the k-fold cross-validation operator is defined as following:

- Divide the data into K roughly equal parts.
- for each k=1,2,...K, fit the model with parameter λ to the other K-1 parts, giving $\beta^{-k}(\lambda)$ and compute its error in predicting the kth part:

$$E_k(\lambda) = \sum_{i \in_{kth} part} (y_i - x_i \hat{\beta}^{-k}(\lambda))^2 \tag{B.1}$$

- This gives cross validation error

$$CV(\lambda) = \frac{1}{K} \sum_{k=1}^{K} E_k(\lambda) \tag{B.2}$$

- Do this for many values of λ and chose the value of λ that makes $CV(\lambda)$ smallest.

Appendix C

List of publications

C.1 Journal Publication

- A. Tidjani, A. Taleb-Ahmed, D. Samai and K. Aiadi, "Deep learning features for robust facial kinship verification". IET Image Processing, 12(12), 2018, 2336-2345.

- A. Tidjani, E. Boutellaa, D. Samai, A. hadid, A. Taleb-Ahmed and K. Ben Sid, "Investigating Feature Extraction and Matching Approaches for Kinship Verification From Facial Images". The special issue of EL MIR'AT SCIENCES magazine ISSN 2170-1555 (print-only).

- K. Bensid, FZ. Laallam, D. Samai and **A. Tidjani**, "Finger Knuckle Print Features Extraction using Simple Deep Learning Method," International Journal of Computer Science, Communication & Information Technology(CSCIT), vol. 5, pp. 12-18, (2017).

C.2 Conferences Publications

- A. Tidjani, A. Taleb-Ahmed, D. Samai and K. Aiadi. "Kinship Verification using Context-Aware Local Binary Feature Learning." IEEE International Conference on Control, Automation and Diagnosis, ICCAD'2018, MRRAKESH, MROCCO. pp.1-5, 2018.

- A. Tidjani, D. Samai, A. Hadid, A. Taleb-Ahmed and K. Bensid "En effective kinship verefication system using Multi-levels features extraction," Confrence Internationale en Automatique & Traitement de Signal (ATS-2017), pp. 1-6, 2017.

- A. Tidjani, D. Samai, A. Hadid, A. Taleb-Ahmed and K. Bensid "Investegating feature extraction and matching approaches for kinship verefication from facial images," Conference on Computing Systems and applications (CSA-2016), pp. 1-6, 2016.
- K. Bensid, D. Samai, FZ. Laallam, **A. Tidjani** and M. Korichi "Multimodal Palmprint Biometric System Using New Variants Of Local Phase Quantization and Support Vector Machine," the 2nd International Conference on Pattern Analysis and Intelligent Systems (PAIS 2016), pp. 1-6, 2016.
- K. Bensid, D. Samai, FZ. Laallam, **A. Tidjani**, "Efficient Person Identification by Finger-Knuckle-Print Based on Discrete Cosine Transform Network and SVM classifier," Proceedings of Engineering and Technology PET (ATS 2017), Vol.22, pp.78-83.

Bibliography

[1] M. Georgopoulos, Y. Panagakis, and M. Pantic, "Modelling of facial aging and kinship: A survey," *Image and Vision Computing*, 2018.

[2] T. R. Alley, *Social and applied aspects of perceiving faces.* Psychology Press, 2013.

[3] J. H. Park, M. Schaller, and M. Van Vugt, "Psychology of human kin recognition: Heuristic cues, erroneous inferences, and their implications." *Review of General Psychology*, vol. 12, no. 3, p. 215, 2008.

[4] I. STRAVINSKY, "Rodney needham remarks on the analysis of kinship and marriage," *Rethinking kinship and marriage*, vol. 11, p. 1, 2004.

[5] D. Jones, "The generative psychology of kinship: Part 1. cognitive universals and evolutionary psychology," *Evolution and Human Behavior*, vol. 24, no. 5, pp. 303–319, 2003.

[6] L. T. Maloney and M. F. Dal Martello, "Kin recognition and the perceived facial similarity of children," *Journal of Vision*, vol. 6, no. 10, pp. 4–4, 2006.

[7] M. F. Dal Martello and L. T. Maloney, "Where are kin recognition signals in the human face?" *Journal of Vision*, vol. 6, no. 12, pp. 2–2, 2006.

[8] L. M. DeBruine, F. G. Smith, B. C. Jones, S. C. Roberts, M. Petrie, and T. D. Spector, "Kin recognition signals in adult faces," *Vision research*, vol. 49, no. 1, pp. 38–43, 2009.

[9] K. Finkler, C. Skrzynia, and J. P. Evans, "The new genetics and its consequences for family, kinship, medicine and medical genetics," *Social Science & Medicine*, vol. 57, no. 3, pp. 403–412, 2003.

[10] J. H. Relethford, "Bioassay of kinship from continuous traits," *Human biology*, pp. 689–700, 1980.

[11] A. Alvergne, C. Faurie, and M. Raymond, "Differential facial resemblance of young children to their parents: who do children look like more?" *Evolution and Human behavior*, vol. 28, no. 2, pp. 135–144, 2007.

[12] L. M. DeBruine, "Resemblance to self increases the appeal of child faces to both men and women," *Evolution and Human Behavior*, vol. 25, no. 3, pp. 142–154, 2004.

[13] H. Dibeklioglu, A. Ali Salah, and T. Gevers, "Like father, like son: Facial expression dynamics for kinship verification," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1497–1504, 2013.

[14] G. Guo and X. Wang, "Kinship measurement on salient facial features," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 8, pp. 2322–2325, 2012.

[15] G. Somanath and C. Kambhamettu, "Can faces verify blood-relations?" in *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on.* IEEE, pp. 105–112, 2012.

[16] P. Alirezazadeh, A. Fathi, and F. Abdali-Mohammadi, "A genetic algorithm-based feature selection for kinship verification," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2459–2463, 2015.

[17] S. Wang, Z. Ding, and Y. Fu, "Cross-generation kinship verification with sparse discriminative metric," *IEEE transactions on pattern analysis and machine intelligence*, 2018.

[18] N. Kohli, M. Vatsa, R. Singh, A. Noore, and A. Majumdar, "Hierarchical representation learning for kinship verification," *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 289–302, 2017.

[19] M. B. López, E. Boutellaa, and A. Hadid, "Comments on the kinship face in the wild data sets," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 11, pp. 2342–2344, 2016.

[20] R. Fang, K. D. Tang, N. Snavely, and T. Chen, "Towards computational models of kinship verification," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, pp. 1577–1580, 2010.

[21] M. Shao, S. Xia, and Y. Fu, "Genealogical face recognition based on ub kinface database," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*. IEEE, pp. 60–65, 2011.

[22] Y. Guo, H. Dibeklioglu, and L. van der Maaten, "Graph-based kinship recognition," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*. IEEE, pp. 4287–4292, 2014.

[23] P. M. Nugent, *KINSHIP," in PsychologyDictionary.org*, May 11, 2013, https://psychologydictionary.org/kinship/(accessedJuly22,2018).

[24] L. S. David, *International Encyclopedia of the Social Sciences*. Macmillan, 1968.

[25] M. Guha, *International Encyclopedia of the Social Sciences (2nd edition)*. Emerald Group Publishing Limited, 2008.

[26] S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: Security and privacy concerns," *IEEE security & privacy*, no. 2, pp. 33–42, 2003.

[27] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, *Handbook of fingerprint recognition*. Springer Science & Business Media, 2009.

[28] A. K. Jain, S. S. Arora, K. Cao, L. Best-Rowden, and A. Bhatnagar, "Fingerprint recognition of young children," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1501–1514, 2017.

[29] A. Kong, D. Zhang, and M. Kamel, "A survey of palmprint recognition," *pattern recognition*, vol. 42, no. 7, pp. 1408–1418, 2009.

[30] D. Samai, K. Bensid, A. Meraoumia, A. Taleb-Ahmed, and M. Bedda, "2d and 3d palmprint recognition using deep learning method," in *2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS)*. IEEE, pp. 1–6, 2018.

[31] A. M. Schuster, J. A. Clark, G. T. Davis, P. A. Ivanov, and R. A. Zurek, "Method and apparatus including parallell processes for voice recognition," Jan. 10 2017, uS Patent 9,542,947.

[32] C. Papagno, G. Mattavelli, A. Casarotti, L. Bello, and G. Gainotti, "Defective recognition and naming of famous people from voice in patients with unilateral temporal lobe tumours," *Neuropsychologia*, vol. 116, pp. 194–204, 2018.

[33] R. P. Wildes, "Iris recognition: an emerging biometric technology," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1348–1363, 1997.

[34] V. d. S. Klehm, W. C. Melo, F. S. Farias, K. V. Santos, and S. Waldir, "Iris recognition using minimum average correlation energy and principal component," in *Consumer Electronics (GCCE), 2015 IEEE 4th Global Conference on.* IEEE, pp. 213–214, 2015.

[35] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, no. 12, pp. 2234–2240, 2007.

[36] J. Galbally, S. Marcel, J. Fierrez *et al.*, "Biometric antispoofing methods: A survey in face recognition." *IEEE Access*, vol. 2, no. 1530-1552, p. 1, 2014.

[37] I. Nnorom, J. Igwe, and C. Oji-Nnorom, "Trace metal contents of facial (make-up) cosmetics commonly used in nigeria," *African Journal of Biotechnology*, vol. 4, no. 10, 2005.

[38] J. R. Maple, K. W. Vig, F. M. Beck, P. E. Larsen, and S. Shanker, "A comparison of providers and consumers perceptions of facial-profile attractiveness," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 128, no. 6, pp. 690–696, 2005.

[39] M. Almuashi, S. Z. M. Hashim, D. Mohamad, M. H. Alkawaz, and A. Ali, "Automated kinship verification and identification through human facial images: a survey," *Multimedia Tools and Applications*, vol. 76, no. 1, pp. 265–307, 2017.

[40] R. Fang, A. C. Gallagher, T. Chen, and A. Loui, "Kinship classification by modeling facial feature heredity," in *Image Processing (ICIP), 2013 20th IEEE International Conference on.* IEEE, pp. 2983–2987, 2013.

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

[42] J. Lu, X. Zhou, Y.-P. Tan, Y. Shang, and J. Zhou, "Neighborhood repulsed metric learning for kinship verification," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 2, pp. 331–345, 2014.

[43] 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.

[44] N. Kohli, R. Singh, and M. Vatsa, "Self-similarity representation of weber faces for kinship classification," in *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on.* IEEE, pp. 245–250, 2012.

[45] H. Dibeklioğlu, A. A. Salah, and T. Gevers, "Are you really smiling at me? spontaneous versus posed enjoyment smiles," in *European Conference on Computer Vision.* Springer, pp. 525–538, 2012.

[46] H. Yan and J. Hu, "Video-based kinship verification using distance metric learning," *Pattern Recognition*, vol. 75, pp. 15–24, 2018.

[47] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

[48] S. Xia, M. Shao, and Y. Fu, "Toward kinship verification using visual attributes," in *Pattern Recognition (ICPR), 2012 21st International Conference on.* IEEE, pp. 549–552, 2012.

[49] T. F. Vieira, A. Bottino, and I. U. Islam, "Automatic verification of parent-child pairs from face images," in *Iberoamerican Congress on Pattern Recognition.* Springer, pp. 326–333, 2013.

[50] H. Yan, J. Lu, W. Deng, and X. Zhou, "Discriminative multimetric learning for kinship verification," *IEEE Transactions on Information forensics and security*, vol. 9, no. 7, pp. 1169–1178, 2014.

[51] J. Hu, J. Lu, J. Yuan, and Y.-P. Tan, "Large margin multi-metric learning for face and kinship verification in the wild," in *Asian Conference on Computer Vision.* Springer, pp. 252–267, 2014.

[52] X. Wang and C. Kambhamettu, "Leveraging appearance and geometry for kinship verification," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, pp. 5017–5021, 2014.

[53] L. Kou, X. Zhou, M. Xu, and Y. Shang, "Learning a genetic measure for kinship verification using facial images," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[54] A. Bottinok, I. U. Islam, and T. F. Vieira, "A multi-perspective holistic approach to kinship verification in the wild," in *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, vol. 2. IEEE, pp. 1–6, 2015.

[55] H. Yan, J. Lu, and X. Zhou, "Prototype-based discriminative feature learning for kinship verification," *IEEE Transactions on cybernetics*, vol. 45, no. 11, pp. 2535–2545, 2015.

[56] X. Zhou, Y. Shang, H. Yan, and G. Guo, "Ensemble similarity learning for kinship verification from facial images in the wild," *Information Fusion*, vol. 32, pp. 40–48, 2016.

[57] X. Zhou, H. Yan, and Y. Shang, "Kinship verification from facial images by scalable similarity fusion," *Neurocomputing*, vol. 197, pp. 136–142, 2016.

[58] M. Xu and Y. Shang, "Kinship verification using facial images by robust similarity learning," *Mathematical Problems in Engineering*, vol. 2016, 2016.

[59] A. Puthenputhussery, Q. Liu, and C. Liu, "Sift flow based genetic fisher vector feature for kinship verification," in *Image Processing (ICIP), 2016 IEEE International Conference on*. IEEE, pp. 2921–2925, 2016.

[60] J. Lu, J. Hu, and Y.-P. Tan, "Discriminative deep metric learning for face and kinship verification," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4269–4282, 2017.

[61] E. Boutellaa, M. B. López, S. Ait-Aoudia, X. Feng, and A. Hadid, "Kinship verification from videos using spatio-temporal texture features and deep learning," *arXiv preprint arXiv:1708.04069*, 2017.

[62] H. Yan, "Kinship verification using neighborhood repulsed correlation metric learning," *Image and Vision Computing*, vol. 60, pp. 91–97, 2017.

[63] B. Patel, R. Maheshwari, and B. Raman, "Evaluation of periocular features for kinship verification in the wild," *Computer Vision and Image Understanding*, vol. 160, pp. 24–35, 2017.

[64] O. Laiadi, A. Ouamane, E. Boutellaa, A. Benakcha, A. Taleb-Ahmed, and A. Hadid, "Kinship verification from face images in discriminative subspaces of color components," *Multimedia Tools and Applications*, pp. 1–23, 2018.

[65] R. Aliradi, A. Belkhir, A. Ouamane, and A. S. Elmaghraby, "Dieda: discriminative information based on exponential discriminant analysis combined with local features representation for face and kinship verification," *Multimedia Tools and Applications*, pp. 1–18, 2018.

[66] Y.-G. Zhao, Z. Song, F. Zheng, and L. Shao, "Learning a multiple kernel similarity metric for kinship verification," *Information Sciences*, vol. 430, pp. 247–260, 2018.

[67] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, pp. 886–893, 2005.

[68] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[69] J.-G. Wang, J. Li, C. Y. Lee, and W.-Y. Yau, "Dense sift and gabor descriptors-based face representation with applications to gender recognition," in *Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on*. IEEE, pp. 1860–1864, 2010.

[70] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[71] X. Wu, E. Boutellaa, M. B. López, X. Feng, and A. Hadid, "On the usefulness of color for kinship verification from face images," in *Information Forensics and Security (WIFS), 2016 IEEE International Workshop on*. IEEE, pp. 1–6, 2016.

[72] S. Xia, M. Shao, and Y. Fu, "Kinship verification through transfer learning," in *IJCAI*, pp. 2539–2544, 2011.

[73] J. Lu, J. Hu, V. E. Liong, X. Zhou, A. Bottino, I. U. Islam, T. F. Vieira, X. Qin, X. Tan, S. Chen *et al.*, "The fg 2015 kinship verification in the wild evaluation," in *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, vol. 1. IEEE, pp. 1–7, 2015.

[74] Y. Fang, Y. Y. S. Chen, H. Wang, and C. Shu, "Sparse similarity metric learning for kinship verification," in *Visual Communications and Image Processing (VCIP), 2016*. IEEE, pp. 1–4, 2016.

[75] J. Hu, J. Lu, Y.-P. Tan, J. Yuan, and J. Zhou, "Local large-margin multi-metric learning for face and kinship verification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2017.

[76] M. Kan, S. Shan, D. Xu, and X. Chen, "Side-information based linear discriminant analysis for face recognition." in *BMVC*, vol. 11, pp. 125–1, 2011.

[77] A. Dehghan, E. G. Ortiz, R. Villegas, and M. Shah, "Who do i look like? determining parent-offspring resemblance via gated autoencoders," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1757–1764, 2014.

[78] L. Li, X. Feng, X. Wu, Z. Xia, and A. Hadid, "Kinship verification from faces via similarity metric based convolutional neural network," in *International Conference Image Analysis and Recognition*. Springer, pp. 539–548, 2016.

[79] K. Zhang12, Y. Huang, C. Song, H. Wu, L. Wang, and S. M. Intelligence, "Kinship verification with deep convolutional neural networks," in *In British Machine Vision Conference (BMVC)*, sep, 2015.

[80] M. Dawson, A. Zisserman, and C. Nellåker, "From same photo: Cheating on visual kinship challenges," in *Asian Conference on Computer Vision*. Springer, 2018, pp. 654–668.

[81] X. Zhou, K. Jin, M. Xu, and G. Guo, "Learning deep compact similarity metric for kinship verification from face images," *Information Fusion*, vol. 48, pp. 84–94, 2019.

[82] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *Josa a*, vol. 4, no. 3, pp. 519–524, 1987.

[83] H. Yan and J. Lu, "Introduction to facial kinship verification," in *Facial Kinship Verification*. Springer, pp. 1–5, 2017.

[84] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," Yale University New Haven United States, Tech. Rep., 1997.

[85] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[86] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[87] L. Wolf, T. Hassner, and Y. Taigman, "Descriptor based methods in the wild," in *Workshop on faces in'real-life'images: Detection, alignment, and recognition*, 2008.

[88] S.-R. Zhou, J.-P. Yin, and J.-M. Zhang, "Local binary pattern (lbp) and local phase quantization (lbq) based on gabor filter for face representation," *Neurocomputing*, vol. 116, pp. 260–264, 2013.

[89] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in *International conference on image and signal processing*. Springer, pp. 236–243, 2008.

[90] J. Kannala and E. Rahtu, "Bsif: Binarized statistical image features," in *Pattern Recognition (ICPR), 2012 21st International Conference on*. IEEE, pp. 1363–1366, 2012.

[91] L. Zhang, R. Chu, S. Xiang, S. Liao, and S. Z. Li, "Face detection based on multi-block lbp representation," in *International conference on biometrics*. Springer, pp. 11–18, 2007.

[92] C. Shu, X. Ding, and C. Fang, "Histogram of the oriented gradient for face recognition," *Tsinghua Science and Technology*, vol. 16, no. 2, pp. 216–224, 2011.

[93] X. Zhou, J. Lu, J. Hu, and Y. Shang, "Gabor-based gradient orientation pyramid for kinship verification under uncontrolled environments," in *Proceedings*

of the 20th ACM international conference on Multimedia. ACM, pp. 725–728, 2012.

[94] J. Hu, J. Lu, Y.-P. Tan, J. Yuan, and J. Zhou, "Local large-margin multi-metric learning for face and kinship verification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2017.

[95] Y. Duan, J. Lu, J. Feng, and J. Zhou, "Context-aware local binary feature learning for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[96] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[97] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision.* Springer, pp. 818–833, 2014.

[98] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *Journal of Big Data*, vol. 2, no. 1, p. 1, 2015.

[99] C. J. Ng and A. B. J. Teoh, "Dctnet: A simple learning-free approach for face recognition," in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2015 Asia-Pacific.* IEEE, pp. 761–768, 2015.

[100] S. Kakarwal and R. Deshmukh, "Wavelet transform based feature extraction for face recognition," *IJCSA Issue-I June*, vol. 9740767, 2010.

[101] V. Vapnik, "The nature of statistical learning theory springer new york google scholar," 1995.

[102] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[103] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "Liblinear: A library for large linear classification," *Journal of machine learning research*, vol. 9, no. Aug, pp. 1871–1874, 2008.

[104] L. Nanni, A. Lumini, and S. Brahnam, "Ensemble of texture descriptors for face recognition obtained by varying feature transforms and preprocessing approaches," *Applied Soft Computing*, vol. 61, pp. 8–16, 2017.

[105] B. Kulis *et al.*, "Metric learning: A survey," *Foundations and Trends® in Machine Learning*, vol. 5, no. 4, pp. 287–364, 2013.

[106] S. Theodoridis, K. Koutroumbas *et al.*, "Pattern recognition," *IEEE Transactions on Neural Networks*, vol. 19, no. 2, p. 376, 2008.

[107] E. Vezzetti and F. Marcolin, *Similarity measures for face recognition.* Bentham Science Publishers, 2015.

[108] G. Roffo, S. Melzi, and M. Cristani, "Infinite feature selection," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec, pp. 4202–4210, 2015.

[109] G. Roffo, S. Melzi, U. Castellani, and A. Vinciarelli, "Infinite latent feature selection: A probabilistic latent graph-based ranking approach," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1398–1406, 2017.

[110] G. Roffo and S. Melzi, "Ranking to learn," in *International Workshop on New Frontiers in Mining Complex Patterns.* Springer, pp. 19–35, 2016.

[111] C. M. Bishop *et al.*, *Neural networks for pattern recognition.* Oxford university press, 1995.

[112] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (roc) curve." *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

[113] A. Ross, K. Nandakumar, and A. Jain, "Handbook of multibiometrics (international series on biometrics). secaucus," 2006.

[114] A. Ross and A. Jain, "Information fusion in biometrics," *Pattern recognition letters*, vol. 24, no. 13, pp. 2115–2125, 2003.

[115] J. Lu, X. Zhou, Y.-P. Tan, Y. Shang, and J. Zhou, "Neighborhood repulsed metric learning for kinship verification," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 2, pp. 331–345, 2014.

[116] R. Kohavi *et al.*, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Ijcai*, vol. 14, no. 2. Montreal, Canada, pp. 1137–1145, 1995.

[117] Y. Bengio and Y. Grandvalet, "No unbiased estimator of the variance of k-fold cross-validation," *Journal of machine learning research*, vol. 5, no. Sep, pp. 1089–1105, 2004.