République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Larbi Tébessi - Tébessa
Faculté des Sciences Exactes et des Sciences de la Nature et de la vie

Département : Mathématiques et Informatique

Mémoire de fin d'étude

Pour l'obtention du diplôme de MASTER

Domaine : Mathématiques et Informatique

Filière : Informatique

Option : Systèmes d'information

Thème

La vérification automatique de parenté basée sur le visage

Présenté Par :
Mohamed Hamidane

Devant le jury :

Dr. Abdeljalil Gattal	*MCA*	*Université Larbi Tébessi*	*Président*
Dr. Salima Bourougaa	*MCB*	*Université Larbi Tébessi*	*Examinateur*
Dr. Issam Bendib	*MCB*	*Université Larbi Tébessi*	*Encadreur*
Dr. Abdallah Meraoumia	*MCA*	*Université Larbi Tébessi*	*Co-Encadreur*

Thème

La vérification automatique de parenté basée sur le visage

Mohamed Hamidane

Université de Larbi Tébessi

Résumé

Dans ces derniers temps, la vérification de la parenté à partir du visage humain a attiré beaucoup d'attention dans les domaines de la vision par ordinateur et de la reconnaissance de formes en raison de ses applications importantes dans de nombreux domaines tels que la recherche historique et généalogique, science médico-légale, recherche des membres de la famille disparus, analyse des médias sociaux, et l'identification des proches à partir d'une collection de photos et récupération d'images. De nombreuses approches ont été proposées mais les performances ne sont pas encore atteintes vu aux difficultés reconnues dans la conception des systèmes de vérification de parenté. Dans ce mémoire, nous avons proposé deux schémas de vérification de parenté. Le premier schéma est basé sur quatre algorithmes d'extraction de caractéristiques à savoir le LBP, HOG, TPLBP et FPLBP. Une phase de classification utilisant le classifieur KPPV basé sur la métrique de similarité cosinus est utilisée afin de tester les caractéristiques extraites. Alors, deux architectures basées sur l'apprentissage profond (CNN) sont proposées dans le deuxième schéma de vérification de parenté. En validant ces schémas de vérification de parenté sur les deux bases de données KinFaceW-I et KinFaceW-II, nous avons dégagé une amélioration considérable des taux de vérification de parenté en comparaison aux récentes approches.

Mots clés : Vérification de parenté, apprentissage automatique, apprentissage profond, prédicteur landmarks, texture, régions principales, motif binaire local, histogramme des gradients orientés, motif binaire local à trois patchs, motif binaire local à quatre patchs, réseaux neuronaux convolutionnelles, similarité cosinus, KinFaceW-I, KinFaceW-II.

Abstrac

In recent times, kinship facial human verification gained a huge attention in the fields of computer vision and pattern recognition due to its important applications in many fields such as historical research and genealogy, forensic science, tracing missing family members, social media analysis, and identifying family members from photo collection and image recovery. Many approaches have been proposed but the performance has not yet been achieved given the difficulties recognized in the design of kinship verification systems. In this thesis, we have proposed two kinship verification schemes. The first scheme is based on three feature extraction algorithms namely LBP, HOG, TPLBP and FPLBP. A classification phase using the KPPV classifier based on the cosine similarity metric is used to test the extracted features. Then, two architectures based on deep learning (CNN) are presented in the second kinship verification scheme. By validating the proposed kinship verification schemes on the two databases KinFaceW-I and KinFaceW-II, we have found a considerable improvement in kinship verification rates compared to recent approaches.

Keywords: Kinship verification, machine learning, deep learning, landmarks predictor, texture, main regions, LBP, HOG, TPLBP, FPLBP, CNN, cosine similarity, KinFaceW-I, KinFaceW-II.

ملخص

في الآونة الأخيرة، اكتسب التحقق من القرابة من الوجه البشري الكثير من الاهتمام في مجالات رؤية الكمبيوتر والتعرف على الأنماط نظرًا لتطبيقاته المهمة في العديد من المجالات مثل علم الأنساب، الطب الشرعي، تتبع أفراد الأسرة المفقودين، تحليل وسائل التواصل الاجتماعي، وتحديد القرابة من الصور المسترجعة. حاليا تم اقتراح العديد من الأنظمة الخاصة بالتحقق من القرابة ولكن النتائج بقيت مقبولة لحد ما نظرًا لصعوبة تحديد القرابة من خلال الصور. في هذه الأطروحة، اقترحنا شكلين للتحقق من القرابة. يعتمد الشكل الأول على ثلاث خوارزميات لاستخراج الميزات وهي النموذج الثنائي المحلي، المدرج البياني للتدرجات الموجهة، النموذج الثنائي المحلي مع ثلاث بقع والنموذج الثنائي المحلي مع أربع بقع. في هذه المرحلة تم استخدام مصنف كي أقرب جار مستعملين مقياس تشابه جيب التمام لاختبار الميزات المستخرجة. بعد ذلك، تم اقتراح بنيتين تستندان على التعلم العميق والشبكات العصبية التلافيفية في الشكل الثاني. من خلال نتائج الطرق المقترحة في كلا الشكلين معتمدين على قاعدتي البيانات KinFaceW-I وKinFaceW-II، وجدنا تحسنًا كبيرًا في نسب التحقق مقارنة بالمقترحات الحديثة.

الكلمات المفتاحية: التحقق من القرابة، التعلم الالي، التعلم العميق، متنبئ المعالم، الملمس، المناطق الأساسية، النموذج الثنائي المحلي، مدرج بياني للتدرجات الموجهة، النموذج الثنائي المحلي مع ثلاث بقع، النموذج الثنائي المحلي مع أربع بقع، الشبكات العصبية التلافيفية، تشابه جيب التمام، KinFaceW-I، KinFaceW-II.

.

Dédicace



Je dédie ce travail :

***À** ceux ; qui m'ont tant donné sans rien demander, qui m'ont toujours offert leur soutient*

m'ont épargné aucun effort pour m'aider, qui m'ont appris mes principes,

à qui tous les mots ne suffisent pas pour les

remercier :

« Mes très chers parents »

Que Dieu vous protège.

***À** mes frères et à ma sœur, spécialement "Youssef" et "Chourouk"*

avec mes souhaits de bonheur et de prospérité.

***À** toute ma famille maternelle et paternelle.*

***À** tous mes chers ami(e)s surtout" Zied & Adnene & Chouaib ". Pour leur aide et*

leur soutien moral durant l'élaboration du travail de fin d'études.

***À** mes encadreurs pour leur aide précieuse.*

***À** mes collègues.*

***À** toute personne, qui m'a aidée à réaliser ce modeste travail, de proche ou de loin.*

Mohamed Hamidane

Remerciement

Mes remerciements vont tout premièrement à dieu tout puissant pour la volonté, la santé et la patience qui m'a donné durant tous ces années d'études.

Je remercie vivement les membres de jury qui ont accepté d'examiner et pour l'honneur qu'ils me font en jugeant mes travaux et pour avoir accepté de participer au jury de ce travail.

Je tiens à présenter mes reconnaissances à mes encadrantes Dr. I. Bendib, Dr. A. Meraoumia pour leurs conseils, leurs orientations, et la disponibilité qu'il m'a accordée pour faire réussir ce travail, tout au long de la période.

Je profite cette occasion pour exprimer ma gratitude à ces personnes pour leurs contributions comme mes collègues, ma famille et mes amis.

Je voudrai exprimer tout d'abord ma gratitude à ma belle-famille et sur tout, à ma mère, mon père, mes sœurs. Ils sont toujours intéressés à mes études avec le succès et m'ont aidé beaucoup plus qu'elles ne peuvent le croire.

Je voudrais également adresser mes plus sincères remerciements à tous mes proches et amis qui m'ont toujours soutenue et encouragée au cours de la réalisation de cette mémoire. Vraiment merci.

Merci à vous tous !

TABLE DES MATIERES

Résumé

Abstract

ملخص

Introduction Générale 1

CHAPITRE 01

L'IA et les liens de parenté

1. Introduction 5

2. Les liens de parenté en théorie 5

2.1. Terminologies utilisées 5

2.2. Les types de liens de parenté 6

2.3. Les degrés des liens de parenté 7

3. Les liens de parenté et vision par ordinateur 8

3.1. Vérification et reconnaissance des liens de parenté 8

3.2. Reconnaissance de la parenté vs reconnaissance faciale 9

3.2.1. L'objectif 9

3.2.2. Le mode de fonctionnement 10

3.2.3. Les performances 11

3.2.4. Les défis 11

3.2.5. Applications 11

4. Processus du système de vérification et reconnaissance des liens de parenté 11

4.1. Phase de détection des visages et prétraitement 12

4.2. Phase d'extraction et sélection des caractéristiques 12

4.3. Phase de reconnaissance et de vérification 13

5. Défis et difficultés de la reconnaissance des liens de parenté 14

5.1. Variations d'âge et de sexe 15

5.2. Changement d'illumination 16

5.3. Variations de pose et de profil 17

5.4. Expressions faciales 17

5.5. Occlusions 18

6. Domaines d'application 18

7. Conclusion 20

CHAPITRE 02

Méthodes et techniques pour la reconnaissance automatique des liens de parenté

1. Introduction 22

2. Les types de systèmes de reconnaissance automatique 22

2.1. Les systèmes de reconnaissance classiques 22

2.1.1. Extraction de caractéristiques 23

2.1.2. La classification 29

2.2. Les systèmes de reconnaissance basés sur l'apprentissage profond 31

2.2.1. Fonctionnement de l'apprentissage profond 31

2.2.2. Les applications du l'apprentissage profond 32

2.2.3. Les modèles de réseaux de neurones profonds (Architecteurs) 33

3. Les réseaux de neurones convolutifs (CNNs) 34

3.1. Principe 34

3.2. La phase de convolution 35

3.2.1. La couche convolution (CONV) 35

3.2.2. La couche pooling (POOL) 36

3.2.3. La couche d'activation 36

3.3. La phase d'évaluation ou reconnaissance 37

3.3.1. La couche de mise à plat « *flattening* » 37

3.3.2. La couche dropout 37

3.3.3. Couche entièrement connectée (Fully-Connected) 38

3.4. Choix des hyperparamètres de CNNs 39

3.4.1. Nombre de filtres 39

3.4.2. Forme de filtres 39

3.4.3. Forme de Max pooling 40

3.5. Quelques réseaux convolutifs célèbres (pré-entraînés) 40

3.5.1. LeNet 40

3.5.2. AlexNet 40

3.5.3. VGGNet 41

3.5.4. ResNet 42

4. Mesures d'évaluation de performance des systèmes 43

3.1. Matrice de confusion 43

3.2. Courbe ROC (*Receiver Operating Characteristic*) 43

3.3. Les métriques d'évaluation 44

5. Travaux connexes pour la vérification automatique des liens de parenté 45
4.1. Les systèmes basés sur les caractéristiques locales 45
4.2. Les systèmes à base d'apprentissage profond 47
4.3. Les systèmes basés sur l'apprentissage des métriques 47
6. Conclusion 51

CHAPITRE 03 Conception et résultats expérimentaux

1. Introduction 53
2. Méthodologie et méthodes proposées 53
2.1. Vérification de parenté basée sur les schémas classiques d'extraction de caractéristiques 53
2.1.1. Extraction du vecteur de caractéristiques 54
2.1.2. Sélection de caractéristiques 54
2.1.3. La classification 56
2.2. Vérification de parenté basée sur l'apprentissage en profondeur 56
2.2.1. Vérification de parenté basée sur le CNN basique 57
2.2.2. Vérification de parenté basée sur les points clés du visage (KP-CNN) 58
3. Expérimentations et résultats 62
3.1. Bases de données 62
3.2. Protocole de tests 64
3.3. Evaluation des performances 65
3.3.1. Performances des méthodes basées sur les techniques d'extraction de caractéristiques classiques 65
3.3.2. Performances des méthodes basées sur l'apprentissage en profondeur 74
4. Etude comparative 77
5. Conclusion 78
Conclusion Générale 79
Bibliographie 80

LISTE DES Tables

Page

TABLEAU 2.1 : Matrice de confusion dans le cas d'une classification binaire 43

TABLEAU 2.2 : Etude comparatifs entre les performances des travaux susmentionnés avec les bases KinFaceW-I et KinFaceW-II 50

TABLEAU 3.1 : Plages d'index des numéros de visages pour 5-folds des ensembles d'images de KinFaceW-I et KinFaceW-II 64

TABLEAU 3.2 : Performance de la vérification de parenté utilisant le descripteur local **LBP** sur la base KinFaceW-I 66

TABLEAU 3.3 : Performance de la vérification de parenté utilisant le descripteur local **LBP** sur la base KinFaceW-II 66

TABLEAU 3.4 : Performance de la vérification de parenté utilisant le descripteur local **HOG** sur la base KinFaceW-I 68

TABLEAU 3.5 : Performance de la vérification de parenté utilisant le descripteur local **HOG** sur la base KinFaceW-II 68

TABLEAU 3.6 : Performance de la vérification de parenté utilisant le descripteur local **TPLBP** sur la base KinFaceW-I 70

TABLEAU 3.7 : Performance de la vérification de parenté utilisant le descripteur local **TPLBP** sur la base KinFaceW-II 71

TABLEAU 3.8 : Meilleurs performances de la vérification de parenté utilisant le descripteur local **FPLBP** sur les deux bases KinFaceW-I et II 73

TABLEAU 3.9 : Performance de la vérification de parenté basée **CNN** basique dans la base KinFaceW-I 74

TABLEAU 3.10 : Performance de la vérification de parenté basée **CNN** basique dans la base KinFaceW-II 75

TABLEAU 3.11 : Performance de la vérification de parenté basée **KP-CNN** dans la base KinFaceW-I 76

TABLEAU 3.12 : Performance de la vérification de parenté basée **KP-CNN** dans la base KinFaceW-II 77

TABLEAU 3.13 : Comparaison des méthodes proposées 78

LISTE DES FIGURES

Page

FIGURE 1.1 : Organigramme des liens de parenté proches 7

FIGURE 1.2 : Comparaison entre (A) système de vérification de parenté et (B) système d'identification de parenté 9

FIGURE 1.3 : Modes d'identification faciale et de parenté 10

FIGURE 1.4 : Système général automatisé de vérification et de reconnaissance de la Parenté... 14

FIGURE 1.5 : Exemples d'images de visage non contrôlées 15

FIGURE 1.6 : Exemple d'images de visage avec des différences d'âge et de sexe 16

FIGURE 1.7 : Exemple de variation d'éclairage 17

FIGURE 1.8 : Exemples de variation de poses de la tête 17

FIGURE 1.9 : Exemples de variation d'expressions faciales 18

FIGURE 1.10 : Exemples d'occlusion du visage 18

FIGURE 2.1 : Les étapes d'un système de reconnaissance automatique classique 22

FIGURE 2.2 : Exemple d'extraction des caractéristiques en utilisant l'opérateur LBP basique 24

FIGURE 2.3 : Trois voisinages pour des R et P différents 24

FIGURE 2.4 : Images réelles des modalités biométriques **(b)** Images LBP basique correspondantes 25

FIGURE 2.5 : L'image réelle de visage. **(b)** Le code TPLBP calculé avec les paramètres S= 8, w= 3 et α= 2. **(c)** La représentation TPLBP d'une image réelle de visage 26

FIGURE 2.6 : L'image réelle de visage. **(b)** Le code FPLBP calculé avec les paramètres S= 8, α= 1. **(c)** La représentation FPLBP d'une image réelle de visage 27

FIGURE 2.7 : Exemple d'extraction l'histogramme avec le descripteur HOG 29

FIGURE 2.8 : Principe de la classification en utilisant la technique de SVM 30

FIGURE 2.9 : La tâche de la classification en utilisant l'algorithme K-Nearest Neighbors (KNN). *Pour k = 1 la classe majoritaire du point central est la classe A* 31

FIGURE 2.10 : Le procède du ML traditionnel comparé à celui du Deep Learning 31

FIGURE 2.11 : Approche d'apprentissage en profondeur 32

FIGURE 2.12 : Les deux parties qui composent l'architecture des réseaux neuronaux convolutifs 34

FIGURE 2.13 : Exemple de calcul de la convolution 35

FIGURE 2.14 : Exemple avec types de pooling (avec un filtre "**2x2**" et un pas de "**2**") 36

FIGURE 2.15 : La couche ReLU 36

FIGURE 2.16 : Illustration de fonctionnement du dropout 37

FIGURE 2.17 : Illustration du rôle de la couche flattening 38

FIGURE 2.18 : Principe de la couche entièrement connectée (FC) 38

FIGURE 2.19 : Architecture de LeNet 40

FIGURE 2.20 : Architecture de AlexNet 41

FIGURE 2.21 : Architecture de VGGNet 42

FIGURE 2.22 : Architecture de ResNet 42

FIGURE 2.23 : Courbe ROC (Receiver Operating Characteristic) 44

FIGURE 3.1 **:** Schéma général de vérification de parenté basé sur les méthodes classiques d'extraction de caractéristiques 54

FIGURE 3.2 **:** Architecture proposée du **CNN** de base pour la vérification de parenté 58

FIGURE 3.3 **:** Les 68 points clés du visage 58

FIGURE 3.4 **:** Exemple des régions utilisées 59

FIGURE 3.5 **:** Architecture proposée du **KP-CNN** pour la vérification de parenté 59

FIGURE 3.6 **:** Architecture proposée du modèle **Eyes_CNN** 60

FIGURE 3.7 : Architecture proposée du modèle **Nose_CNN** 61

FIGURE 3.8 **:** Architecture proposée du modèle **Mouth_CNN** 62

FIGURE 3.9 **:** Exemples d'image du visage : (a) KinFaceW-I et (b) KinFaceW-II 63

FIGURE 3.10 : Les quatre types de relations : père-fille, père-fils, mère-fils et mère-fille, qui définit les deux bases des images utilisées 63

FIGURE 3.11 : Courbes ROC représentant les précisions de la vérification de parenté utilisant le descripteur **LBP** dans les bases (a) KinFaceW-I et (b) KinFaceW-II 67

FIGURE 3.12 : Courbes ROC représentant les précisions de la vérification de parenté utilisant le descripteur **HOG** dans les bases (a) KinFaceW-I et (b) KinFaceW-II 69

FIGURE 3.13 : Courbes ROC représentant les précisions de la vérification de parenté utilisant le descripteur **TPLBP** dans les bases (a) KinFaceW-I et (b) KinFaceW-II 72

FIGURE 3.14 : Courbes ROC représentant les précisions de la vérification de parenté utilisant le descripteur **FPLBP** dans les bases (a) KinFaceW-I et (b) KinFaceW-II 73

Liste Des Abréviations

ADN: Aacide Désoxyribo Nucléique

IA: Intelligent Artificial

M-S: Mother-Son

F-S: Father-Son

LBP: Local Binary Patterns

TPLBP: Three-Patch Local Binary Patterns

FPLBP: Four-Patch Local Binary Patterns

CNN: Convolutional Neural Network

SVM: Support Vector Machine

ML: Machine Learning

DNN: Deep Neural Networks

LSTM: Long Short-Term Memory

AlexNet: Alex Network

ResNet: Residual Network

PDFL: Prototype-based Discriminative Feature Learning

MPDFL: Multi-view Prototype-based Discriminative Feature Learning

OSL-A: online similarity learning algorithm

DDMML: Discriminative Deep Metric Multidimensional Learning

MNRML: Multi-view Neighborhood repulsed correlation metric learning

ESL: Ensemble Similarity Learning

LLMMML: Local Large-Margin Multi-Metric Learning

KinFaceW-II: Kin Face in the Wild-II

KinFaceW-I: Kin Face in the Wild-I

M-D: Mother-Daughter

F-D: Father-Daughter

KIN: Kinship

HOG: Histogram of oriented gradients

ACP : Analyse en Composantes Principales

KPPV : K Plus Proches Voisins

ANN: Artificiel Neural Networks

KP-CNN: Key Points Convolutional Neural Network

KNN: K Nearest Neighbors

RNN: Recurrent Neural Network

LeNet: LeCun Network

VGGNet: Visual Geometry Group Neural Network

ROC: Receiver Operating Characteristic

IFVF: Inheritable Fisher Vector feature

DDML: Discriminative Deep Metric Learning

NRML: Neighborhood repulsed correlation metric learning

DMML: Discriminative Multi Metric Learning

LMMML: Local Margin Multi-Metric Learning

KML: Kinship Metric Learning

Introduction Générale

Le progrès technologique, notamment dans les nouvelles techniques de traitement de l'information et les algorithmes intelligents ainsi que les capacités de calculs ont permis de développer des solutions informatiques basées sur l'intelligence artificielle pour la résolution d'un grand nombre de problèmes causés par l'insécurité et les accès non autorisés ainsi que la gestion d'immense volume de multimédia éparpillée dans les réseaux sociaux et les plateformes de partages de données.

Au cours des dernières années de nombreuses technologies biométriques ont été développées qui sont basées sur les identificateurs biométriques physiologiques et comportementaux tels que les empreintes digitales, le visage, l'iris, la voix, et les signatures numériques. La reconnaissance des visages est parmi les technologies biométriques les plus utilisées et les plus adaptées ainsi qu'elle est très utilisée dans le domaine de vision par ordinateur. En effet, elle joue un rôle central en biométrie, interface homme-machine, contrôle d'accès, indexation multimédia, etc. En littératures, plusieurs méthodes de reconnaissance des paramètres biométriques comme le visage ont été proposées durant ces dernières années. Ces applications exploitent généralement deux modalités : la reconnaissance à partir de séquence d'images (vidéo) et la reconnaissance à partir d'images fixes.

Malgré que la reconnaissance de visages basée sur la vidéo est la plus utilisée à celle basée sur des images fixes, puisqu'elles exploitent les informations temporelles et spatiales dans le processus de reconnaissance automatique. Nous trouvons plusieurs domaines d'applications qui intéressent à la reconnaissance basée sur les images fixes comme l'identification et la reconnaissance automatique des individus. On peut répartir l'ensemble des techniques de reconnaissance de visages basées sur les images fixes, en trois grandes catégories : les méthodes globales, les méthodes locales et les méthodes hybrides. L'analyse des images faciales est l'un des principaux thèmes de recherche dans le traitement d'images, la vision par ordinateur. En effet, les images faciales contiennent de nombreuses informations utiles sur la personne à laquelle ils appartiennent, notamment l'identité, sexe, ethnicité, âge, et expressions émotionnelles, etc.

Entre autres, l'utilisation à l'échelle des réseaux sociaux et les plateformes de communications et de partages de données ont engagé des nouvelles problématiques dans le

traitement des images faciales telles que la recherche des tendances de relation entre les groupes ou les personnes. Ces relations potentielles dans les photos incluent celles des parents, des collègues et des amis. La vérification et la reconnaissance des liens de parenté sont de nouvelles pistes de recherche en vision par ordinateur qui ont fait l'objet d'une attention croissante ces dernières années.

En revanche, motivés par les découvertes psychologiques, plusieurs scientifiques ont étudié la capacité des évaluateurs humains ou la perception de reconnaître la parenté à partir d'images faciales humaines. De plus, la relation biologique et les similitudes entre les traits trouvés dans la même famille ont inspiré à les exploiter pour développer des systèmes informatiques qui seront capables de reconnaître l'existence d'un lien de parenté et de l'appliquer dans une variété de domaines d'applications.

Les systèmes de détection des liens de parenté établissent des relations entre individus et groupes sur le modèle des relations biologiques entre parents et enfants, entre frères et sœurs et entre partenaires conjugaux. Encore une fois, en raison des progrès de la psychologie, du développement biologique unique et d'autres facteurs, l'attention des chercheurs s'est déplacée pour proposer un algorithme qui détermine si une paire d'images faciales appartenant à une catégorie de lien de parenté ou non. L'analyse des liens de la parenté à travers l'image faciale humaine est une tâche complexe et les travaux de recherches effectués dans cette piste sont prometteurs. Fondamentalement, les difficultés à reconnaître la parenté peuvent être liées à la parenté elle-même et aussi liées à l'environnement de la base de données utilisée.

Dans ce contexte, dans cette étude nous avons opté pour le développement d'un système intelligent qui permet la vérification automatique de l'existence d'un lien de parenté. À cet effet, nous avons proposé d'utiliser les nouvelles méthodes d'apprentissage automatique pour l'élaboration d'un modèle informatique intelligent pour la détection des liens de parenté. En pratique, il n'existe pas des formules ou des caractéristiques fixes qui permettent la détection de l'existence d'un lien de parenté. Pour cela, nous exploitons les capacités des modèles de la « Machine Learning » pour surmonter cette problématique.

Entre autres, l'exploitation des informations du contexte pour la recherche d'un lien de parenté entre deux images faciales est un choix ambitieux. En effet, les similitudes entre deux images avec un lien de parenté sont généralement dans la forme de la bouche, les yeux, le nez et aussi la couleur de la peau. Dans ce contexte, nous avons utilisé ces informations de contexte dans le développement de notre système de détection automatique de lien de parenté.

Outre l'introduction générale dans laquelle nous mettons en relief le contexte et la problématique principale et les objectifs traités dans ce mémoire, le présent document est structuré en trois chapitres :

- Le premier chapitre traite les notions fondamentales des liens de la parenté et passer en revue les différents types et les degrés de la parenté en théorie. Il décrit également la vision de l'ordinateur et les liens de la parenté, ensuit nous introduisons une présentation sur le système général automatisé de vérification et de reconnaissance de la parenté et également déterminer les défis et les difficultés qui diminue les performances du système et nous mentionnons finalement certains domaines d'applications avec des motivations sur ce sujet.

- Dans le deuxième chapitre nous présentons une description détaillée sur les techniques de l'apprentissage automatique et l'apprentissage profond où nous exposera ses différentes architectures des réseaux de neurones et en particulièrement les réseaux de neurones convolutionnels (CNNs), ses applications et ses fondements. Nous introduisons ensuite une étude bibliographique sur les travaux connexes dans cette problématique avec une synthèse sous forme d'un tableau récapitulatif.

- Dans le troisième chapitre, nous présentons notre contribution pour le développement d'un système automatique de détection des liens de parenté. Dans ce contexte, nous avons proposé trois variantes d'architectures, une basée sur un modèle d'apprentissage classique et les deux autres sur un modèle d'apprentissage profond. Ensuite, nous avons effectué une série de tests empiriques pour détecter les paramètres optimaux des systèmes proposés et nous discutons les résultats obtenus. Enfin, nous dressons une conclusion générale avec les perspectives futures.

Chapitre I

1. Introduction

En domaine de psychologie, les travaux de recherches ont indiqué que les traits du visage sont des indices fiables et critiques pour mesurer la similarité génétique par rapport aux autres mesures telles que les mesures biologiques. En effet, les tests ADN, par exemple, sont largement utilisés dans la détermination de la paternité et de la scène du crime des enquêtes, mais ce processus est fastidieux, lent et couteux. Entre temps, l'intelligence artificielle « IA » et les techniques de vision par ordinateurs permettent de trouver des solutions intelligentes pour la détection des liens de parenté à base de visage. Ces possibilités ont motivé les travaux de recherches dans les domaines de biométrie, reconnaissances faciales et la vision par ordinateur de développé des modèles informatiques pour ajouter une nouvelle dimension intelligente, rapide et économique pour la vérification de la similarité génétique des personnes via des images faciales.

Le visage est l'élément le plus important que les gens utilisent pour se reconnaître. Les humains peuvent s'identifier rapidement et facilement par leurs images de visage. Cependant, ces compétences sont affectées par les changements importants dans les caractéristiques du visage, ou même par les modalités de capture comme l'angle de vision, l'éclairage, etc. Le visage reste une approche de reconnaissance dynamique de l'homme. L'analyse de l'image faciale a été largement étudiée dans la communauté de l'IA et l'imagerie et un grand nombre de ces méthodes ont été proposées pour diverses applications pratiques et les grands succès sont enregistrés dans les applications de la reconnaissance faciale. Cependant, nous trouvons de nouvelles pistes de recherche prometteuse dans le domaine d'applications pour le visage tel que la vérification automatique de la parenté et les systèmes d'estimation d'âge.

Dans ce chapitre nous essayons de présenter les généralités de base des liens de parenté et ces modalités d'applications dans le domaine de l'IA. Ainsi, nous focalisons sur les défis relatifs à la détection des liens de parenté à partir des images de visages et leurs domaines d'applications.

2. Liens de parenté en théorie

Nous présentons dans cette section la terminologie et les différents concepts de parenté utilisés dans cette étude.

2.1. Terminologies utilisées

- **En général**

Le terme parenté indique une similitude, une familiarité ou une proximité entre les

entités sur la base de certains ou de tous les traits ou caractéristique de base, peut être formulé ou qualifié de relation de parenté. La parenté est un terme large qui a différents moules de sens en fonction du cadre ou du contexte dans lequel elle opère [1].

Dans le contexte anthropologique

La parenté fait référence au réseau de relations sociales qui constituent une partie importante de la vie de la plupart des humains dans la plupart des sociétés. L'étude de la parenté est l'étude de ce que l'homme fait des faits fondamentaux de la vie - l'accouplement, la gestation, la parentalité, la socialisation, la fratrie, etc. Néanmoins, dans d'autres contextes, la parenté peut avoir différentes significations [2].

En biologie

La parenté se réfère généralement au degré de parenté génétique ou au coefficient de relation entre les membres individuels d'une même espèce [3, 4]. De plus, différentes sociétés classent différemment les relations de parenté et utilisent donc différents systèmes de terminologie de parenté [5].

2.2. Types de liens de parenté

Généralement nous trouvons que la majorité des spécialistes des sciences sociales classent les liens de parenté dans la vie humaine en trois aspects ou types.

- **Liens de sang (Consanguineal) :** Cette parenté est basée sur le sang ou les liens génétiques, la relation signifiant la relation entre parents et enfants également entre frères et sœurs, dit le groupe de sociologie. C'est le type de parenté fondamental et le plus universel. Également connu sous le nom de parenté primaire, il implique des personnes directement liées [6].
- **Liens de mariage (Affinal) :** Cette parenté est basée sur le mariage. La relation entre mari et femme est également considérée comme une forme fondamentale de parenté [6].
- **Liens sociaux (Social) :** Cette parenté concerne les personnes qui partagent des liens de parenté à travers des affiliations religieuses ou les groupes sociaux. La différence majeure de ce lien par rapport aux deux précédents se réside dans la capacité de mettre en fin cette relation de manière absolue sans recours juridique [6].

Dans notre projet, nous concentrerons sur le premier type qui est le lien de sang, parce que dans notre objectif est la recherche d'un système intelligent qui permet la détection de l'existence des liens de parenté basée sur le sang et particulièrement celles qui sont entre les

parents et leurs enfants c.-à-d. les quatre relations : père-fils, père-fille, mère-fille, mère-fils.

2.3. Les degrés des liens de parenté

Toute relation entre deux individus est basée sur le degré de proximité ou de distance de cette relation. Le degré de proximité ou de distance de toute relation dépend de la façon dont les individus sont liés entre eux. Les liens de parenté sont essentiellement sous trois classes (ou degré), qui sont :

- **Parenté primaire :** Ce degré fait référence à des parents qui sont étroitement et directement liés les uns aux autres, est appelé les principaux parents. Les huit types de parenté primaire comprennent la mère-fils, le mari-femme, le père-fils, la mère-fille, le père-fille, la sœur-frère, la sœur cadette-sœur aînée et le frère cadet-frère aîné [7].
- **Parenté secondaire :** Ce degré fait référence aux parents principaux des parents primaires. En d'autres termes, ceux qui sont directement liés aux parents primaires (également appelé parents secondaires). Il y a 33 parents secondaires, comme le frère du père, le mari de la sœur et la femme du frère [7].
- **Parenté tertiaire :** Ce degré se réfère à la parenté secondaire de la parenté primaire (également connue sous le nom de parent tertiaire). Environ 151 parents tertiaires ont été identifiés, comme le frère du mari de la sœur et la femme du beau-frère [7].

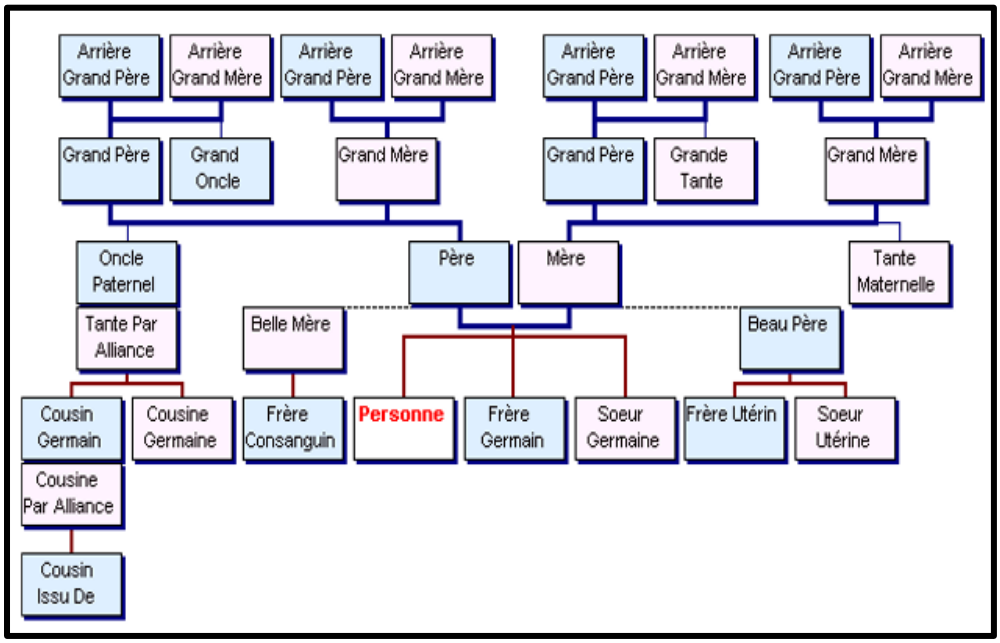


Figure 1.1 : Organigramme des liens de parenté proches (Liens de sang - ou la naissance) [8].

3. Liens de parenté et vision par ordinateur

Dans les systèmes de la reconnaissance automatique de parenté, l'objectif est d'apprendre la machine à reconnaître les relations à base de caractéristiques entre des paires d'images de visages. En effet, cette tâche de reconnaissance peut être une tâche de détection binaire à partir de traits extraits d'images faciales ou une tache de reconnaissance pour déterminer le degré exact de cette relation [9].

3.1. Vérification et reconnaissance des liens de parenté

La détection des liens de parenté peut fonctionner dans deux modes qui sont :

- **Mode de vérification (authentification)** : nommée aussi mode d'authentification, le système agit par vérification de l'existence d'une relation de parenté entre une paire d'images.
- **Mode de reconnaissance (identification)** : nommée aussi mode d'identification, son but est de deviner la relation de parenté des deux images d'entrées (classe de relation).

Jusqu'à présent, de nombreux travaux de recherches sont concentrés sur le développement des méthodes pour déterminer la meilleure solution pour vérifier la relation de parenté entre les individus [10, 11, 12]. En effet, peu de travaux qui sont intéressés sur l'identification de degré des liens de parenté [13].

Afin de cerner la relation entre la vérification et l'identification automatique, nous supposons les situations suivantes : « Si nous utilisons une photo qui contient un groupe de personnes (ou un individu) comme galerie et que nous utilisons une autre photo comme requête », alors, à ce stade, trois cas sont générés (**Figure.1.2**), comme suit :

- ✓ Y a-t-il une ressemblance entre la photo de requête et les membres situés sur la photo ? (***Vérification***).
- ✓ Si oui, existe-t-il une relation entre la photo de requête et les membres situés sur la photo ? (***Vérification***)
- ✓ Si oui, identifiez le type de cette relation (***Reconnaissance*** *ou* ***Identification***).

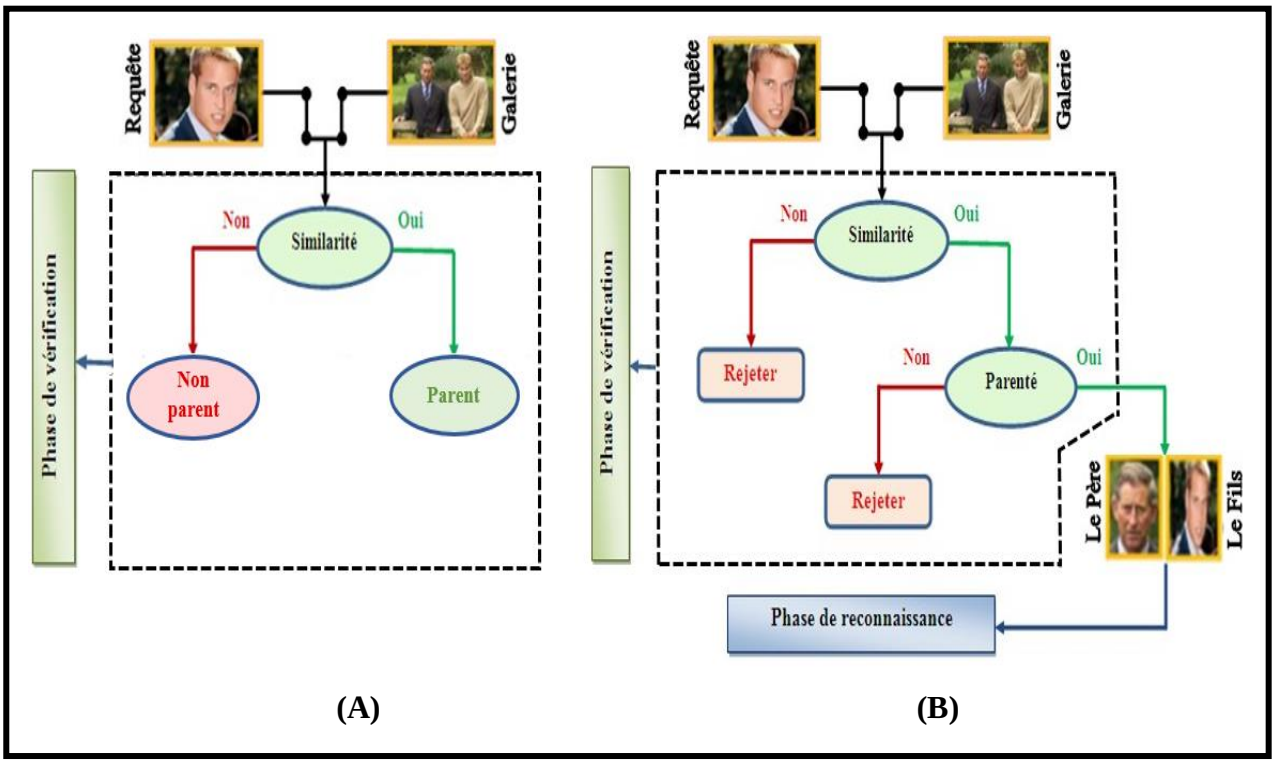


Figure 1.2 : Comparaison entre (A) système de vérification de parenté et (B) système d'identification de parenté [1].

3.2. Reconnaissance de la parenté VS reconnaissance faciale

Malgré que le processus de reconnaissance faciale puisse sembler similaire à celui de la de la reconnaissance de parenté, alors que ces deux processus sont deux choses différentes. On pourrait dire que la reconnaissance de la parenté est le plus haut niveau que la reconnaissance faciale. Dans le même temps, nous ne pouvons pas ignorer l'existence de facteurs communs entre eux, tels que les défis d'extractions de caractéristiques discriminantes et les techniques utilisées de classification.

La différence entre la reconnaissance faciale et la reconnaissance de parenté peut être exprimé en fonction de critères suivants :

3.2.1. L'objectif

- ✓ L'objectif du système de reconnaissance faciale basée sur des images peut être défini comme suit : étant donné une image faciale d'entrée et une base de données d'images faciales d'individus connus, la tâche du système est de distinguer les caractéristiques faciales de l'image faciale d'entrée et l'image faciale d'un individu connu de la base de données ainsi que de vérifier ou identifier si l'image appartient à la même personne ou non [14, 15].
- ✓ En revanche, l'objectif d'un système de reconnaissance de parenté consiste tout d'abord

de trouver des caractéristiques permettant à la fois de vérifier les liens de parenté entre deux images et de les distinguer aux images de la base [16, 12].

3.2.2. Le mode de fonctionnement

Les deux systèmes englobent le mode de vérification et d'identification, avec une différence dans le style de mise en œuvre. La **figure. 1.3** montre les modes de reconnaissance faciale et de parenté.

- ✓ Vérification du visage : est-ce la même personne ? (**1 : 1**)
- ✓ Identification du visage : qui est cette personne ? (**1 : N**)
- ✓ Vérification de parenté : vérifie l'existence d'une relation de parenté entre une paire d'images ou plus (**1 : N**)
- ✓ Identification de la parenté : identifie le degré de parenté (**1 : N**).

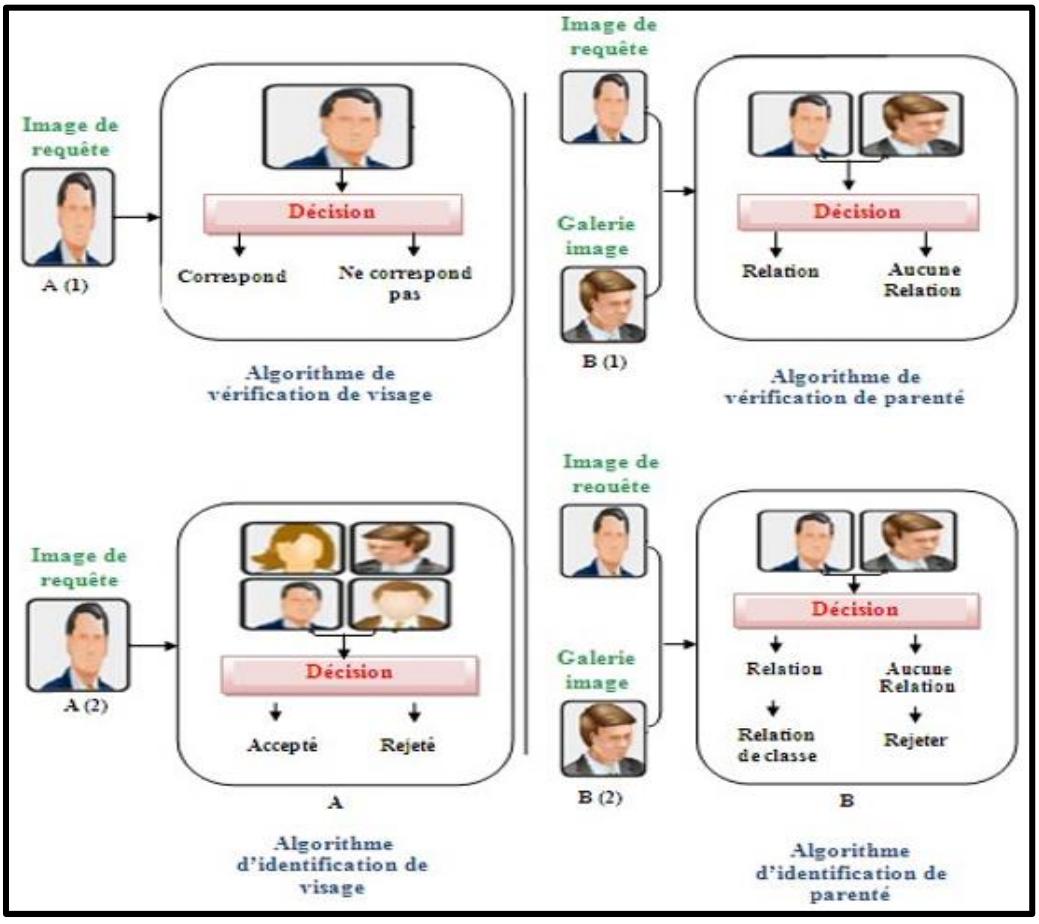


Figure 1.3 : Modes d'identification faciale et de parenté : **A** identité (A(1) vérification faciale, A(2) identification faciale), et **B** parenté (B(1) vérification de parenté, B(2) identification de parenté) [1].

Dans la **figure 1.3**, on observe l'existence d'un lien entre la vérification et l'identification dans le système de reconnaissance de parenté, ce qui signifie que le système ne peut

procéder à la détermination du degré de parenté sans avoir recours à vérifier l'existence de la relation de parenté.

3.2.3. Les performances

Les performances des machines en termes de reconnaissance faciale sont devenues très précises, pouvant atteindre **97,25%** [17]. Bien que la précision actuelle pour la reconnaissance de la parenté soit de **88,40%** [18], cette précision n'est pas satisfaisante et a maintenant ouvert un nouveau domaine de recherche.

3.2.4. Les défis

Généralement les défis majeurs sont communs entre la reconnaissance de l'identité et la détection des liens de parenté. Ces systèmes peuvent être affectés par un ensemble d'influences telles que l'éclairage, l'absence et la présence de composants structurels, les occlusions, les émotions, l'expression faciale, le maquillage, la chirurgie esthétique, le vieillissement, les rotations, les variations de pose, le bruit, l'échelle, l'encombrement et la résolution, des jumeaux, etc. [19, 20, 21]. Il s'agit des défis les plus importants et les plus courants dans ces deux domaines.

3.2.5. Applications

Les applications les plus courantes et familières pour la reconnaissance faciale dans divers domaines sont la surveillance, la sécurité, le divertissement, le contrôle d'accès, le suivi des suspects et la médecine légale. La reconnaissance de parenté est l'aspect le plus intéressant lié à la relation des personnes dans les images et a une influence significative sur les applications réelles et de nombreux domaines.

La différence que nous pouvons observer est que la sécurité et la surveillance dominent la plupart des applications de reconnaissance faciale, tandis que le la parenté vise à étudier la relation biologique entre les personnes [22].

4. Processus du système de vérification et reconnaissance des liens de parenté

Le processus de reconnaissance automatisée de la parenté sur les images faciales est effectué en examinant des caractéristiques importantes trouvées dans les traits du visage et en les comparant avec ceux qui ont été trouvés dans d'un autre visage. Dans la plupart des cas, l'entrée dans le système de vérification et de reconnaissance de la parenté est toujours des images fixes. Deux images servent d'entrées (une *paire de visages*), l'une est une image pour un parent (*mère ou père*) et l'autre est une image de l'un des enfants (*fils ou fille*). Le système

automatisé de vérification et de reconnaissance de la parenté passe par six phases pour vérifier l'existence de toute relation de parenté, et reconnaître la classe de cette relation.

4.1. Phase de détection des visages et prétraitement

Cette phase est souvent utilisée pour la détection des zones d'intérêts (le visage). En effet, cette tâche utilise les techniques de reconnaissance faciale afin de détecter automatiquement le contour ou l'objet de visage. Cependant, les dataset ou les corpus utilisés dans les systèmes de détection de lien de parenté contiennent généralement des images prétraitées. Des images contiennent uniquement le visage avec orientation uniforme. Ces ressources sont utilisées pour la création et l'évaluation des performances des systèmes de détection de lien de parenté. Tandis que, la phase de détection du visage est primordiale lors de l'implémentation de ces systèmes en mode réel.

4.2. Phase d'extraction et sélection des caractéristiques

L'extraction et la sélection des caractéristiques représentatives, pertinentes et discriminantes dans une tache de classification automatique est le grand défi pour les systèmes de reconnaissances automatiques. En littératures, nous trouvons plusieurs techniques utilisées pour cette phase qui seront détaillées si dessous. Ces techniques sont définies sur la base d'une seule entrée, mais les systèmes de détection de lien de parenté possèdent une paire d'images en entrée ce qui rend cette phase plus complexe. En effet en cherchant des caractéristiques communes et discriminantes pour deux images très différentes dans plusieurs dimensions telles que l'âge sexe.

Généralement, les méthodes d'extraction de caractéristiques peuvent être divisé en trois groupes : basé sur les méthodes d'extraction classique (Handcraft), caractéristiques appris (learned Handcraft) et basé sur l'apprentissage en profondeur (deep learning).

- **Méthodes classiques d'extraction de caractéristiques**

L'image contient de nombreuses caractéristiques qui peuvent être utilisées pour analyser une telle scène. Le type de caractéristiques utilisé varie en fonction de l'application prévue. Par exemple, dans une image contenant une scène composée de formes A, B et C, plusieurs méthodes doit être utilisé pour extraire chaque forme séparément. Par conséquent, il est nécessaire de savoir le type de scène pour choisir efficacement la méthode appropriée. Par conséquent, comme ces les techniques nécessitent une expertise, les caractéristiques obtenues sont considérées comme fabriquées à la main. L'une des faiblesses de ces méthodes est que les caractéristiques de l'image sont calculées selon un ensemble d'étapes

préconçues et ne sont pas appris le contexte dans lequel l'image appartient. Comme exemples de certaines fonctionnalités artisanales méthodes d'extraction, DCT, DWT, LBP, PCA, ICA et autres.

❖ **Méthodes d'extraction de caractéristiques apprises**

Pour améliorer l'efficacité des méthodes d'extraction de caractéristiques classique, de nombreux chercheurs ont tenté d'incorporer l'idée d'apprentissage dans l'extraction de caractéristiques pour extraire des vecteurs de caractéristiques plus précis. L'idée générale est de donner des connaissances (informations) préalables pour extraire efficacement vecteurs de caractéristiques. À ce jour, la façon dont ces informations sont présentées varie de chercheur à l'autre. En fait, de nombreux travaux dans ce sens se sont concentrés sur la représentation faciale ont été relativement bien exécutés. À titre d'exemples de ces méthodes, modèle quantifié local (LQP), descripteur de visage discriminant (DFD), contexte d'apprentissage de caractéristiques binaires locales (CA-LBFL) et autres.

❖ **Méthodes basées sur l'apprentissage en profondeur**

L'apprentissage profond ou apprentissage en profondeur (en anglais : deep learning, deep structured learning, hierarchical learning) est un ensemble de méthodes d'apprentissage automatique tentant de modéliser avec un haut niveau d'abstraction des données grâce à des architectures articulées de différentes transformations non linéaires. Ces techniques ont permis des progrès importants et rapides dans les domaines de l'analyse du signal sonore ou visuel et notamment de la reconnaissance faciale, de la reconnaissance vocale, de la vision par ordinateur, du traitement automatisé du langage. Il existe un grand nombre de variables d'architectures profondes, citons par exemple CNN, PCANet, DCTNet…

L'étape qui suit est l'étape de sélection, dont ses techniques sont utilisées pour réduire les dimensions des caractéristiques pour les calculs ainsi que l'élimination des bruits et des informations redondantes.

4.3. Phase de reconnaissance et de vérification

Dans cette phase, nous récurons aux algorithmes de classification automatiques ou les systèmes d'apprentissage approfondi pour la recherche de l'existence d'éventuelle relation de lien de parenté entres les paires d'images d'entrées. Certes, les performances des deux phases précédentes influent fortement sur les résultats de classification, cependant le choix des classifieurs approprié en tenant compte de la spécificité de la tâche améliore les performances de ces systèmes.

De point de vue des résultats des systèmes de détection des liens de parenté à partir de visage, elle se varie une classification binaire : existence ou non-existence de lien de parenté vers une classification multi-classes qui permet de détecter le type de relation. Ce choix est souvent basé sur les caractéristiques de corpus de données utilisées pour le développement.

L'ensemble des étapes du système de la parenté peut ainsi être résumé dans la **figure 1.4** suivante :

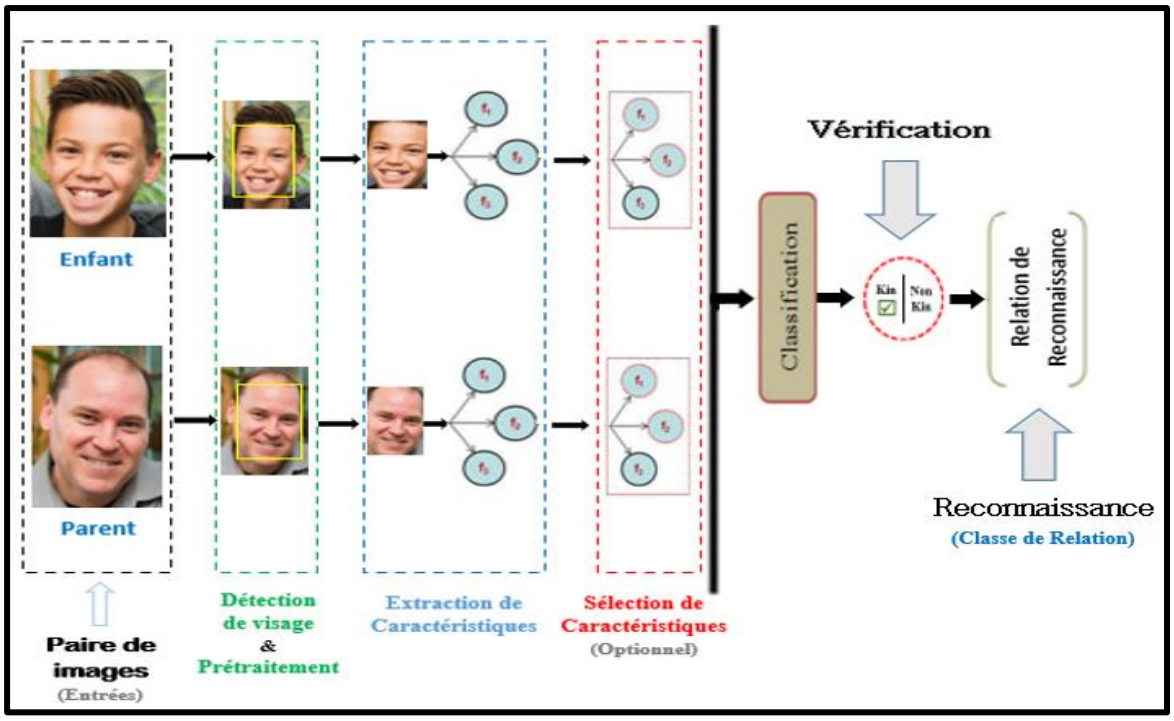


Figure 1.4 : Système général automatisé de vérification et de reconnaissance de la parenté.

5. Défis et difficultés de la reconnaissance des liens de parenté

La recherche des solutions efficaces pour la reconnaissance des liens de parenté est limitée par les défis et les difficultés qui entravent la réalisation de ces systèmes. Bien que le visage humain soit un sujet avec une grande variété de changements intrinsèques et extrinsèques, les systèmes de traitement automatique de visage ou bien spécifiquement le système de reconnaissance de lien de parenté à travers des images faciales sont construits à un nombre limité d'échantillons de visage généralement capturés dans des conditions contrôlées. La **figure 1.5** illustre certaines images de visage difficiles capturées dans la nature (conditions incontrôlées). Il est pratiquement impossible de couvrir tous les changements faciaux pendant la phase de l'apprentissage, ce qui empêche les systèmes de gérer les visages invisibles avec de nouveaux changements. De plus, des études sont effectuées [23] ont démontré que des variations en termes (changement d'illumination, pose de la tête, occultation… etc.) dans différentes images de visage d'une même personne peut être plus importantes que les variations

de visages de personnes différentes. Par conséquent, les performances d'analyse de visage se dégradent remarquablement dans des environnements défavorables.

Cependant, les défis caractérisent le domaine de la reconnaissance de lien de parenté comme un problème plus large et plus difficile que la vérification et l'identification du visage. De plus, des études [24] ont démontré spécifiquement dans ces systèmes que les défis apparaissent en raison des variations (âge, sexe, changement d'illumination, pose de la tête, expressions faciales, occultations, similitude minimale entre parents, etc.). Dans les sections suivantes, nous présentons les principaux difficultés et défis.



Figure 1.5 : Exemples d'images de visage non contrôlées [25].

5.1. Variations d'âge et de sexe

L'écart d'apparence rencontré dans le problème de la vérification de parenté est beaucoup plus grand que celui dans le cadre de vérification classique du visage. Etant donné deux images avec des genres différents et des âges différents, **Figure 1.6**, l'objectif est la vérification d'existence d'un lien de parenté « *Père-fille* ». La similitude faciale dépend de plusieurs facteurs, y compris l'âge et le sexe. Dans des études antérieures en tant que proposition et solution au problème de la différence d'âge, le grand écart d'âge entre les aînés et leurs jeunes enfants a été réduit en exploitant l'image des parents âgés lorsqu'ils étaient jeunes (en extrayant les traits lorsqu'ils étaient vieux qui correspondent aux traits lorsqu'ils étaient jeunes) en tant qu'ensemble de données intermédiaires [26].



Figure 1.6 : Exemple d'images de visage avec des différences d'âge et de sexe [1].

Dans le cas du sexe, les résultats d'expériences dans des études antérieures montrent que le pourcentage de précision augmente quand un père est comparé à son fils, et une mère avec sa fille et diminue quand un père est comparé à sa fille, et une mère avec son fils. Cette constatation explique l'effet du sexe sur les performances de l'algorithme. Cette observation n'est pas une règle absolue, mais nous devons en tenir compte. De plus, cela a motivé les recherches d'identification des traits distinctifs qui représentent les différences entre les sexes [1, 27].

5.2. Changement d'illumination

Les variations d'illumination viennent entraîner des variations considérables dans l'apparence d'un visage. Des changements d'illumination peuvent faire apparaître sur le visage des ombres accentuant ou, au contraire, masquant certaines caractéristiques faciales. Deux types d'éclairage peuvent influencer : l'illumination globale (ou ambiante) et l'illumination locale. Alors que l'illumination globale affecte tout le visage de manière uniforme, l'illumination locale entraîne la création de zones d'ombres et de zones éclairées, la **figure 1.7** présente un exemple de visage dont la source lumineuse l'éclairant se déplace. Parmi les solutions proposées dans la littérature au problème de l'éclairage, il y a le développement de descripteurs de visage robustes au changement d'illumination [28].

L'extraction de caractéristiques invariantes aux changements de luminosité est également une approche largement décrite dans la littérature [29]. Notons enfin qu'un certain nombre d'approches traitent le problème de la luminosité dans les systèmes de reconnaissance par une étape de prétraitement dont l'objectif principal est bien souvent de corriger les artefacts dus aux variations de luminosité.



Figure 1.7 : Exemple de variation d'éclairage [30].

5.3. Variations de pose et de profil

Le taux de reconnaissance de lien de parenté se dégrade considérablement dans les cas de variations et de changements des angles d'inclinaison de la tête dans les images à traiter. Cette difficulté a été démontrée par des tests d'évaluation élaborés sur les bases FERET et FRVT [31, 32]. La variation de pose est considérée comme un problème majeur pour le système de reconnaissance de lien de parenté. Quand le visage est de profil dans le plan image (orientation< 30 °), il peut être normalisé en détectant au moins deux traits faciaux (passant par les yeux). Cependant, lorsque la rotation est supérieure à 30°, la normalisation géométrique n'est plus possible. La **figure 1.8** illustre quelques exemples de variation de pose et de profil.



Figure 1.8 : Exemples de variation de poses de la tête.

5.4. Expressions faciales

L'apparence d'un visage varie grandement en présence d'expressions faciales. Les éléments faciaux tels que la bouche ou encore les yeux peuvent alors subir des déformations importantes, pouvant faire échouer un système de reconnaissance de lien de parenté fondée par exemple sur des points d'intérêt. La bouche est en général l'élément facial qui est le plus varie, mais l'aspect des contours peut par exemple être grandement modifié [33]. La **figure 1.9** présente quelques exemples de variation d'expressions faciales.



Figure 1.9 : Exemples de variation d'expressions faciales [30].

5.5. Occlusions

Les occlusions partielles apparaissent fréquemment dans des applications réelles. Elles peuvent être causées par une main cachant une partie du visage, par des cheveux longs, ou par le port d'accessoires tels que des lunettes de vue, de soleil, par tout autre objet (foulard, chapeau . . .), ou encore par une autre personne. Il arrive également qu'une partie du visage cache une autre, comme dans le cas d'une rotation hors plan par exemple. Il est important de savoir reconnaître des visages partiellement occlus. Par conséquent, de nombreux travaux de recherches ont tenté de gérer de telles situations en proposant des approches robustes pour l'occlusion partielle du visage. Des méthodes de sous-espace ont été utilisées pour projeter le visage dans un nouvel espace et éliminer les parties obstruées. Certaines autres méthodes ont tenté de reconstruire les parties du visage occluses, tandis que d'autres détectent l'occlusion et n'utilisent que des parties non occluses pour l'analyse du visage [26]. La **figure 1.10** présente quelques exemples de variation de contraintes d'occlusions.



Figure 1.10 : Exemples d'occlusion du visage.

6. Domaines d'application

Les relations biologiques et les similitudes entre les traits dans les visages dans la même famille ont encourager de trouver des méthodes de reconnaissance automatique des liens de parenté afin de développer des systèmes informatiques qui seront capables de détecter et même

de reconnaître les degrés des liens de parenté et de les appliquer dans des variétés domaines d'applications. Bien qu'il existe d'autres mesures biologiques modernes comme l'ADN, mais elles ne sont pas applicables à grande échelle [10, 34].

La détection de la parenté à travers les visages est devenue un sujet intéressant d'une grande importance scientifique parce que les systèmes de reconnaissance faciale sont devenus plus fiables, actifs et largement appliqués dans des situations en monde réel. Entre autres, avec les progrès technologiques, tels que les appareils photo numériques de haute qualité, des dispositifs mobiles et Internet, les images numériques deviennent une nouvelle dimension d'identité des personnes. L'identification des personnes et de leurs relations à partir des images faciales a des valeurs sociales et commerciales importantes en tenant compte les critères, les approches et les algorithmes adoptés dans les études d'identification de parenté pour chaque domaine d'application. Par conséquent, le choix d'une application qui existe dans le monde réel, comme l'organisation d'albums de famille, peut être le point de départ et la pierre angulaire de base pour résoudre le problème de la parenté.

Dans d'autres domaines de ce sujet, d'autres projets potentiels comme la recherche d'informations ou d'images et l'analyse des médias sociaux, qui s'appuient sur les constatations et les résultats concernant les solutions aux problèmes liés aux relations familiales.

Nous résumons les domaines d'application possibles comme suit :

- Retrouver les membres de la famille disparus.
- Médecine légale.
- Traite / contrebande d'enfants.
- Analyse des réseaux sociaux
- Déterminer les parents d'une victime ou d'un suspect par les organismes d'application de la loi.
- Gestion automatique et étiquetage ou annotation des bases de données d'images.
- Identification des proches à partir d'une collection de photos et récupération d'images.
- Recherches historiques et généalogiques.
- Construction d'un arbre relationnel dynamique à partir de photos et organisation d'albums familiaux.

Les motifs et applications susmentionnés attirent un nombre important de chercheurs pour présenter leurs contributions dans ce domaine, et pour en revenir à nos débuts, l'étude de la parenté est un sujet de grand intérêt scientifique. Plusieurs applications dans ce domaine telles

que l'analyse des réseaux sociaux ou la recherche de membres de la famille disparus peuvent être améliorées avec les développements importants des techniques de vision par ordinateur. Cependant, la recherche sur ce sujet reste complexe et difficile. Et de toute évidence, la recherche dans ce domaine est toujours active [25].

7. Conclusion

Dans ce chapitre, nous avons introduit les notions théoriques des liens de parenté ainsi que les concepts d'intelligence artificielle qui nous intéresse et nécessaire à la bonne compréhension de la reconnaissance automatique des liens de parenté à travers des images faciales. Ensuite, on a donné aussi une vision générale sur le processus du système de vérification et reconnaissance des liens de parenté. Ainsi, nous avons présenté les difficultés et les défis qui affectent l'efficacité et les performances de la reconnaissance de ces systèmes. Enfin, nous avons terminé ce chapitre par la représentation des domaines d'applications et les motivations concernant l'évolution dans ce sujet. Dans le chapitre suivant, nous détaillons les techniques d'apprentissages automatiques utilisées dans notre travail ainsi que les travaux connexes reliés à ce domaine.

Chapitre II

1. Introduction

Les systèmes de reconnaissance automatique basés sur les techniques de *Machine Learning* « ML » ont enregistré des performances remarquables. Ces performances sont relatives aux modalités d'entrées utilisées telles que les textes, les images, les flux de parole et les vidéos, ainsi que la complexité de la tâche a exécuté.

Entre autres, l'évolution importante enregistrée dans les techniques et les algorithmes utilisés dans pour l'apprentissage automatique et celle de l'apprentissage profond couplé avec les puissances de calculs actuelles ont permet d'entamer de problématiques plus complexes comme la reconnaissance automatique des liens de parenté.

Dans ce contexte, nous présentons dans ce chapitre les techniques de ML utilisées pour le développement des systèmes de reconnaissance automatique des liens de parenté à base des images faciales. Ainsi, nous présentons une étude bibliographique sur les travaux réalisés dans ce contexte avec un tableau de synthèse récapitulatif.

2. Les types de systèmes de reconnaissance automatique

Le développement des systèmes de reconnaissance automatique en ML est basé sur la stratégie déployée pour la phase d'apprentissage automatique ainsi que les algorithmes de reconnaissance utilisés. Généralement, dans le domaine de reconnaissance faciale et les liens de parenté nous trouvons trois catégories : les systèmes classiques basés sur les algorithmes d'apprentissage classique de ML, les systèmes basés sur les algorithmes de l'apprentissage profond et les systèmes basés sur l'apprentissage des mesures « *Metric Learning* ». Dans ce mémoire et tenant compte de nos objectifs, nous intéressons sur les deux premiers types de systèmes.

2.1. Les systèmes de reconnaissance classiques

Généralement, un système de reconnaissance automatique qui utilise les techniques d'apprentissage automatique classique comme les de vérification de la parenté est constituée de deux phases principales :

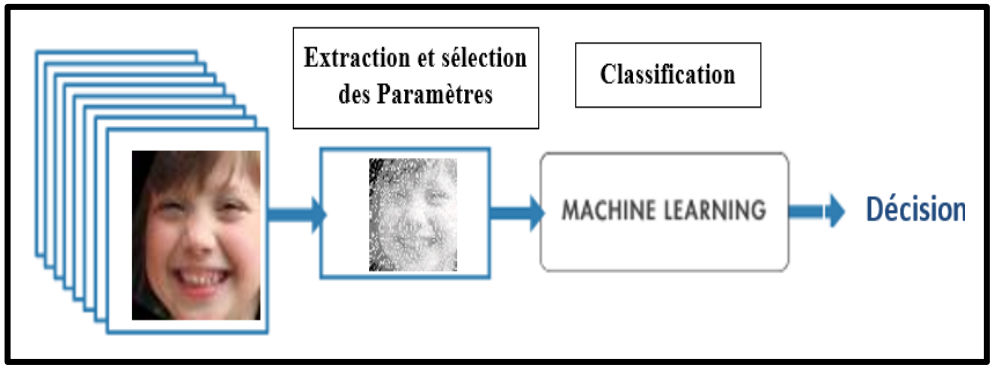


Figure 2.1 : Les étapes d'un système de reconnaissance automatique classique.

2.1.1. Extraction de caractéristiques

Cette étape représente le cœur du système de vérification après la localisation du visage dans l'image faciale. Elle permet l'extraction des informations discriminantes, utiles et non redondantes (vecteur de caractéristiques), à travers la description de points d'intérêts associés aux caractéristiques faciales, telles que le skin, les yeux, les sourcils, le nez et la bouche. Ces derniers sont utilisés pour la construction des modèles de représentation pour les algorithmes de classification et décision. L'efficacité de cette étape à une influence directe sur les performances des systèmes de vérification des liens de parenté. L'extraction des caractéristiques relatives aux visages utilise plusieurs méthodes qui sont classées généralement en trois approches (globales, locales et hybrides), parmi lesquelles, nous citons celles que nous avons utilisés lors de développement de notre contribution pour la détection automatique des liens de parenté, qui sont : LBP « *Local Binary Pattern* », TPLBP « *Three-Patch Local Binary Patterns* », FPLBP « *Four-Patch Local Binary Patterns* », HOG « *Histogram of oriented gradients* ».

a. Motif binaire local (LBP)

Cette technique était initialement proposée par *Ojala* [35]. Afin de caractériser les textures présentes dans des images en niveaux de gris. Elle était largement utilisée dans diverses applications notamment la reconnaissance faciale. Le principe général de cette technique est de comparer le niveau de gris de chaque pixel avec les valeurs des niveaux de ses huit voisins dans un voisinage 3 × 3 en soustrayant la valeur du pixel central. Afin de générer un motif binaire. Les valeurs strictement négatives obtenues sont codées avec un "zéro" (0) et les autres avec un "un" (1). Les pixels de ce motif binaire sont alors multipliés par des poids donnés par les puissances de deux (2 poids) et sommés afin d'obtenir un code LBP du pixel central ou courant. Ainsi, nous obtenons pour chaque image avec des pixels dont l'intensité est comprise entre 0 et 255 et elles utilisent les histogrammes pour le choix des descripteurs [36]. La **figure 2.2** illustre le principe de cette méthode.

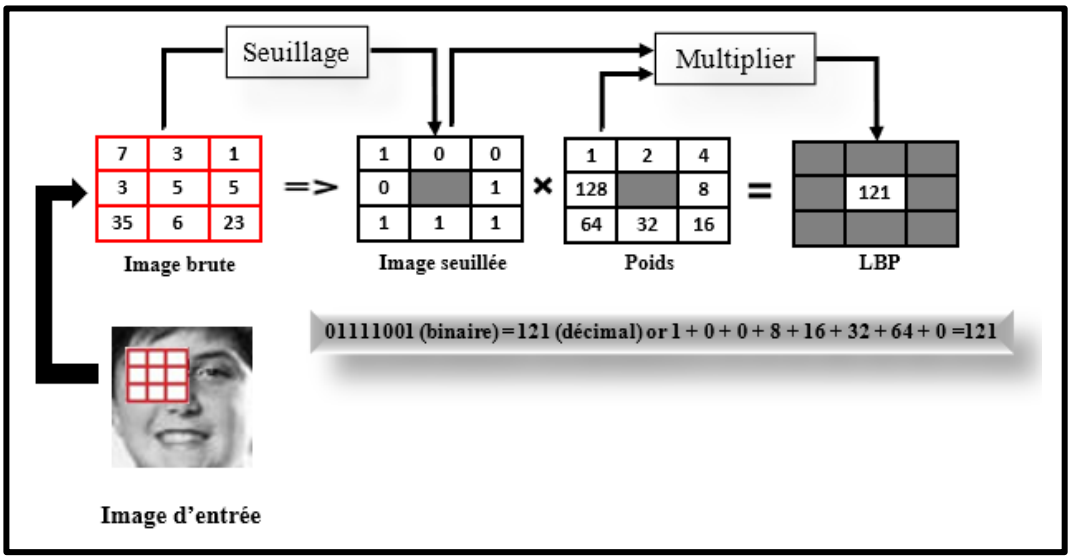


Figure 2.2 : Exemple d'extraction des caractéristiques en utilisant l'opérateur LBP basique.

Cependant, afin de traiter la texture à différentes échelles et de capturer les caractéristiques dominantes avec des structures à grande échelle. La technique LBP a été étendue en utilisant des voisinages de taille déférente (multi-échelle) [37]. Dans ce cas, un cercle de rayon R autour du pixel central et les valeurs des P points échantillonnés sur le bord de ce cercle sont prises et comparées avec la valeur du pixel central. La **figure 2.3**, illustre trois voisinages pour des valeurs de R et P différentes.

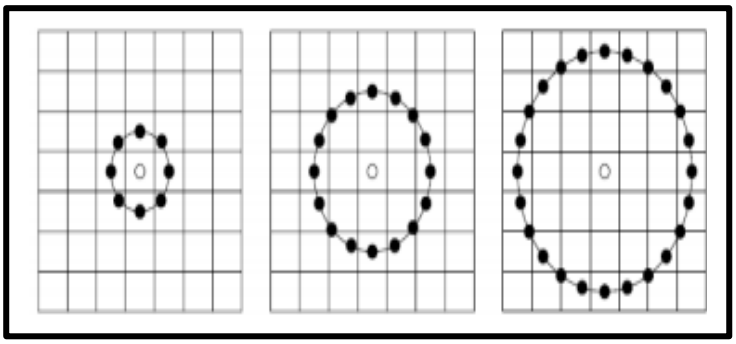


Figure 2.3 : Trois voisinages pour des R et P différents [36].

Formellement, en donnant un pixel central en coordonnées (x_c, y_c) avec le niveau de gris g_c dans l'image d'entrée, ses niveaux gris qui correspondent à P pixels voisins sur un cercle de rayon R sont $(g0, g1, \dots, gP-1)$. Le LBP résultant peut être exprimé sous forme décimale du pixel g_c comme suit :

$$\boldsymbol{LBP(x_c, y_c)_{P,R} = \sum_{i=0}^{p-1} u(g_i{}^{P,R} - g_c).2^i} \qquad (2.\ 1)$$

Où ***u(x)*** est la fonction de seuillage, donnée par :

$$\boldsymbol{u(x) = \begin{cases} 1 & si\ x \geq 0 \\ 0 & autrement \end{cases}} \qquad (2.\ 2)$$

La **figure** ci-dessous illustrée des exemples de quelques modalités biométriques avec leurs images LBP basique.

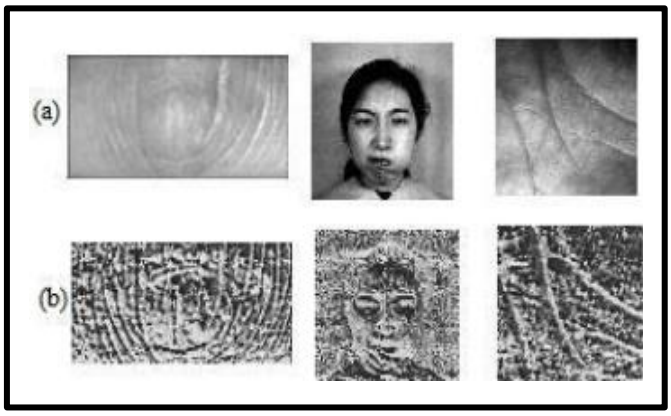


Figure 2.4 : (a) Images réelles des modalités biométriques **(b)** Images LBP basique correspondantes [38].

En effet, pour les résultats obtenus dans les applications du monde réel, la propriété la plus importante du descripteur LBP c'est qu'il est invariant contre les changements monotones du niveau gris causés comme les changements d'éclairage. Aussi sa facilité de calcul, qui permet l'analyse des images complexe.

b. LBP à trois patchs (TPLBP)

Le LBP à trois patches TPLBP « *Three-Patch Local Binary Patterns* » est une variante de LBP proposée par *Wolf, et al* [39]. Cette méthode est basée sur la comparaison des valeurs de trois patchs pour produire une valeur d'un seul bit dans le code attribué à chaque pixel. Pour chaque pixel de l'image, elle considère un patch (w × w) centré sur le pixel, et S patches supplémentaires distribuées uniformément dans un anneau de rayon r. Pour le paramètre α, elle utilise des paires de patchs α (le diamètre du cercle). La valeur d'un seul bit est définie la similitude des deux patches par rapport au patch central. Le code résultant a des bits S par pixel. Le TPLBP donne des résultats prometteurs en reconnaissance faciale. Cette méthode est donnée par l'équation suivante :

$$\boldsymbol{TPLBP}_{r,S,w,\alpha}(p) = \sum_i^S f\left(d(C_i, C_p) - d(C_{i+\alpha \ mod \ S}, C_p)\right).2^i \qquad (2.\ 3)$$

Où C_i et $C_{i+\alpha \ mod \ S}$ sont deux patchs le long de l'anneau et C_p est le patch central, d est toute fonction de distance entre deux patchs, et f est une fonction de seuillage qui est calculée comme suit :

$$f(x) = \begin{cases} 1 & if \ x \geq \tau \\ 0 & if \ x < \tau \end{cases} \qquad (2.4)$$

La valeur τ : est le seuil de comparaison, et généralement supérieure à zéro (p. ex., $\tau = 0{,}01$) pour assurer une certaine stabilité dans les régions uniformes [40]. Le principe de la méthode d'opérateur TPLBP est illustré à la **figure 2.5**.

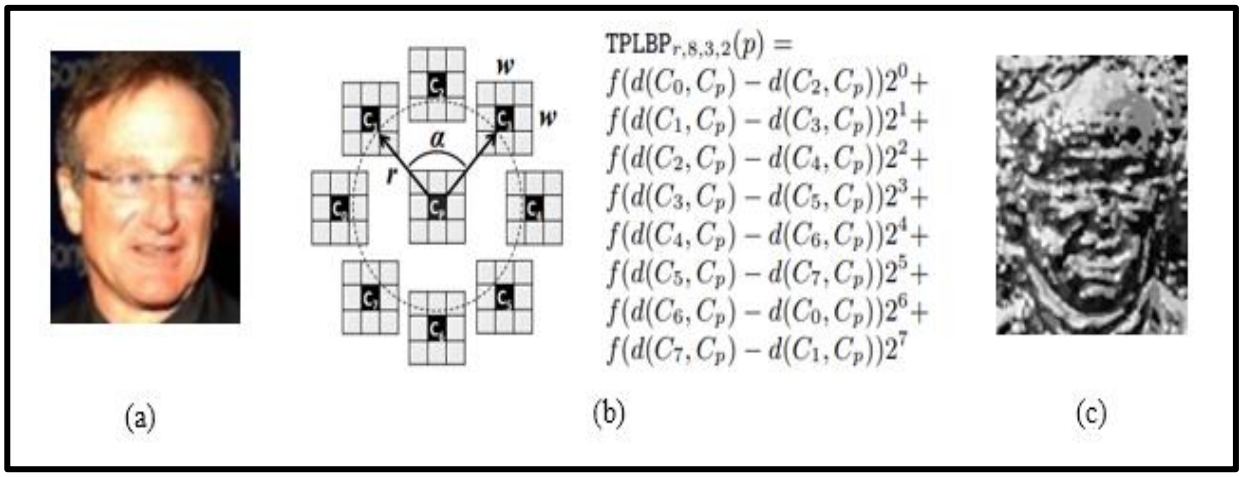


Figure 2.5 : (a) L'image réelle de visage. **(b)** Le code TPLBP calculé avec les paramètres S= 8, w= 3 et α= 2. **(c)** La représentation TPLBP d'une image réelle de visage [39].

Chaque image est divisée en une grille de régions qui ne se chevauchent pas et un histogramme de chaque code binaire est calculé pour chacune de ces régions. Chacun de ces histogrammes a été normalisées à la longueur de l'unité, leurs valeurs ont été tronquées à 0.2, puis normalisées de nouveau à la longueur unitaire. Finalement, l'image est représentée par ces histogrammes concaténés en un seul vecteur.

c. LBP à quatre patchs (FPLBP)

La technique LBP à quatre patchs FPLBP « *Four-Patch Local Binary Patterns* » est une autre variante de LBP, elle est presque la même que TPLBP avec une différence dans le nombre d'anneaux. Avec FPLBP, deux anneaux avec rayon r1 et r2 sont utilisés et centrés dans le pixel, et S patchs de taille (w × w) distribués autour de ces deux anneaux. Pour produire les codes (FPLBP) à quatre patchs, elle compare deux patchs symétriques centraux dans l'anneau intérieur avec deux patchs symétriques centraux dans l'anneau extérieur. Après la comparaison, un bit dans chaque pixel est pris en compte en fonction de laquelle des deux patchs sont les plus similaires. Ainsi, pour les S patchs le long de chaque cercle, S/2 est le nombre de paires symétriques centrales et qui est aussi la longueur des codes binaires produits [39]. La définition formelle d'opérateur FPLBP est la suivante :

$$\boldsymbol{FPLBP}_{r_1,r_2,S,w,\alpha}(p) = \sum_i^{S/2} f\left(d\left(C_{1i}, C_{2,i+\alpha \bmod S}\right) - d\left(C_{1,i+S/2}, C_{2,i+\frac{S}{2}+\alpha \bmod S}\right)\right).2^i \qquad (2.5)$$

Le principe de la méthode FPLBP est illustré à la **figure 2.6**.

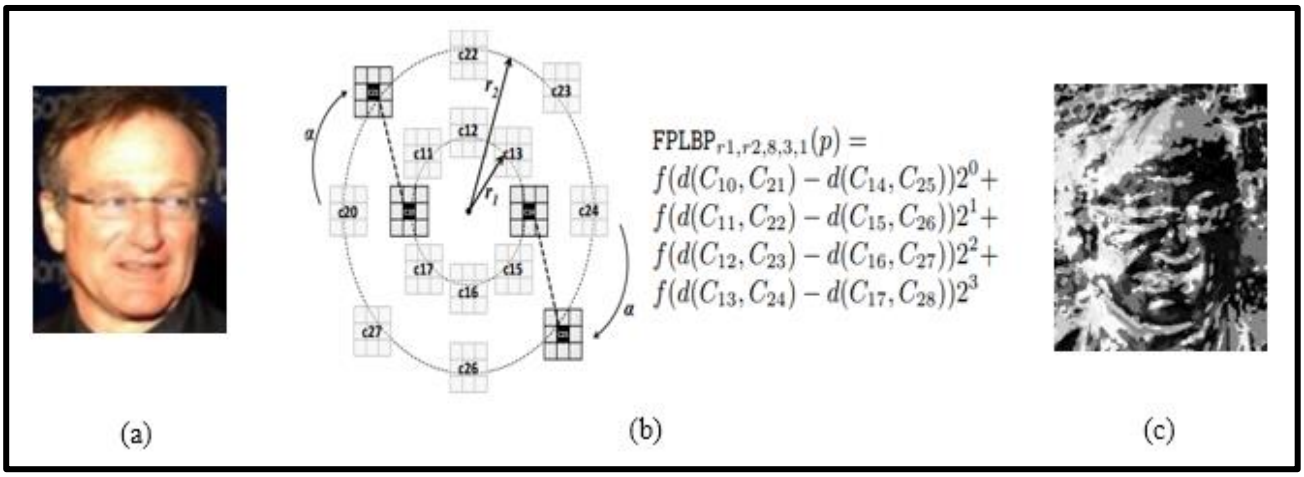


Figure 2.6 : (a) L'image réelle de visage. **(b)** Le code FPLBP calculé avec les paramètres S= 8, α= 1. **(c)** La représentation FPLBP d'une image réelle de visage [39].

d. Histogramme des gradients orientés (HOG)

L'histogramme des gradients orientés HOG « *Histogram of oriented gradients* » introduit par *Dalal et Triggs* [41], est actuellement l'une des méthodes les plus performantes et les plus utilisées pour l'apparence et la forme locale d'un objet dans une image. Il peut être décrit par la distribution de l'intensité du gradient ou la direction des contours. Ceci peut être fait par la division de l'image en petites régions adjacentes, appelées cellules, et le calcul pour chaque cellule l'histogramme des directions du gradient (orientations des contours) pour les pixels à l'intérieur de cette cellule. La concaténation de ces histogrammes représente alors le vecteur de caractéristiques du descripteur « HOG ». Pour de meilleurs résultats, les histogrammes locaux sont normalisés en contraste, par le calcul d'une mesure de l'intensité sur des groupes de cellules adjacentes, appelées des blocs. Cette dernière est utilisée pour normaliser toutes les cellules du bloc. L'algorithme de cette technique contient les étapes suivantes :

1- **Prétraitement :** dans cette étape les images d'entrées sont premièrement converties en niveaux de gris après une normalisation selon l'information de luminance. Ensuite, l'image I (x, y) est divisée en N cellules régulières (N petites régions) et des blocs. Les valeurs de

gradient sont calculées pour chaque pixel en avec un filtre dérivatif 1-D centré, dans les directions horizontales et verticales. Pour cela les masques suivants sont utilisés :

$$\boldsymbol{D_x = [-1 \quad 0 \quad 1]} \qquad (2.6)$$

$$\boldsymbol{D_Y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}} \qquad (2.7)$$

$$\begin{cases} \boldsymbol{G_x(x,y) = I(x,y) * D_x} \\ \boldsymbol{G_y(x,y) = I(x,y) * D_y} \end{cases} \qquad (2.8)$$

2- **Calcul du gradient :** Le gradient d'un pixel (x, y) d'une image est un vecteur caractérisé par son amplitude et sa direction. Le calcul du gradient est une étape critique dans la formation de descripteurs.

L'amplitude d'une image se calcule comme suit :

$$\boldsymbol{|G(x,y)| = \sqrt{G_x(x,y)^2 + G_y(x,y)^2}} \qquad (2.9)$$

Et la direction du gradient est donnée par :

$$\boldsymbol{\alpha(x,y) = arctan\left(\frac{G_x(x,y)}{G_y(x,y)}\right)} \qquad (2.10)$$

Où G_x (x, y) et G_y (x, y) sont les gradients dans la direction horizontale et verticale, au pixel (x, y).

3- **Construction de l'histogramme :** La troisième étape est la création des histogrammes de l'orientation des gradients par des cellules carrées de petite taille de 4×4 à 12×12 pixels. Chaque pixel de la cellule vote alors pour une classe de l'histogramme, en fonction de l'orientation du gradient à ce point. Le vote du pixel est pondéré par l'intensité du gradient en ce point.

4- **Formation et normalisation des blocs :** afin d'éviter les disparités dues aux variations d'illumination, ainsi que l'introduction de redondance dans le descripteur. Les cellules sont regroupées par bloc (concaténation des histogrammes des cellules), le vecteur de valeur du bloc est ensuite normalisé. Le vecteur de HOG pour chaque cellule est normalisé en utilisant des vecteurs dans leurs blocs respectifs. En utilisant l'équation de normalisation (L2-norme) suivant :

$$\boldsymbol{V_n = \frac{V}{\sqrt{\|V\|_2^2 + e^2}}} \qquad (2.11)$$

Où : $\boldsymbol{V}$ est le vecteur contenant tous les histogrammes d'un bloc donné, $\boldsymbol{\|V\|_2}$est la norme $L2$ du vecteur $\boldsymbol{V}$ du descripteur et $\boldsymbol{e}$ est une petite constante.

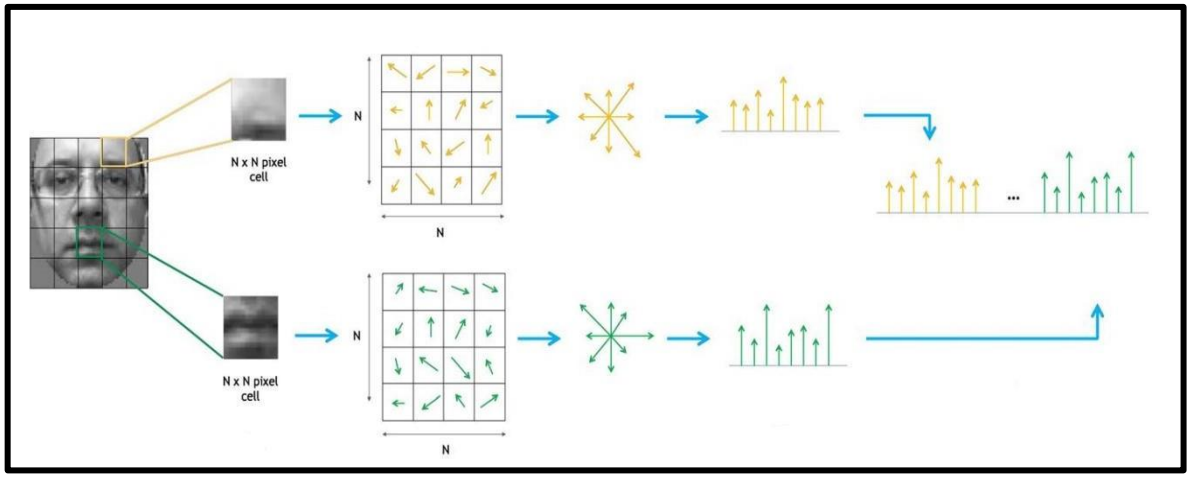


Figure 2.7 : Exemple d'extraction l'histogramme avec le descripteur **HOG** [42].

2.1.2. La classification

La classification automatique de l'image est le processus d'attribution automatique d'une classe à une image avec un système de classification. C'est ainsi que nous trouvons la classification des objets, des scènes, des textures, de la reconnaissance faciale.

La classification c'est l'étape finale dans le processus de vérification de lien de parenté, elle consiste à modéliser les paramètres extraits d'une modalité (le visage) des deux individus en se basant sur leurs caractéristiques communes. Un modèle est un ensemble d'informations utiles, discriminantes et non redondantes qui caractérise deux ou plusieurs individus ayant des similarités, ces derniers seront regroupés dans la même classe (classé les deux personnes en deux classes « classe parenté » ou « classe non-parenté »). Selon les caractéristiques extraites précédemment, les algorithmes de classification diffèrent, récemment beaucoup ont été proposées pour la classification dans la littérature, on trouve des techniques basées sur la mesure de la similarité ou du calcul de la distance tels que distance euclidienne, Minkowski et la métrique cosinus qui mesure la similarité de deux vecteurs de caractéristique.

Il y a d'autres méthodes se basent sur la classification des caractéristiques par un seul classificateur comme SVM « *Support Vector Machine* », KNN « *K-Nearest Neighbors* », toutes ces méthodes consistent à déterminer le modèle le plus proche des visages d'entrées parmi les images de la base de données.

a. Support Vector Machine (SVM)

La technique « *Support Vector Machine* (SVM) », également appelé Séparateur à Vaste Marge, développé par Vapnik en 1995 [43], est l'une des méthodes les plus populaires dans la famille des approches supervisé. La technique est destinée à résoudre des problèmes de classification ou de régression. Son but est de sépare linéairement avec un hyperplan optimal les données d'apprentissage (appartenant à des classes différentes) et maximiser la distance

entre les classes à partir d'un classificateur linéaire [44]. Cette technique est appliquée dans divers domaines, à savoir la reconnaissance des caractères, la reconnaissance d'images et le diagnostic médical, etc.

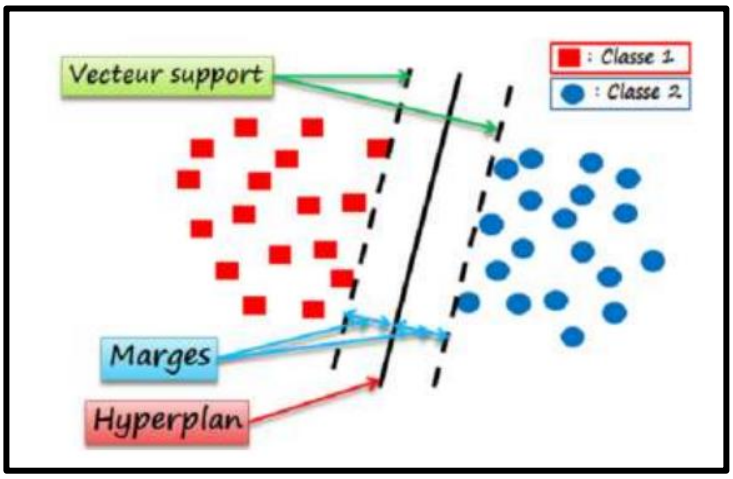


Figure 2.8 : Principe de la classification en utilisant la technique de SVM.

b. K-Nearest Neighbors (KNN)

L'algorithme « *K-Nearest Neighbors* (KNN) », également appelé K plus proche voisins (K-PPV) est un algorithme supervisé non paramétrique pour la classification [45]. Cet algorithme est largement utilisé pour plusieurs tâches d'intelligence en classification, segmentation d'images et vidéo. Le principe de l'algorithme est de regrouper les échantillons selon leur voisinage, c'est-à-dire que chaque point est affecté à la classe la plus représentée parmi ses k voisins les plus proches. Cet algorithme est basé sur les deux éléments principaux suivants :

- **K** : le nombre de cas les plus proches à utiliser.
- Une métrique pour mesurer le plus proche voisin.

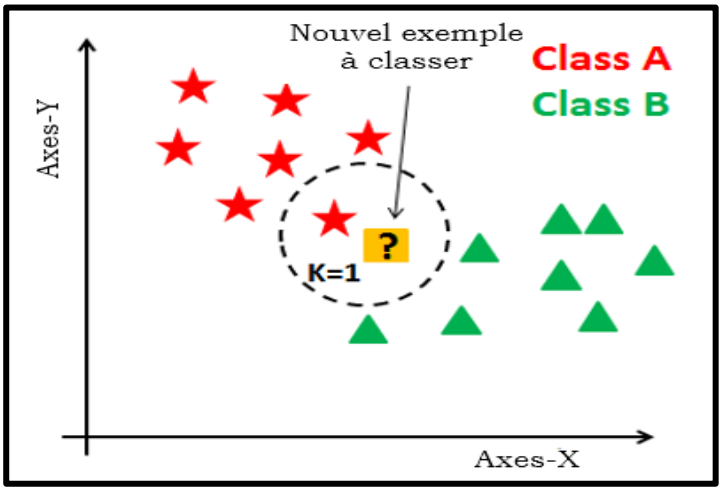


Figure 2.9 : La tâche de la classification en utilisant l'algorithme K-Nearest Neighbors (KNN). *Pour k = 1 la classe majoritaire du point central est la classe A* [46].

2.2. Les systèmes de reconnaissance basés sur l'apprentissage profond

L'apprentissage profond « Deep Learning » peut être considéré comme un sous-ensemble des algorithmes de Machine Learning (ML) et l'intelligence artificielle (IA). L'apprentissage profond est une technique d'apprentissage basée les réseaux de neurones artificiels ANN « *Artificiel Neural Networks* », ils sont conçus pour imiter la façon dont les humains pensent et apprennent. Ces réseaux de neurones artificiels, comprenant de nombreuses couches et stimulent l'apprentissage en profondeur pour aider les systèmes intelligents à comprendre les tâches de perception complexes avec une précision maximale. L'apprentissage profond définit ses propres paramètres de base, ce qui permet à la machine d'apprendre par elle-même sans le recours aux techniques classiques d'extraction des paramètres. Les techniques d'apprentissage profond sont utilisées dans la classification des images, la traduction des langues et la reconnaissance vocale...etc. Il peut être utilisé pour résoudre tout problème de reconnaissance de formes [47, 48].

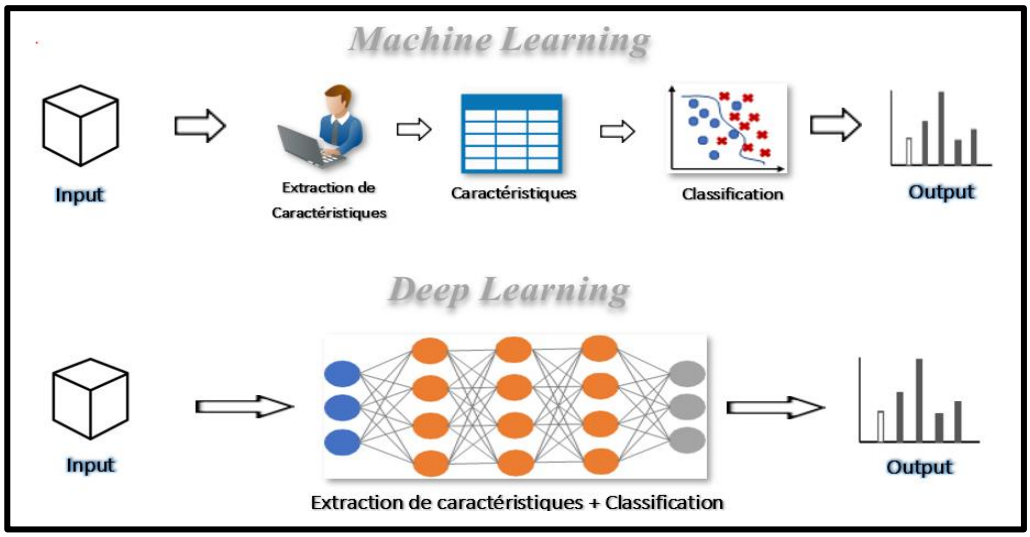


Figure 2.10 : Le procède du ML traditionnel comparé à celui du Deep Learning [49].

2.2.1. Fonctionnement de l'apprentissage profond

L'apprentissage profond se différencie de l'ensemble des algorithmes d'apprentissages de la Machine Learning par une architecture de réseaux de neurones artificiels multicouches complètement intégrés qui lui offre une meilleure autonomie dans son apprentissage. Car dans sa version « classique ». Les algorithmes de la Machine Learning nécessitent généralement des données préalablement traitées, pour son entraînement. Ces techniques peuvent donner des systèmes prédictifs, où leur exactitude dépend de la qualité des caractéristiques extraites. Les algorithmes d'apprentissage automatique « classique » ont besoin d'être guidés afin de pouvoir reconnaître les propriétés d'un élément donné. Au contraire, dans l'apprentissage profond, le

programme repère lui-même les caractéristiques nécessaires à son objectif. Ce qui le rend particulièrement pertinent pour l'apprentissage non supervisé et pour l'exploitation du grand volume de données [50].

Le principe d'apprentissage profond est sur l'entrainement sur un grand nombre d'exemples étiquetés afin de concevoir des modèles pour chaque classe dans l'espace d'étiquetage. Ainsi, après la mise à jour des différents paramètres de chaque couche neuronale, le système compare les résultats de prédiction obtenus pour un ensemble d'exemples de test par rapport aux étiquetages. Ainsi, avec un processus itératif, il ajuste les poids placés sur les différentes couches neuronales pour corriger son erreur en utilisant l'algorithme de rétropropagation « *back-propagation* ». Le processus est répété des milliers de fois jusqu'à l'atteint des performances optimales du système de reconnaissance automatique. Cette technique d'apprentissage est appelée « *supervised learning* » ou apprentissage supervisé. L'autre technique d'apprentissage non supervisé. Les réseaux de neurones doivent reconnaître des patterns au sein des ensembles de données pour apprendre par eux-mêmes quels éléments d'une photo peuvent être pertinents. La **figure 2.11** représente un exemple d'un système de reconnaissance automatique basée sur l'apprentissage profond avec un étiquetage trinaire : trois classes « Cat », « Dog » et « Raccoon ».

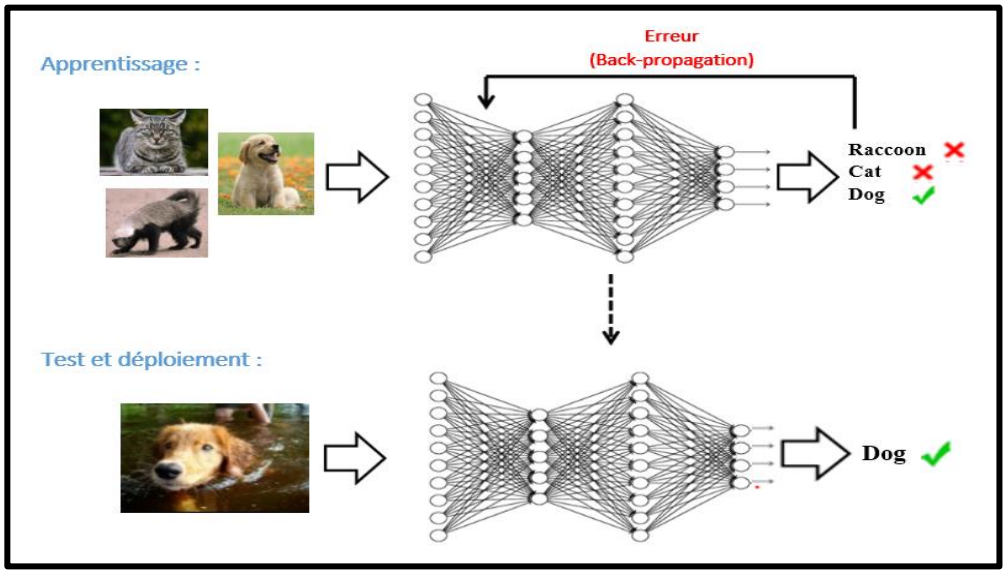


Figure 2.11 : Approche d'apprentissage en profondeur [51].

2.2.2. Les applications du l'apprentissage profond

L'apprentissage en profondeur ou plus communément appelé apprentissage hiérarchique est une discipline qui et évolue rapidement et prolifère et change notre façon de voir les technologies. Ainsi, il est encore plus difficile d'avoir une vision de ses futures applications. En particulier, Les principales avancées ont été dans l'application de l'apprentissage en profondeur à l'analyse

multimédia, y compris le traitement de l'audio et du langage naturel, traitement de l'image. Il a émergé dans de nombreux autres domaines, notamment les systèmes automatiques de diagnostics médicaux dans le domaine de la santé en général et dans celui de la bio-informatique pour l'étude de l'ADN, ainsi que dans le domaine de la sécurité informatique, etc. Dans la section suivante, nous citons certains domaines d'application des algorithmes d'apprentissage profond.

- Reconnaissance d'image.
- Traduction automatique.
- Traitement du langage naturel.
- La reconnaissance faciale.
- La reconnaissance vocale.
- La conduite autonome.
- Le diagnostic médical et les soins de santé (Health-care).
- La modération automatique des réseaux sociaux [47, 52, 53].

2.2.3. Les modèles de réseaux de neurones profonds (Architecteurs)

Il existe une variété d'architectures des réseaux de neurones basées apprentissage profond proposé dans les littératures. En effet, ces architectures dépendent étroitement par la complexité de la tâche et les caractéristiques des données d'entrées. Ces données d'entrées peuvent être des images isolées, des séquences d'images ou du vidéo, du texte et même des données dans un mode d'acquisition en temps réel. Pour cela, l'architecture de ces réseaux doit être adaptée selon les caractéristiques de données d'entrées. En général, nous trouvons trois grandes catégories d'architectures :

- ✓ Les réseaux neuronaux profonds **DNN** « Deep Neural Networks ».
- ✓ Les réseaux de neurones convolutifs **CNN** « *Convolutional Neural Network* ».
- ✓ Les réseaux de neurones récurrents **RNN** « *Recurrent Neural Network* ».
- ✓ Les réseaux de neurones récurrents à mémoire court-terme et long terme **LSTM** « *Long ShortTerm Memory* ».

Cependant, dans notre problématique les entrées d'un système de reconnaissance automatique des liens de parenté sont des paires d'images indépendantes (isolés). Ainsi, l'architecture d'apprentissage profond utilisée dans ces cas est les réseaux convolutif. Pour cela, nous intéressants dans les sections suivantes seulement sur l'architecture des réseaux de neurones convolutifs (CNN).

3. Les réseaux de neurones convolutifs (CNNs)

3.1. Principe

Les réseaux de neurones convolutifs (ou réseaux de neurones à convolutions, ou CNNs ou ConvNet) sont des types de réseaux de neurones artificiels acycliques « *feed-forward* ». Leur fonctionnement est inspiré par les processus biologiques, ils consistent en un empilage multicouche de perceptrons, dont le but est de prétraiter de petites quantités d'informations. Le nom « réseaux de neurones convolutifs » indique que le réseau emploie une opération linéaire mathématique entre matrices appelée convolution.

Les réseaux de neurones convolutifs sont actuellement parmi les modèles les plus performants pour la reconnaissance automatique des images. Ils comportent deux phases :

- La première est la phase de convolution, elle fonctionne comme un extracteur de caractéristiques des images à travers une succession de filtres, ou noyaux de convolution. Elle permet la création de nouvelles images intermédiaire appelée cartes de convolutions « *Feature maps* ». Au final, les cartes de convolutions sont alignées et concaténées dans des vecteurs de caractéristiques.
- La deuxième représente la phase de reconnaissance, elle est constituée de couches entièrement connectées « *Fully-connected* ». Le rôle de cette phase est la classification des entrées en se basant sur les résultats de la première phase. La sortie est généralement des vecteurs de probabilités selon l'étiquetage de données utilisées [54].

La **figure 2.12** présente une illustration du principe du réseau de neurones convolutifs. Les détails des éléments de chaque phase sont l'objectif des sections suivantes.

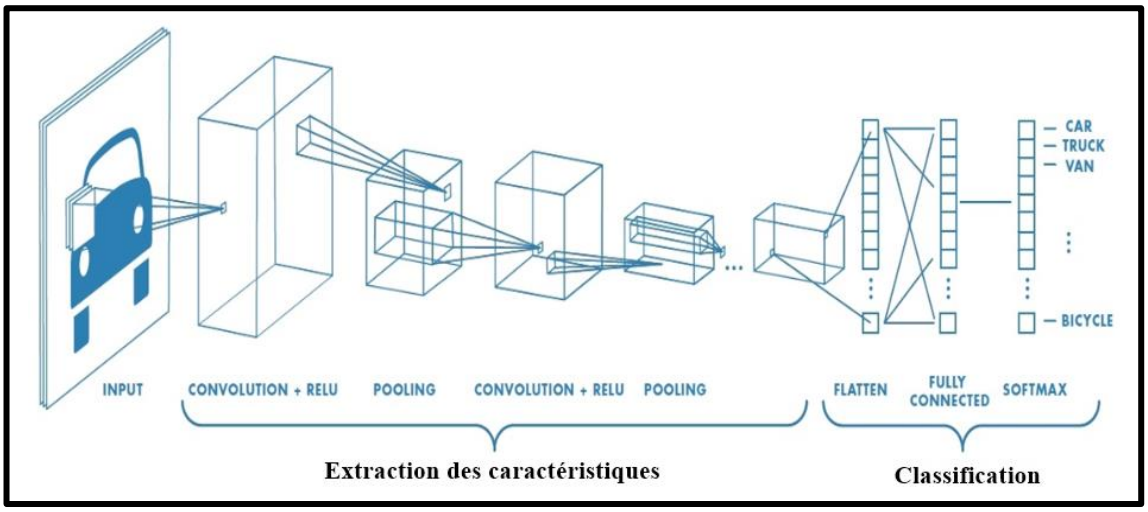


Figure 2.12 : Les deux parties qui composent l'architecture des réseaux neuronaux convolutifs [55].

3.2. La phase de convolution

La convolution est l'un des principaux éléments constitutifs d'un CNN. Elle fonctionne comme un processus de filtrage, c'est-à-dire appliquée un balayage sur les données d'entrée avec un filtre « *kernel* » de taille fixe. Pour chaque position du filtre, les valeurs des deux matrices en superposition (filtre et la partie de l'image à traiter) sont multipliées. Chaque valeur ainsi inférée est projetée dans une nouvelle matrice. Cette matrice représente une nouvelle image qui fait ressortir les caractéristiques recherchées au travers du filtre. La sortie de la convolution, que l'on peut appeler **«** carte de caractéristiques **»**, à des dimensions plus petites que celle de l'image en entrée [56]. La **figure 2.13** illustre le principe de convolution.

La phase de convolution pour l'architecture CNN est formée par un ensemble de couches de traitement en cascades. Dans la section suivante, nous présentons les éléments de cette phase :

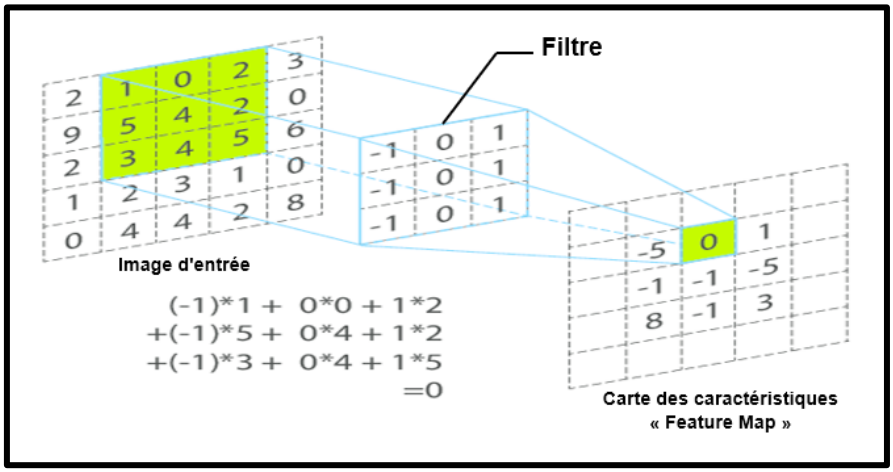


Figure 2.13 : Exemple de calcul de la convolution [57].

3.2.1. La couche de convolution (CONV)

La couche de convolution est la couche de construction de base d'un CNN, elle permet de créer des cartes de caractéristiques pour chaque paire (image, filtre). Elle permet de grader les informations discriminantes de chaque filtre appliqué à l'image. Trois paramètres essentiels pour la configuration de cette couche : la profondeur, le pas et la marge.

- **Profondeur de la couche (nombre de filtres) :** représente le nombre de filtres pour chaque couche de convolution. Elle représente la profondeur d'espace de recherche des caractéristiques discriminante pour les images d'entrées.
- **Le pas (stride) :** représente le pas de chevauchement des filtres dans l'image d'entrée. Il permet d'assurer la continuité des informations intra filtres pour une couverture globale des caractéristiques extraites.

- **La marge (zero-padding) :** permet de contrôler les dimensions de sortie des images filtrées pour chaque couche de convolution de sortie. En particulier, il est parfois souhaitable de garder la même dimension pour les images filtrées [58, 59].

3.2.2. La couche de pooling (POOL)

La couche de « Pooling » est souvent placée entre deux couches de convolution, elle reçoit en entrée plusieurs cartes de caractéristiques « Feature maps », et applique à chacune une opération de regroupement « pooling ». Il permet de réduire la taille de l'image filtrée tout en conservant les caractéristiques les plus importantes qu'elle contient, la **figure 2.14** illustre le principe de cette couche. Les deux fonctions les plus utilisées sont :

- **Max-pooling :** Elle garde la valeur la plus élevée de la zone définie.
- **Mean-pooling :** Elle garde la moyenne des valeurs de la zone définie.

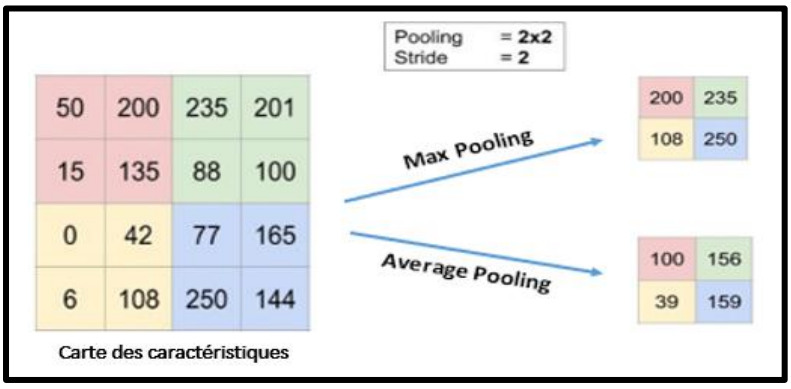


Figure 2.14 : Exemple avec types de pooling (avec un filtre "**2x2**" et un pas de "**2**") [60].

3.2.3. La couche d'activation

Comme toute architecture basée les réseaux de neurones artificiels, la fonction d'activation est le noyau de décision pour chaque couche de convolution et elle permet la correction des erreurs dans le processus de rétropropagation. En effet, il existe plusieurs fonctions d'activation, mais la plus utilisée dans la phase de convolution dans les architectures CNN est la fonction « Relu ».

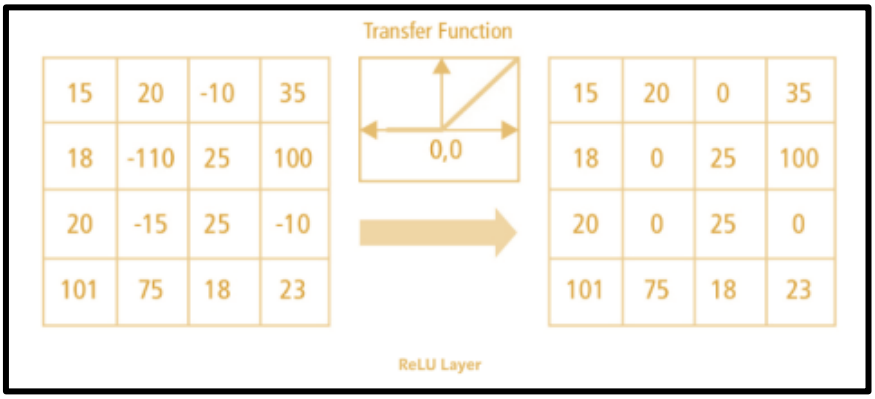


Figure 2.15 : La couche ReLU [61].

Cette fonction est définie par la fonction $f(x) = max\ (0, x)$. Appelée aussi « fonction d'activation non saturante », elle remplace toutes les valeurs négatives reçues en entrées par des zéros. Cependant, les autres fonctions utilisées sont :

- La fonction sigmoïde $f(x) = \ (1 + e\ - x)\ - 1$.
- La fonction tangente hyperbolique $f(x) = tanh(x)$.
- La fonction tangente hyperbolique saturante $f(x) = |tanh(x)|$ [58].

3.3. La phase d'évaluation ou reconnaissance

3.3.1. La couche de mise à plat « *flattening* »

Le « *flattening* » consiste à créer le vecteur de caractéristiques qui représente l'image d'entrée. C'est une étape de regroupement de toutes les images filtrées obtenues en deux dimensions à partir d'une image source « x » dans un espace d'une seule dimension. Le résultat de cette couche devient l'entrée du processus de reconnaissance ou de prédiction.

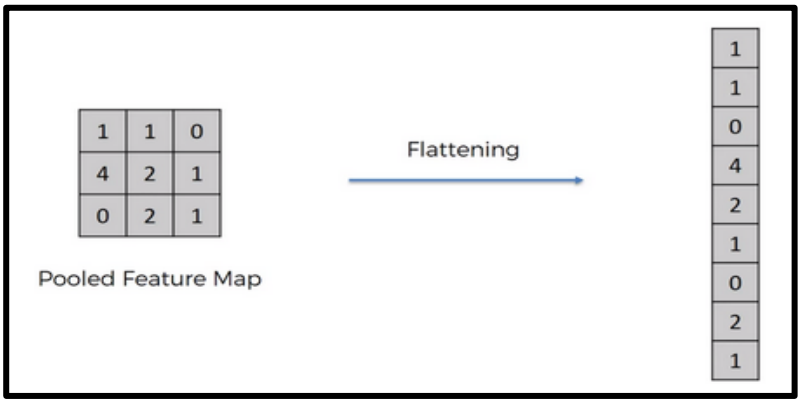


Figure 2.16 : Illustration du rôle de la couche flattening [61].

3.3.2. La couche dropout

La couche du dropout consiste à surmonter le problème de surapprentissage par la désactivation d'un ensemble aléatoire des connexions dans le réseau. Au moment de l'entraînement, l'opération consiste à l'abandon temporaire d'une unité choisie au hasard et de toutes ses connexions du réseau pendant l'entraînement. Ainsi, pour chaque nouvel élément d'entrée, le réseau sélectionne de manière aléatoire un sous-ensemble de neurones, formant une architecture de couche différente. Ces architectures utilisent des poids communs, mais étant donné que l'apprentissage ne dépend pas de neurones et de connexions spécifiques, la couche dropout peut aider à prévenir le surapprentissage « *Over fitting* » [62, 63].

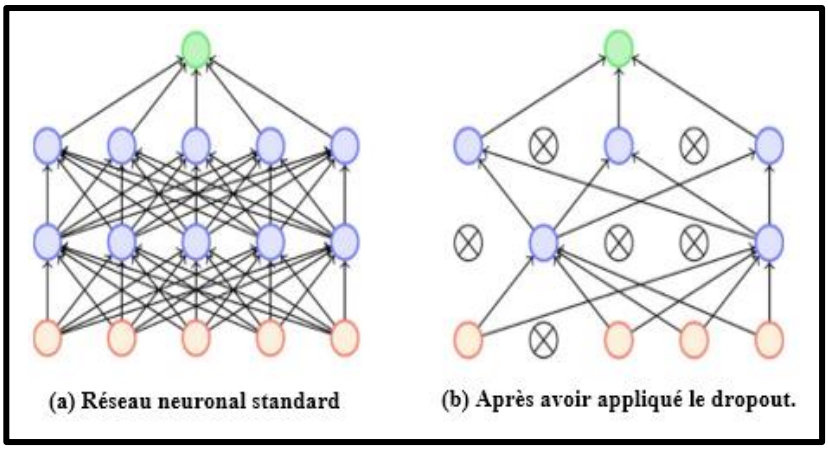


Figure 2.17 : Illustration de fonctionnement du dropout [64].

3.3.3. Couche entièrement connectée (Fully-Connected)

La couche entièrement connectée « dense » ou FC constitue toujours la dernière phase d'un CNN après plusieurs couches de convolution et de max-pooling. Les neurones d'une couche entièrement connectée ont des connexions à toutes les sorties de la couche précédente (comme on le voit régulièrement dans les réseaux de neurones réguliers).

Cette couche reçoit un vecteur plat en entrée résultant de l'étape de flattening et crée un nouveau vecteur en sortie. Pour ce faire, elle applique une combinaison linéaire puis éventuellement une fonction d'activation aux valeurs reçues en entrée. L'intérêt de la couche FC à classifier l'image lors de l'entrée du réseau, elle renvoie un vecteur de taille N, Où N est le nombre de classes. Chaque élément du vecteur indique la probabilité pour l'image en entrée d'appartenir à une classe [58].

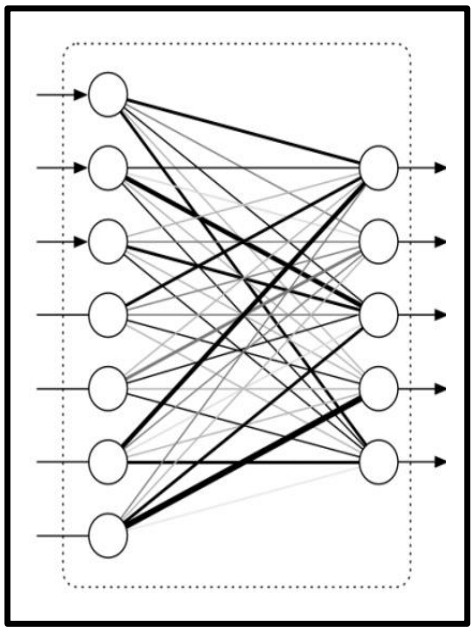


Figure 2.18 : Principe de la couche entièrement connectée (FC) [65].

3.4. Choix des hyperparamètres de CNNs

Malheureusement, tous les aspects définis dans l'architecture des CNN, il y a toujours une longue liste de paramètres qui doivent être définis manuellement pour permettre aux CNN d'avoir de meilleurs résultats. Les CNN utilisent plus d'hyperparamètres même si les règles habituelles pour les taux d'apprentissage et des constantes de régularisation s'appliquent toujours, il faut prendre en considération les notions suivantes :

- ✓ Pour chaque couche convolutive, combien de caractéristiques doit-on choisir ? (**Nombre de filtres**)
- ✓ Combien de pixels doit-on considérer dans chaque caractéristique ? (**Forme de filtres**)
- ✓ Pour chaque couche de pooling, quelle taille de fenêtre doit-on choisir ? Quel pas ?
- ✓ Pour chaque couche entièrement connectée supplémentaire, combien de neurones cachés doit-on définir ?

3.4.1. Nombre de filtres

Comme la taille des images intermédiaires diminue avec la profondeur du traitement, le nombre de filtres dans les couches proches de l'entrée est réduit, tandis que l'inverse se retrouve dans les couches plus proches de la sortie où elles sont plus grandes. Pour égaliser le calcul à chaque couche, le produit du nombre de caractéristiques et le nombre de pixels traités sont généralement choisis pour être à peu près constants à travers les couches. Pour préserver l'information en entrée, il est très important de conserver le nombre de sorties intermédiaires (c.-à-d. de multiplier le nombre d'images intermédiaires par le nombre de positions de pixels) afin d'augmenter (au sens large) d'une couche à l'autre. Le nombre d'images intermédiaires contrôle directement la puissance du système, dépend du nombre d'exemples disponibles et la complexité du traitement.

3.4.2. Forme de filtres

Les formes de filtre varient grandement dans la littérature. Ils sont souvent choisis en fonction de l'ensemble de données. Les meilleurs résultats sont ceux obtenus sur les images de MNIST (28x28) avec la taille de 5x5 sur la première couche, tandis que les ensembles de données d'images naturelles (souvent avec des centaines de pixels dans chaque dimension) ont tendance à utiliser de plus grands filtres de première couche de 12x12, et aussi 15x15 [66]. Donc, elle est relative à la tâche et la taille et le nombre des images en entrée et généralement elle est obtenue par les tests empiriques.

3.4.3. Forme de Max pooling

Les valeurs typiques sont 2 × 2. Des volumes d'entrée très importants peuvent justifier une mise en commun de 4 × 4 dans les couches inférieures (premières couches). Cependant, le choix de formes plus grandes réduira considérablement la dimension du signal et peut entraîner une perte excessive d'informations.

3.5. Quelques réseaux convolutifs célèbres (pré-entraînés)

Il est plus usuel d'utiliser des modèles de **CNNs** déjà entrainer et de les réadapter pour le problème de classification ou autre, c'est ce que l'on appelle l'apprentissage par transfert « *Transfer learning* ». Il s'agit de transférer la connaissance acquise et l'apprentissage d'un modèle traitant un problème de classification général vers un autre type de problème particulier. Il existe un grand nombre de réseaux **CNN** pré-entrainés, les plus connus sont :

3.5.1. LeNet

Les premières applications réussies des réseaux convolutifs ont été développées par *LeCun et al* dans les années 1990 [67]. Parmi ceux-ci, le plus connu est l'architecture « LeNet » utilisée pour lire les codes postaux, les chiffres, la reconnaissance de caractères manuscrits, etc.

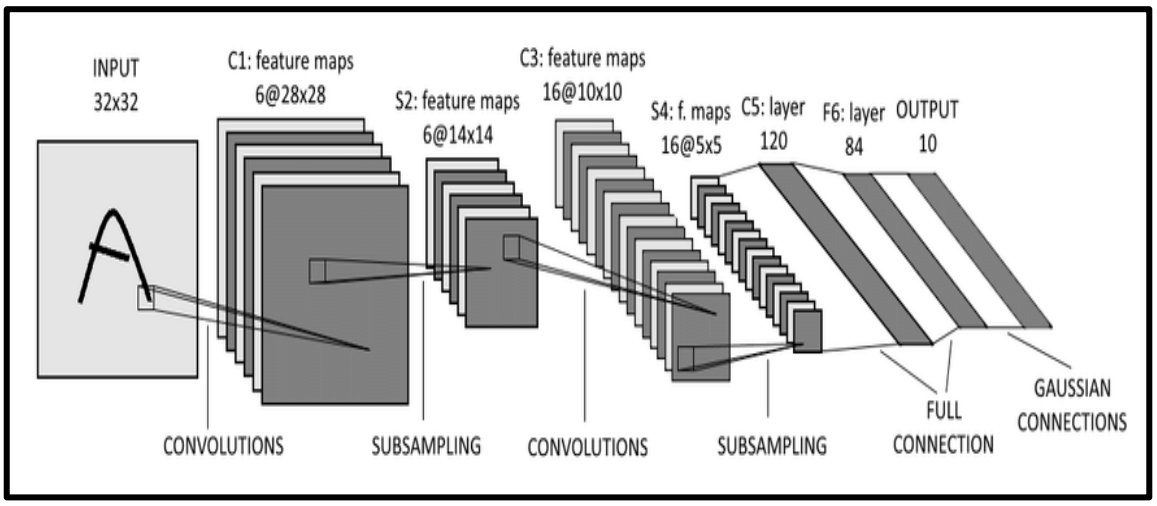


Figure 2.19 : Architecture de **LeNet** [68].

3.5.2. AlexNet

Le premier travail qui a popularisé les réseaux de neurones convolutifs dans la vision par ordinateur, développé par *Alex Krizhevsky et al* [69]. L'AlexNet a été soumis au challenge ImageNet ILSVRC en 2012 et a nettement surperformé tous ses concurrents précédents et a remporté le défi en réduisant l'erreur de top-5 de 26% à 15 ,3%. Le réseau avait une architecture très similaire à celle de LeNet, mais il était plus profond, plus grand et contenait huit couches ; les cinq premières étaient des couches de convolutions empilées les unes sur les autres et les trois dernières étaient des couches entièrement connectées et ils ont utilisé la fonction d'activation

non saturante de Relu, qui a montré une amélioration des performances d'entrainement. Il a été entraîné avec 60 millions de paramètres et 650 000 neurones pendant 5 à 6 jours sur deux GPUs GTX 580 de 3 Go.

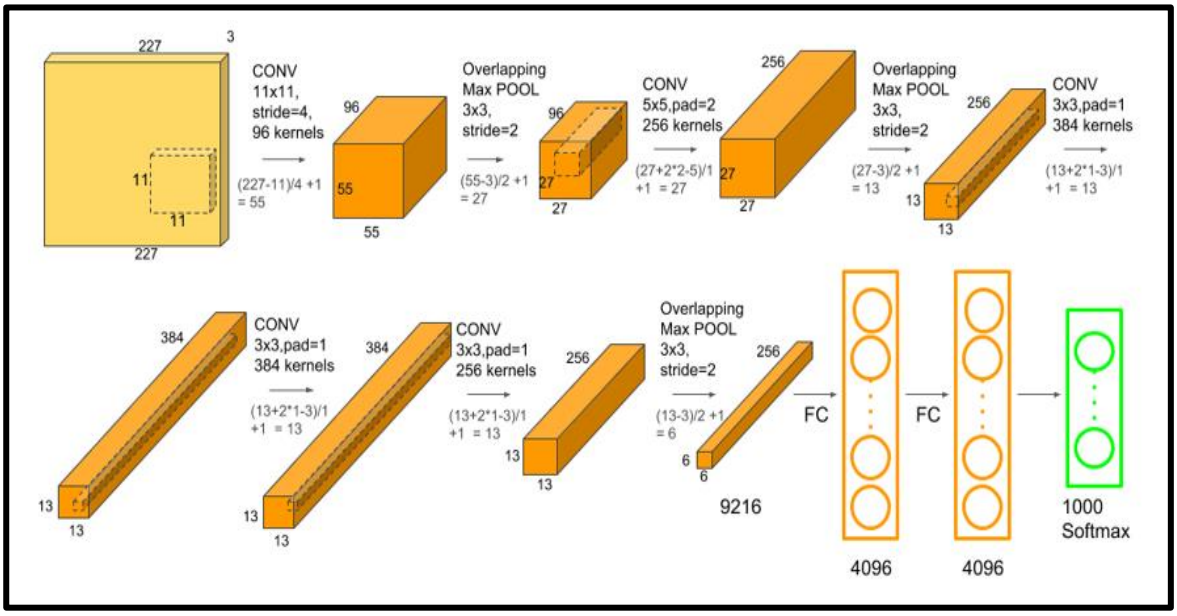


Figure 2.20 : Architecture de **AlexNet** [70].

3.5.3. VGGNet

C'est un réseau de neurones convolutif très performant dans le cadre du challenge de reconnaissance visuelle à grande échelle Image Net, où sa principale contribution a été de montrer que la profondeur du réseau est un composant essentiel pour de bonnes performances. VGGNet a été développé par le Visual Geometry Group de l'Université d'Oxford, d'où le nom VGG, en 2014 par *K. Simonyan et al* [71]. C'était l'un des modèles célèbres soumis à l'ILSVRC-2014 en atteignant le modèle à 92,7% de précision de test top-5 dans la base d'ImageNet. VGGNet, également connu sous le nom de VGG-16, se compose de 16 couches de poids ; 13 couches convolutives plus trois couches entièrement connectées et, de manière attrayante, présente une architecture extrêmement homogène, car il est très facile à comprendre et qui ne réalise que 3 x 3 convolutions et un regroupement 2x2 du début à la fin. VGG-16 a été formé pendant des semaines et utilisait des GPU NVIDIA Titan Black, et l'inconvénient de VGG est que ce réseau est généralement gros, plus coûteux à évaluer et utilise beaucoup plus de mémoire et de paramètres (environ 138 M).

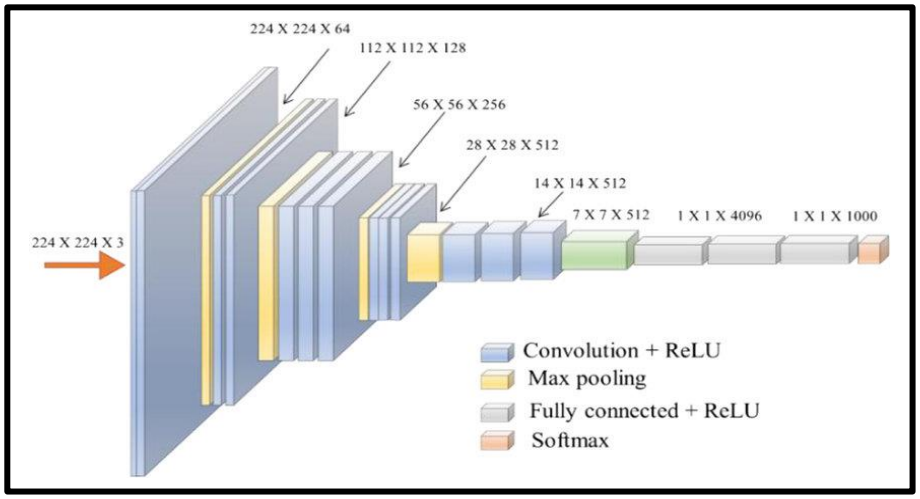


Figure 2.21 : Architecture de **VGGNet** [72].

3.5.4. ResNet

Le réseau *Residual Neural Network* (ResNet) développé par *Kaiming He et al* [73]. Il a été le gagnant de la 1ère place à la compétition de classification ILSVRC 2015. ResNet a introduit une architecture innovante avec des « sauts de connexions » et une lourde normalisation par lots. Ces connexions sont également appelées unités « *Gated* » ou unités « *Recurrentes gated* » et ont une forte similitude avec les éléments récents réussis appliqués dans les RNNs. L'intérêt de derrière la création ce type de réseaux de neurones convolutifs est de résoudre le problème de descente de gradients « *Vanishing* » dans les réseaux de neurones convolutifs très profond. Grâce à cette technique, ils ont pu former un réseau de neurones avec 152 couches et environ 60 millions paramètres, tout en ayant une complexité inférieure à celle de VGGNet. Il atteint un taux d'erreur de 3,57% parmi les cinq premiers (top-5), ce qui est supérieur à la performance au niveau humain sur le corpus utilisé.

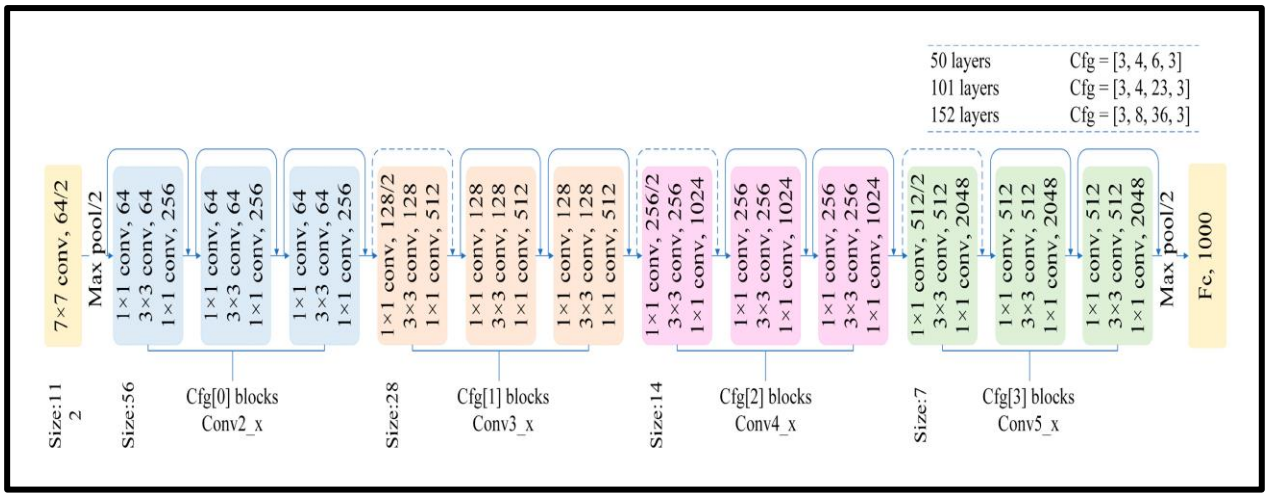


Figure 2.22 : Architecture de **ResNet** [74].

4. Mesures d'évaluation de performance des systèmes

Afin de mesurer les performances des systèmes de reconnaissance automatique, nous trouvons dans les littératures plusieurs métriques et courbes qui permettent l'évaluation de ces systèmes. Dans ce contexte, nous citons dans les sections suivantes les métriques utilisées dans les systèmes de vérification automatiques des liens de parenté.

4.1. Matrice de confusion

Cette matrice est définie dans le domaine de recherche d'information « *Information Retrieval* », elle permet d'évaluer des résultats de prédictions d'un classifieur par la comparaison des valeurs réelles d'étiquetage Y des données d'entrées avec les valeurs prédites Ŷ fournies par le modèle de classification. La matrice de confusion est une matrice dont les colonnes sont les classes obtenues et les lignes sont les classes désirées (Valeurs réelles) pour tout l'échantillon de test utilisé. Ainsi, la diagonale contient l'évaluation des échantillons bien classée (prédictions correctes), hors diagonale les éléments mal classés (prédictions incorrectes) comme le présente le Tableau 3.2 [75].

Classe réelle	**Classe prédite**	
	C1	C2
C1	***VP***	***FN***
C2	***FP***	***VN***

Tableau 2.1 : Matrice de confusion dans le cas d'une classification binaire.

Avec :

- **VP** sont les vrais positifs, c'est-à-dire les observations qui ont été classées positives et qui le sont réellement.
- **FP** sont les faux positifs, c'est-à-dire les individus classés positifs et qui sont en réalité des négatifs.
- De la même manière, les **FN** sont les faux négatifs et **VN** sont les vrais négatifs.

4.2. Courbe ROC (*Receiver Operating Characteristic*)

La courbe ROC est un graphe qui représente les performances d'un modèle de classification pour tous les seuils de classification. Cette courbe trace le taux des vrais positifs en fonction de celui des faux positifs. La courbe ROC montre ces taux pour différents seuils et plus cette courbe est éloignée de la première bissectrice, plus le modèle de classification à une forte capacité à discriminer. L'indicateur synthétique associé à la courbe ROC est la surface située sous la

courbe, c'est l'AUC « *Aera Under the Curve* ». Un modèle de classification est performant si l'AUC est proche de 1. À l'inverse, un modèle de classification n'est pas discriminant si l'AUC est proche de 0,5.

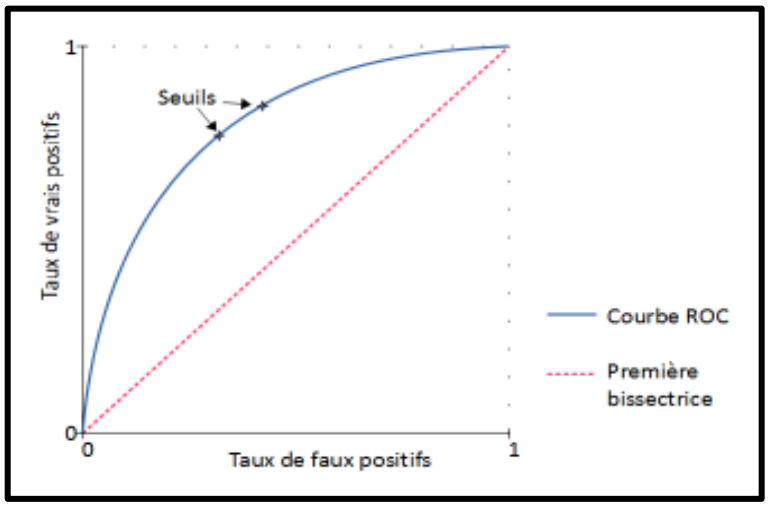


Figure 2.23 : Courbe ROC (Receiver Operating Characteristic) [76].

4.3. Les métriques d'évaluation

Il existe plusieurs métriques numériques d'évaluation utilisées dans les systèmes de reconnaissance automatiques, parmi eux nous citons :

- **Accuracy :** désigne la proportion des prédictions correctes effectuées par le modèle de classification.

$$\mathbf{Accuracy} = \frac{\mathbf{VP} + \mathbf{VN}}{\mathbf{VP} + \mathbf{VN} + \mathbf{FP} + \mathbf{FN}}$$

- **Précision :** de ceux que l'algorithme a pu classer, combien sont corrects.

$$\mathbf{Précision} = \frac{\mathbf{VP}}{\mathbf{VP} + \mathbf{FP}}$$

- **Rappel :** de ceux qui existent, combien l'algorithme a pu trouver.

$$\mathbf{Rappel} = \frac{\mathbf{VP}}{\mathbf{VP} + \mathbf{FN}}$$

- **F-mesure :** moyenne harmonique entre la précision et le rappel.

$$\mathbf{F-mesure} = \frac{2 * (\mathbf{Précision} * \mathbf{Rappel})}{(\mathbf{Précision} + \mathbf{Rappel})}$$

5. Travaux connexes pour la vérification automatique des liens de parenté

Les systèmes de reconnaissance des liens de parenté fonctionnent sur deux modes : la vérification et l'identification. Cependant, il existe également des systèmes basés uniquement sur la vérification de parenté, c'est-à-dire sans définir la classe de la relation de parenté entre les deux images qui sont traitées. En effet, la majorité des travaux proposés dans les littératures qui seront cités dans les sections suivantes traitent le problème de vérification de l'existence des liens de parenté.

Entre autres, les architectures d'apprentissage automatique et profond ont révélé leurs efficacités dans de nombreux problèmes de reconnaissance automatique complexes avec l'utilisation des algorithmes intelligents et de réseaux de neurones artificiels pour extraire et apprendre des informations discriminantes. Par conséquent, divers travaux et contributions anciennes et récentes ont été proposés comme une solution à ce sujet de la vérification de la parenté à partir d'images faciales. Ces travaux peuvent être classés en trois approches principales : basés sur les caractéristiques, basés sur l'apprentissage profond et basés sur l'apprentissage des métriques. Dans ce qui suit, nous en avons mentionné quelques travaux récents sur chaque approche.

5.1. Les systèmes basés sur les caractéristiques locales

Lu et al [77]**,** ont proposés une nouvelle méthode d'apprentissage de caractéristiques discriminantes (PDFL) basée sur un prototype pour la vérification de parenté, vise à apprendre des caractéristiques discriminantes de niveau intermédiaire pour mieux caractériser la relation des images de visage pour la vérification de parenté, contrairement à des méthodes qui utilisent des descripteurs de caractéristiques artisanaux de bas niveau. Dans ce travail, ils ont appliqué trois descripteurs de caractéristiques différents, pour extraire des informations différentes de chaque image faciale, avec les deux bases « KinFaceW-I » et « KinFaceW-II ». Ils ont enregistré les résultats suivants : les caractéristiques du descripteur SPLE ont donné de bien meilleurs résultats que l'autre descripteur (LBP, SIFT) avec une précision de 65,6% pour la base « KinFaceW-I ». Tandis que, pour la base « KinFaceW-II », ils ont obtenu une précision de 76.4% comme le résultat le plus élevé ainsi avec les caractéristiques SPLE et l'autres caractéristiques SIFT, LBP ils ont enregistré les taux de reconnaissance 64% et 69.9% respectivement. Ainsi, afin d'utiliser plusieurs caractéristiques de bas niveau pour l'apprentissage des caractéristiques de niveau intermédiaire, ils ont proposé en outre une méthode PDFL multi-vue « MPDFL » afin d'améliorer les performances de vérification. Ils ont enregistré meilleur score 70.1% pour la base « KinFaceW-I » et 77% avec la base « KinFaceW-II ».

Liu et al [78]**,** ont proposé une méthode innovante de vecteur de caractéristique de Fisher héréditaire « IFVF » pour la vérification de parenté. Cette approche proposée vise à découvrir les relations génétiques entre les enfants et les parents afin de conserver à la fois la capacité discriminante et la capacité héréditaire. Plus précisément, le vecteur de Fisher est dérivé pour chaque image en agrégeant les caractéristiques de descripteur SIFT échantillonnées de manière dense dans l'espace de couleur de l'adversaire, puis une transformation héréditaire est apprise, qui maximise la similitude entre les images de parenté tout en minimisant celle entre les images de non-parenté pour chaque paire d'images simultanément. Enfin, ils ont utilisé une mesure de similitude cosinus pour le calcul de la similarité entre les vecteurs. La méthode proposée est évaluée sur les deux bases « KinFaceW-I » et « KinFaceW-II » où les résultats expérimentaux ont montré les taux de vérification de 73.45% et 81.60% respectivement.

Duan et al [79]**,** ont proposé une méthode de soustraction de caractéristiques pour supprimer la partie sans lien de parenté de la caractéristique locale pour la vérification de la parenté. Ceci est fait en utilisant une méthode de descente de gradient pour réduire simultanément la distance caractéristique entre les paires d'images de visage avec la parenté et maximiser la distance entre les paires de non-parenté. Sur la base de la caractéristique soustraite, la vérification est réalisée au moyen de comparaison de distance « Gaussian ». Les tests expérimentaux ont montré parmi les trois descripteurs : LPQ, SIFT et LBP utilisées dans l'étude, que le descripteur LPQ a donné les meilleurs résultats avec des taux de vérification de 70.9% et 77,1% sur les bases de données « KinFaceW-I » et » KinFaceW-II » respectivement.

Xu et al [80]**,** ont suggéré un modèle de similarité robuste pour apprendre une mesure de similarité linéaire clairsemée selon lequel le score de similarité de chaque paire d'images de visage positive est supérieur au score de similarité moyen de certaines paires d'images de visage négatives. Ils ont aussi développé un algorithme d'apprentissage de la similitude en ligne « OSL-A » pour une application plus évolutive. Ils ont évalué empiriquement la méthode proposée sur les bases « KinFaceW-I » et « KinFaceW-II » où ils ont également appliqué deux descripteurs de visage HOG et SIFT, pour extraire les traits du visage, les résultats étaient comme suit : Avec la base « KinFaceW-I », avec les caractéristiques du descripteur HOG le score est 77.92% ainsi avec les caractéristiques SIFT le score est 76.35%. Pour la base « KinFaceW-II », ils ont enregistré les scores 77.10% pour les caractéristiques HOG et de 74.90% pour les caractéristiques SIFT.

Alirezazadeh et al [81], ont suggéré une méthode basée sur la détection des parties principales du visage telle que les yeux, le nez et la bouche avec l'image complète du visage, comme une première étape. Ces parties seront utilisées pour l'extraction des caractéristiques :

LBP, SPLE et LPQ. Ensuite, pour réduire la redondance des vecteurs de caractéristiques et améliorer le taux de vérification, la méthode de sélection de caractéristiques nommer KinFS « *Kinship Feature Selection* » est proposée pour la sélection des caractéristiques discriminantes. Cette méthode est basée sur l'algorithme RSFS « *Random Subset Feature Selection* ». Pour la classification, ils ont utilisé le classifieur SVM, pour déterminer la relation de parenté ou de non-parenté. Les résultats obtenus ont enregistré les taux de 78.9% et 84,5% sur les ensembles de données « KinFaceW-I » et « KinFaceW-II » respectivement.

5.2. Les systèmes à base d'apprentissage profond

Zhang et al [18], proposent d'extraire les caractéristiques de haut niveau pour la vérification de parenté basée sur des réseaux de neurones convolutifs profonds (**CNN**). Compte tenu de l'importance des points clés faciaux, ils ont également extrait des caractéristiques basées sur les points clés pour la vérification de la parenté. Avec les bases « KinFaceW-I » et « KinFaceW-II », les résultats sont 74.8% et 77.5% pour la première et 85.3%, 88.4% pour la deuxième base. Ainsi, ils ont montré que leur approche proposée a apporté des améliorations significatives par rapport aux méthodes précédentes.

Lu et al [82]**,** ont présenté une méthode discriminante d'apprentissage métrique profond « **DDML** » pour la vérification du visage et de la parenté, l'objectif de cette méthode est de capturer la variété non linéaire. Ils ont utilisé un réseau de neurones profond afin d'apprendre un ensemble de transformations non linéaires hiérarchiques. Pour chaque image de visage de la base de données « KinFaceW-I »et « KinFaceW-II », ils ont extrait quatre descripteurs de caractéristiques. Avec la base « KinFaceW-I », le résultat le plus élevé était obtenu avec les caractéristiques LPQ avec une précision de 81.4% et les autres caractéristiques : DSIFT, HOG et LBP ont donné moins avec 78.4%, 77.5% et 75.5% respectivement. Concernant la base « KinFaceW-II », qui a atteint presque les mêmes résultats en obtenant également la plus grande précision de 82.2% avec les caractéristiques LPQ, puis suivie par les caractéristiques DSIFT avec 79.1% et HOG, LBP avec 78.1% et 77.6% respectivement. Ils ont aussi développé une méthode d'apprentissage multidimensionnel profond discriminant (DDMML) pour apprendre conjointement plusieurs réseaux de neurones et qui a donné comme la meilleure précision de 83.5% avec « KinFaceW-I » et 84.3% avec « KinFaceW-II ».

5.3. Les systèmes basés sur l'apprentissage des métriques

Yan et al [83], ont proposé une méthode d'apprentissage métrique repoussé du voisinage (NRML) pour la vérification de parenté, la mesure proposée vise à apprendre une mesure de distance sous laquelle les échantillons intra-classes (avec une relation de parenté) sont

rapprochés le plus possible et pousser des échantillons inter-classes (sans relation de parenté) dans un voisinage, simultanément, de sorte que des informations plus discriminantes peuvent être exploitées pour la vérification. Les résultats des tests empiriques sur les deux bases « KinFaceW-I » et « KinFaceW-II » obtenus avec le classificateur SVM portaient à quatre caractéristiques : avec la première base, les caractéristiques SPLE ont donné une précision de 64.3%, SIFT, LBP et TPLBP ont donné de moins de résultats avec une précision de 63.8%, 63.7% et 62.9% respectivement. Cependant, avec la seconde base la plus grande précision est de 75.7% avec les caractéristiques SPLE, et les caractéristiques TPLBP, LBP et SIFT ont donné une précision de 70.1%, 69.5% et 62.8% respectivement. Leur objectif était aussi la meilleure utilisation de plusieurs descripteurs de caractéristiques afin d'extraire des informations complémentaires, et pour cela ils ont proposé une autre méthode NRML multi-vue « MNRML » pour rechercher une métrique de distance commune pour effectuer une fusion des caractéristiques multiples afin d'améliorer les performances de vérification de parenté, où il a donné de bons résultats avec une précision de 76.5% pour la base « KinFaceW-II » et de 69.9% pour la base « KinFaceW-I ».

Yan et al dans [84], ont proposé est une méthode d'apprentissage multi métrique discriminante « DMML » pour la vérification de parenté via l'analyse d'images faciales afin d'apprendre plusieurs métriques de distance avec des multiples caractéristiques extraites par des différents descripteurs de caractéristiques : LBP, SPLE et SIFT. Elle prouve son efficacité par rapport aux méthodes d'apprentissage mono-métriques et multi-métriques existantes. La méthode proposée avec le classifieur SVM était plus efficace qu'avec d'autres classifieur comme KNN. Pour les résultats obtenus, ils ont enregistré des taux de vérification de 72.25%, 70.62% et 69.5% respectivement pour la base « KinFaceW-I » ont donné. Tandis que pour la base « KinFaceW-II », ils ont enregistré les taux 78.25%, 76.88% et 74,75% respectivement.

Kou et al [85]**,** ont proposé une nouvelle approche de vérification de la parenté en apprenant une mesure de similitude clairsemée d'une manière en ligne, qui se concentre sur l'apprentissage d'une mesure de similitude génétique pour augmenter l'évolutivité pour les applications pratiques avec une quantité sans cesse croissante de données. Ils ont mené des expériences pour vérifier l'efficacité de leur approche les bases de données de parenté disponibles « KinFaceW-I » et « KinFaceW-II ». Pour les représentations caractéristiques de l'image du visage, ils ont utilisé trois descripteurs : SIFT, LBP, et SPLE et ils ont donné les résultats suivants : Avec la base « KinFaceW-I », le résultat le plus élevé était obtenu avec le descripteur SPLE avec une précision de 65.7% et les autres descripteurs : SIFT et LBP ont donné moins avec 65.5% et 63.5% respectivement. Concernant l'autre base « KinFaceW-II », qui a atteint presque les mêmes

résultats en obtenant également la plus grande précision de 74.8% avec le descripteur SPLE, puis suivie par les descripteurs LBP et SIFT avec 68.9% et 66.7% respectivement.

Zhou et al [86]**,** ont proposé une technique d'apprentissage de similarité d'ensemble « ESL », elle est fondée sur une fonction de similarité bilinéaire clairsemée pour modéliser les caractéristiques relatives aux images faciales avec des liens de parenté. La proposition exploite l'ensemble des paires d'images qui ont des liens de parenté en générant un ensemble de modèles de similitude afin d'accroitre la capacité de généralisation du système. Pour tester leur méthode, ils ont utilisé les bases « KinFaceW-I » et « KinFaceW-II ». Ils ont choisi deux descripteurs de caractéristiques : LBP et HOG pour extraire les caractéristiques discriminantes. Le taux de vérification atteint par les caractéristiques HOG a dépassé le taux de LBP avec 78.6% et 74.1% respectivement pour la base « KinFaceW-I » et les taux de 75.7% et 74.3% pour la base « KinFaceW-II ».

Hu et al, dans [87], ont présenté une méthode d'apprentissage multi-métrique à grande marge « LM^3L » pour la vérification du visage et de la parenté, qui apprend conjointement plusieurs métriques de distance globale en vertu desquelles les corrélations des différentes représentations de caractéristiques de chaque échantillon sont maximisées. Les expériences sur la base « KinFaceW-II » et avec les deux descripteurs : LBP, SIFT ont enregistré un taux de vérification de 78.7%. Pour mieux exploiter les structures locales des images de visages, ils ont proposé également une métrique locale d'apprentissage « LML » et une méthode d'apprentissage multi-métrique à grande marge locale « L^2M^3L » pour apprendre un ensemble de métriques locales, où il atteint un taux de vérification de 80.0%.

Jiwen Lu et al [88], ont proposé une méthode d'apprentissage métrique de parenté « KML » avec un modèle couplé de réseau neuronal profond « DNN ». KML apprend une métrique de similitude profonde couplée de sorte que les paires d'images avec la relation de parenté sont rapprochées, tandis que celles sans relation de parenté sont poussées aussi loin que possible, et les résultats montrent la supériorité de ce modèle. Ils ont extrait LBP, HOG, SIFT, VGG comme caractéristiques, des deux bases « KinFaceW-I » et « KinFaceW-II ». Ils ont obtenu les résultats suivants : Les caractéristiques VGG ont donné un résultat beaucoup mieux que d'autres caractéristiques : HOG, SIFT et LBP avec un taux de vérification 82.8% contrairement au 78.9%, 78.7% et 76.6% respectivement, et c'est ce qu'elle a donné à partir résultats de la base « KinFaceW-I ». Alors que la base « KinFaceW-II » a donné, comme résultats et comparaisons entre eux, avec les caractéristiques VGG, ont a également donné le meilleur résultat avec un taux de vérification de 85.7%, par rapport à d'autres caractéristiques : SIFT, LBP et HOG qui a donné moins avec un taux de 81%, 80.6% et 80.3%, respectivement.

Classe	Méthode	Caractéristiques	Classifieur	Base de données	Taux de verification (%)				
					F-S	F-D	M-S	M-D	Moyennes
Caractéristiques	PDFL & MPDFL [77]	LBP+SPLE+ SIFT	SVM	KinFaceW-I	68.2	63.5	61.3	**69.5**	65.6
					73.5	67.5	66.1	73.1	**70.1**
				KinFaceW-II	77.0	74.3	77.0	**77.2**	76.4
					77.3	74.7	77.8	**78.0**	**77.0**
	IFVF [78]	SIFT	Cos-sim	KinFaceW-I	73.3	71.7	71.1	**77.5**	73.4
				KinFaceW-II	**85.6**	75.4	82.8	82.6	**81.6**
	A feature subtraction method [79]	LBP+LPQ+ SIFT	Distance euclidienne (Gaussian)	KinFaceW-I	**75.4**	63.8	69.9	74.6	70.9
				KinFaceW-II	**82.4**	76.2	76.6	73.2	**77.1**
	OSL-A [80]	HOG+SIFT	SVM	KinFaceW-I	**82.7**	75.0	71.8	82.3	**77.9**
				KinFaceW-II	**82.0**	77.2	75.6	73.6	77.1
	Feature in High-dimensional (KinFS) [81]	LPQ+SPLE+ LBP	SVM	KinFaceW-I	77.8	73.5	79.6	**84.6**	**78.9**
				KinFaceW-II	**87.5**	79.8	83.9	86.7	**84.5**
Apprentissage profond	CNN [18]	CNN	CNN	KinFaceW-I	76.1	71.8	78.0	**84.1**	**77.5**
				KinFaceW-II	89.4	81.9	89.9	**92.4**	**88.4**
	DDML & DDMML [82]	LPQ+DSIFT+ LBP+HOG	SVM	KinFaceW-I	83.8	77.0	78.1	**86.6**	81.4
					86.4	79.1	81.4	**87.0**	**83.5**
				KinFaceW-II	**84.8**	82.6	79.4	81.8	82.2
					87.4	83.8	83.2	83.0	**84.3**
Apprentissage métrique	NRML & MNRML [83]	LBP+SPLE+SIFT +TPLBP	SVM	KinFaceW-I	64.1	59.1	63.9	**71.0**	64.3
					72.5	66.5	66.2	72.0	**69.9**
				KinFaceW-II	76.8	73.1	76.8	**77.0**	75.7
					76.9	74.3	77.4	**77.6**	**76.5**
	DMML [84]	LBP+SPLE+ SIFT	SVM	KinFaceW-I	74.5	69.5	69.5	**75.5**	**72.2**
				KinFaceW-II	78.5	76.5	76.5	**79.5**	**78.2**
	Learning a Genetic Measure [85]	SPLE+SIFT+ LBP	SVM	KinFaceW-I	66.1	62.2	64.3	**70.0**	65.7
				KinFaceW-II	74.9	71.0	**76.9**	76.4	**74.8**
	ESL [86]	HOG+LBP	KNN	KinFaceW-I	**83.9**	76.0	73.5	81.5	**78.6**
				KinFaceW-II	**81.2**	73.0	75.6	73.0	**75.7**
	LMMML & LLMMML [87]	LBP+SIFT	SVM	KinFaceW-II	**82.4**	74.2	79.6	78.7	78.7
					82.4	78.2	78.8	80.4	**80.0**
	KML [88]	LBP+HOG+ SIFT+VGG	KNN	KinFaceW-I	83.8	81.0	81.2	**85.0**	**82.8**
				KinFaceW-II	**87.4**	83.6	86.2	85.6	**85.7**

Tableau 2.2 : Etude comparatifs entre les performances des travaux susmentionnés avec les bases **KinFaceW-I** et **KinFaceW-II**.

6. Conclusion

Dans ce chapitre, nous avons présenté les techniques d'apprentissage utilisées dans les systèmes de reconnaissance automatique pour les images. En effet, la vérification automatique des liens de parenté est une application de ces systèmes. Nous avons présenté les deux grandes catégories de ces systèmes : les systèmes qui utilisent l'architecture dite classique en deux phases : extraction de caractéristiques et reconnaissance ou classification. Ainsi que les systèmes basés sur l'apprentissage profond. Ensuite, nous avons effectué une étude bibliographique sur les travaux effectués dans cette problématique avec un tableau de synthèse.

Cependant, les résultats obtenus dans ces travaux montrent que l'utilisation des architectures CNN est prometteuse tout en tenant compte des contraintes de calculs et espace mémoire. Car la particularité des systèmes de vérification automatique des liens de parenté est dans leurs données d'entrées avec une paire d'images ce qui rend la tâche plus complexe. En revanche, les plateformes de développement en cloud comme *Google Colab* ou *Kaggle* nous encouragent d'utiliser ces architectures dans notre problématique. Dans le chapitre suivant, nous présentons notre contribution pour un système de vérification automatique des liens de parenté.

Chapitre III

1. Introduction

Les systèmes de vérification de parenté sont considérés comme un système de reconnaissance de formes qui comprend plusieurs étapes à savoir le prétraitement des images, extraction de caractéristiques et la classification. L'un des problèmes les plus critiques de la reconnaissance de formes est la sélection de caractéristiques appropriées pour une représentation compacte et optimale de l'image. Pour cela, de nombreuses approches pour l'étape d'extraction de caractéristiques ont été proposées dans la littérature. En général, ces techniques peuvent être divisées en deux catégories différentes : techniques classiques (par exemple, **LBP**,**HOG**,**DCT**) et les techniques d'apprentissage en profondeur (par exemple, **CNN**, **PCANet** …).

Dans ce travail, nous avons proposé deux schémas de vérification de parenté. Le premier schéma comprend quatre méthodes de vérification de parenté basées sur les techniques classiques d'extraction de caractéristiques : **LBP**, **HOG** et deux variantes la technique **TPLBP** et **FPLBP**. Alors, le deuxième schéma de vérification de parenté est basé sur l'apprentissage profond (**CNN**) appliqué sur l'image globale et des régions locales.

L'objectif principal de cette étude est d'obtenir les vecteurs de caractéristiques qui peuvent représenter efficacement les caractéristiques discriminantes. Les deux bases de données KinFaceW-I et KinFaceW-II sont utilisées pour évaluer les performances des méthodes proposées. Enfin, nous avons analysé les résultats et les avons comparés à de multiples recherches de vérification de parenté.

2. Méthodologie et méthodes proposées

Dans cette section, nous présentons les différentes méthodes de vérification de parenté qui sont regroupées dans de deux classes de méthodes à savoir les méthodes basées sur les schémas classiques d'extraction de caractéristiques et les méthodes basées sur l'apprentissage en profondeur.

2.1. Vérification de parenté basée sur les schémas classiques d'extraction de caractéristiques

Le schéma général de système de vérification de parenté basé sur les schémas classiques d'extraction de caractéristiques est présenté dans la **figure 3.1**. La première étape est l'étape d'extraction de caractéristiques ou nous avons utilisé les techniques : **LBP**, **HOG**, **TPLBP** et **FPLBP**. La deuxième étape est l'application de l'Analyse des Composants Principales (**ACP**)

pour la sélection de caractéristiques. Enfin, une étape de classification, utilisant le classifieur KPPV basée sur la métrique de similarité ***Cosinus***, est utilisée pour tester les performances de différents schémas d'extraction de caractéristiques.

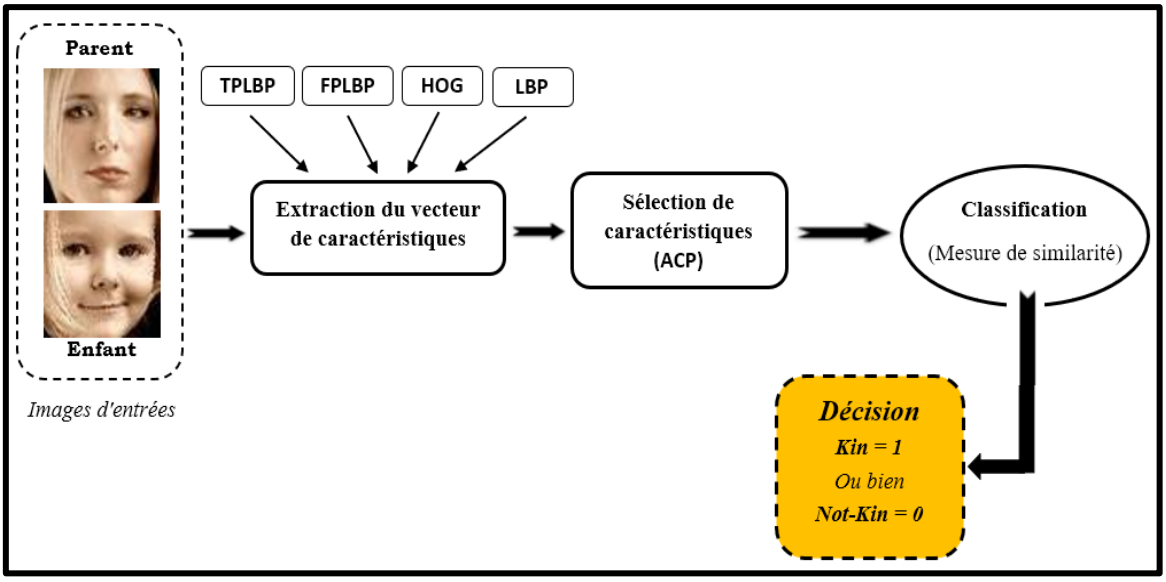


Figure 3.1 : Schéma général de vérification de parenté basé sur les méthodes classiques d'extraction de caractéristiques.

Dans ce qui suit, nous allons détailler chacune de ces étapes.

2.1.1. Extraction du vecteur de caractéristiques

L'extraction de caractéristiques est une étape cruciale dans toute application de reconnaissance de formes, car de bons résultats de classification sont directement liés à l'unicité et à la variabilité des caractéristiques extraites utilisées pour distinguer les différents modèles. Dans nos méthodes proposées, les techniques **LBP**, **HOG**, **TPLBP** et **FPLBP** sont utilisées. Notons que la description de ces méthodes est présentée dans le chapitre 2.

2.1.2. Sélection de caractéristiques

La représentation brute des informations de l'images de visage conduit à des vecteurs de caractéristiques de grandes dimensions, ce qui peut poser des problèmes de complexité de calcul et de capacité de stockage. Ainsi, il est parfois préférable de chercher à réduire la dimensionnalité et de sélectionner les caractéristiques discriminatoires et non redondantes d'un problème de classification pour améliorer ses performances. Plusieurs solutions sont préconisées dans la littérature pour sélectionner et réduire la taille de sa dimension [89]. Dans ce travail, nous avons utilisé l'analyse en composantes principales (**ACP**).

❖ *Analyse en Composantes Principales (ACP)*

L'Analyse en Composantes Principales (**ACP**) ou Principal Component Analysis (**PCA**) en anglais, est une méthode mathématique de réduction de dimensionnalité qui est souvent utilisée pour simplifier et réduire la dimensionnalité de grands ensembles de données et pour représenter efficacement les données (par exemple, images de visages).

L'idée principale de **PCA** consiste à mettre l'accent sur la variation et mettre en évidence les motifs forts dans le jeu de données, ceci en trouvant une nouvelle base de l'espace des données plus petit que l'ensemble original des données, qui néanmoins contient la plupart de l'information de l'échantillon sans perte, dans lequel tous les vecteurs sont indépendants les uns des autres (dites "non corrélées") et orthogonaux entre eux. Ces nouveaux vecteurs, appelés composantes principales ou axes, sont ordonnés dans le sens décroissant par fraction de l'information totale que chacun contient. Le premier de ces vecteurs correspond à la direction de variance maximale des données d'apprentissage. Les autres composantes sont déterminées par la contrainte d'orthogonalité entre les vecteurs tout en respectant une direction de variance maximum. Ainsi, au lieu de stocker une vectrice image de taille N, on stocke un vecteur de taille R ≤ nombre d'images d'apprentissage. Les étapes principales pour l'algorithme de **PCA** sont :

- Chaque image est représentée par un vecteur de taille N.

$$x^i = \left[x_1^i,\ x_2^i,\ x_3^i, \dots, x_N^i\right]^T \qquad (3.1)$$

- Toutes les images sont centrées en calculant la moyenne de toutes les images d'apprentissage et soustrayant l'image moyenne de chaque vecteur.

$$M = \frac{1}{p}\sum_{i=1}^{p} x^i \ \ , \bar{x}^i = x^i - M \qquad (3.2)$$

  Où M : moyenne de toutes les images d'apprentissage.

  Ces vecteurs sont combinés, côte-à-côte, pour créer une matrice $\boldsymbol{X}$ de données d'apprentissage de taille N × P (où P est le nombre d'images d'apprentissage, N la taille de la vectrice image).

- Le calcul de la matrice de covariance (C) qui est définie par la formule :

$$\boldsymbol{C} = \bar{X}\bar{X}^T \qquad (3.3)$$

- On calcule ensuite les vecteurs propres V_i $(i = 1 \dots N)$ etleurs valeurs propres λ_i de la matrice de covariance **C**, où ***i*** = *1…N.*

  Les valeurs propres de la matrice **C** représentent le taux de variation le long de l'axe du vecteur propre associé.

- Ne conservez que les premiers vecteurs propres ordonnés par ordre décroissant des valeurs propres correspondantes (les composants principaux), $k(k \ll N)X^k = V^k.X$

☞ Résoudre pour **PCA**

$$\lambda V_{X^T} = C_x V_{X^T} \quad (3.4)$$

La **PCA** est une technique rapide, simple et populaire dans l'identification de modèle, c'est l'une des meilleures techniques. Les projections de la **PCA** sont optimales pour la reconstruction d'une base de dimension réduite et son fonctionnement peut être considéré comme révélant la structure interne des données de manière à mieux expliquer la variance dans les données [90].

2.1.3. La classification

L'étape de classification est considérée comme on le sait, c'est la phase où la décision est prise, c'est ce qui la rend très importante dans les systèmes biométriques notamment en reconnaissance faciale. Lorsque notre objectif de classification est de comparer deux vecteurs de caractéristiques obtenus à partir du module d'extraction de caractéristiques d'une paire d'images de visage, nous pouvons effectuer deux mesures principales, soit la mesure de similarité (ressemblance), soit la distance (divergence), telle que Euclidienne, Hamming, Minkowski, Cosinus et autres. Dans nos tests, en raison de la spécificité des descripteurs de caractéristiques basés sur des histogrammes, nous sommes intéressés par une technique de mesure de similarité plus spécialisée. La mesure de similitude la plus fréquente pour l'appariement des histogrammes est la mesure du similarité cosinus [91]. Elle sert à mesurer la similarité entre deux vecteurs de N dimensions en trouvant l'angle entre eux.

La mesure de similarité cosinus entre deux vecteurs $\mathbf{V_1}$ et $\mathbf{V_2}$ est défini comme :

$$Cos_{sim}(V_1, V_2) = \frac{V_1 . V_2}{\|V_1\| . \|V_2\|} = \frac{\sum_{i=0}^{n} x_i \, y_i}{\sqrt{\sum_{i=0}^{n} (x_i)^2} . \sqrt{\sum_{i=0}^{n} (y_i)^2}} \quad (3.5)$$

Où ‖. ‖ est la norme euclidienne. Une valeur élevée du score produit entre les deux vecteurs signifie une forte probabilité que la paire de visages appartienne à la même famille, cela signifie l'existence d'une relation de parenté.

2.2. Vérification de parenté basée sur l'apprentissage en profondeur

Les algorithmes classiques sur le sujet de la vérification de la parenté plus précisément l'extraction des caractéristiques à partir d'images faciales représentent généralement des visages avec plusieurs caractéristiques de bas niveau, suivis d'un modèle d'apprentissage superficiel. Cependant, ces caractéristiques manuelles générales ne peuvent pas bien découvrir les informations impliquées dans les images faciales pour la vérification de la parenté, et donc même les meilleurs résultats de ces algorithmes ne sont pas satisfaisants.

Dans la deuxième partie de nos travaux, nous avons proposé deux variantes de méthodes d'extraction des caractéristiques de haut niveau pour la vérification de parenté basée sur les réseaux neuronaux convolutionnells profonds. La première variante est basée sur l'extraction de caractéristiques des images globales de visage. Tandis, dans le deuxième schéma proposé, compte tenu de l'importance de points-clés (land-marks) faciaux sur la vérification de parenté, nous avons modifié la structure de **CNN** basique en une nouvelle structure nommé **KP-CNN** (Key Points CNN) basé sur les points clés de différentes régions du visage.

2.2.1. Vérification de parenté basée sur le CNN basique

La structure de base de CNN (**CNN-Basic**) utilisée dans ce travail comprend deux couches convolutives, suivies d'une couche entièrement connectée et d'une couche **sigmoïde**. Comme le montre la **figure 3.2**, l'entrée est une paire d'images **64 x 64** avec trois canaux (RVB). Suite à l'entrée, la première couche convolutive est générée après convolution de l'entrée via $\mathbf{N_1}$ filtres (kernels) avec un pas de **1** (Stride =**1**). Chaque filtre est de taille $\mathbf{(k_1 \times k_1 \times 6)}$. La deuxième couche convolutive filtre l'entrée de la couche précédente avec $\mathbf{N_2}$ filtres de taille $\mathbf{(k_2 \times k_2 \times N_1)}$. Après les couches convolutionnelles, une couche entièrement connectée (fully-connected) projette les caractéristiques extraites dans un sous-espace avec **F** unités (neurones). Les couches de **Max-pooling** qui sont caractérisées par un pooling de taille (**2** × **2**) et un pas de **1**, suivent les première et deuxième couches convolutionnelles pour réduire la dimension des caractéristiques résultantes (Features Map) et éviter le overfitting. Nous adoptons la fonction **ReLU** comme fonction d'activation des couches de convolution. Enfin, ce réseau est formé via un classifieur **sigmoïde** bidirectionnel (**2 neurones**) à la dernière couche pour vérifier la parenté de deux personnes.

L'architecture du **CNN** que nous avons proposée dans notre travail est la même pour les différentes relations de parenté (**F-S**, **F-D**, **M-S**, et **M-D**). Afin de choisir les meilleurs paramètres du **CNN** (nombre de filtres, taille de filtre et nombre d'unités dans la couche entièrement connectée), nous avons réalisé plusieurs tests avec différentes taille et nombres des filtres et nombres unités ($\mathbf{k_1}$, $\mathbf{k_2}$, $\mathbf{N_1}$, $\mathbf{N_2}$ et **F**).

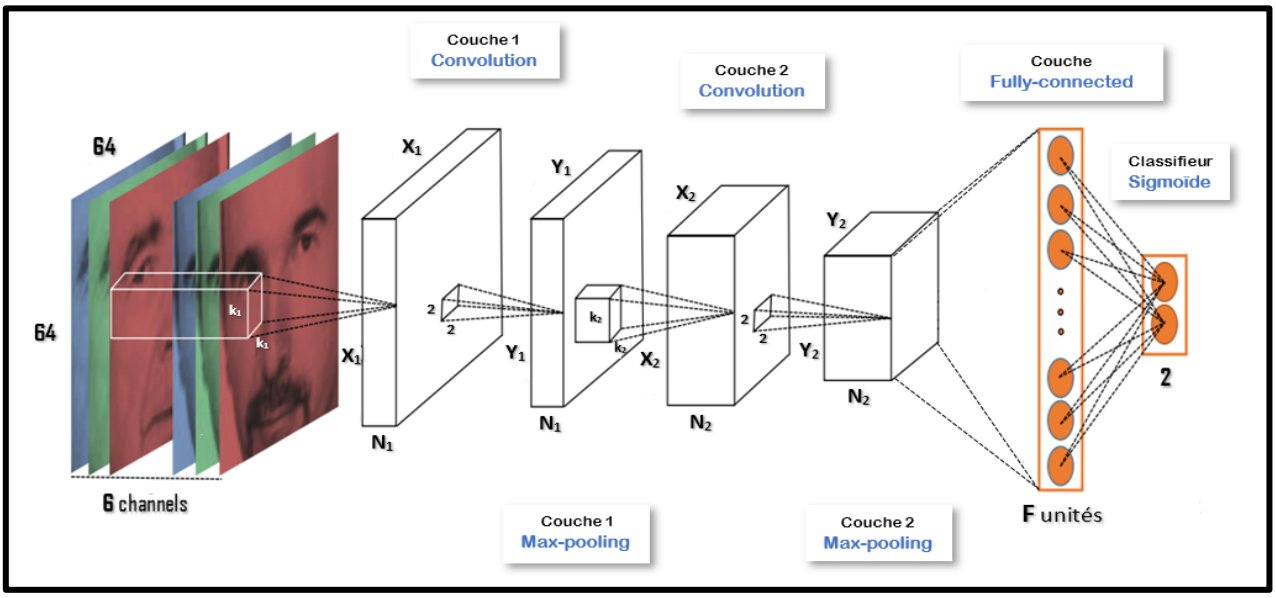


Figure 3.2 : Architecture proposée du **CNN** de base pour la vérification de parenté.

2.2.2. Vérification de parenté basée sur les points clés du visage (KP-CNN)

Dans la deuxième variante de vérification de parenté, nous basons sur des régions locales du visage tel que les yeux, le nez et la bouche. Nous avons considéré que des parties du visage ont un impact significatif sur l'analyse de parenté, et conçoivent ainsi une représentation des caractéristiques pour la vérification de la parenté.

Le principe de base de notre proposition est basé sur l'utilisation d'un algorithme de détection des points faciales nommé « **68 Landmarks predictor** » [92]. Cet algorithme est utilisé pour capturer les coordonnées de différentes régions comme les coins des yeux, bouche et nez, joues, menton, haut du nez et du front… (**Figure 3.3**).



Figure 3.3 : Les 68 points clés du visage.

Dans notre travail, nous avons utilisé uniquement les coordonnées correspondantes aux coins de chaque œil, les coins de la bouche et du nez afin de localiser ces trois régions dans le visage. Une fois ces régions sont définies, chaque région est rognée sur la zone couverte par les points

clés, la taille des images des trois régions a été redimensionnée pour avoir la même résolution comme suit : la région des yeux avec les dimensions (**60** × **25**), le nez avec (**25** × **35**) et la bouche avec (**40** × **30**), comme illustré à la **figure 3.4**.

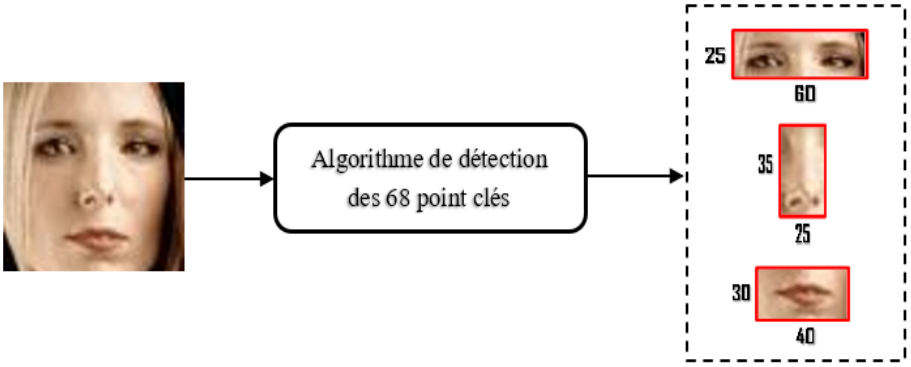


Figure 3.4 : Exemple des régions utilisées.

Afin d'améliorer la vérification de parenté avec ces régions du visage, nous proposons une nouvelle structure (**KP-CNN**) qui est illustré dans la **Figure 3.5**. La nouvelle structure contient 3 modèles **CNN** (**Eyes_CNN**, **Nose_CNN** et **Mouth_CNN**), dont chacun reçoit une paire de régions du visage, l'objectif de ces trois modelés est seulement d'extraire les caractéristiques de ces régions.

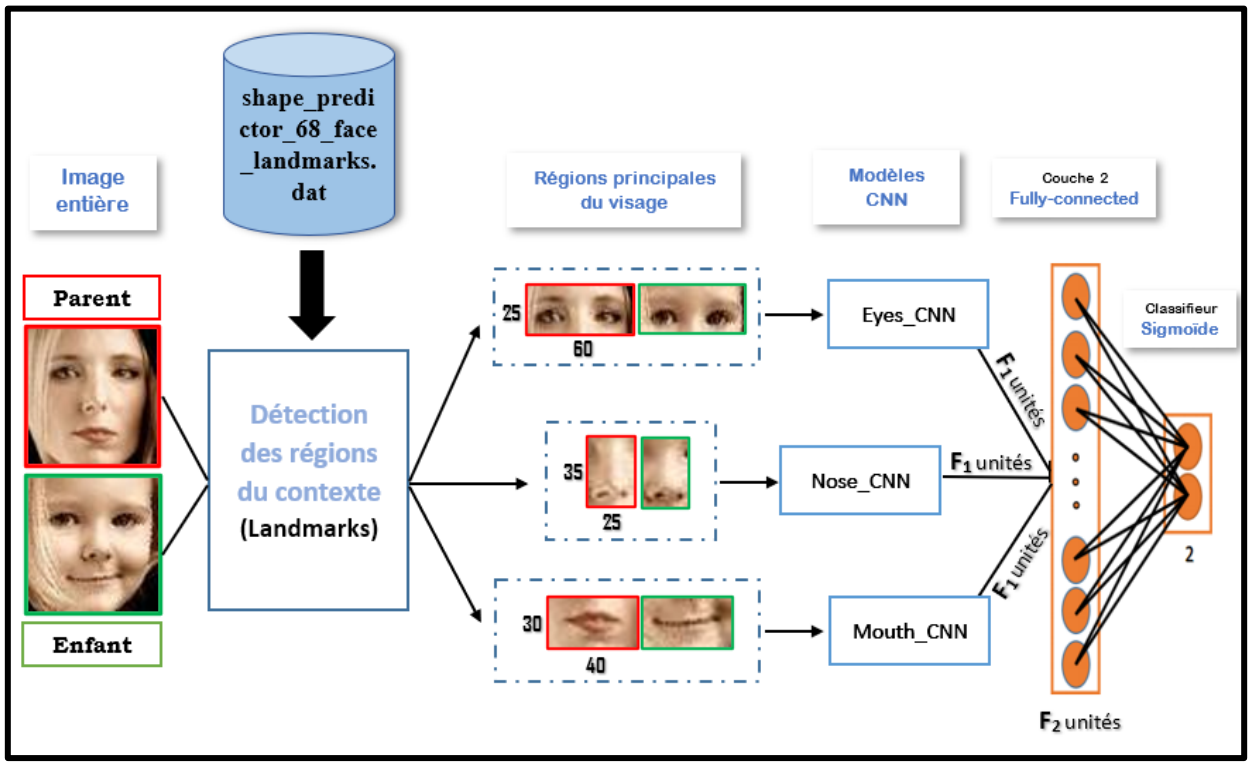


Figure 3.5 : Architecture proposée du **KP-CNN** pour la vérification de parenté.

A) Modèle proposé pour la région des yeux (Eyes_CNN)

L'architecture proposée dans ce modèle, illustrée dans la **figure 3.6**, l'entrée du modèle sera les deux régions des yeux du père et de l'enfant générés par l'algorithme de détection des points faciales (**landmarks**), ce modèle est composé de 6 couches réparties : 3 couches convolutions, 2 couches pooling et une couche entièrement connectée.

- La première couche convolutionnelle est destiné à l'extraction d'identité (feature extraction).il effectue une convolution avec $\mathbf{N_1}$ filtres de taille $\mathbf{k_1 \times k_1}$ pour filtrer les deux régions d'entrée de taille **60 × 25 × 6**. Cette couche est effectuée avec une stride de (**1,1**) avec un padding "same", la fonction d'activation utilisée est ReLU.
- La première couche de maxpooling occupe la deuxième position dans le modèle, elle se caractérise par un pooling de taille **2 × 2** et un stride (**1,1**).
- La deuxième couche convolutionnelle occupe la troisième position dans le modèle. Cette couche prend en entrée la sortie de la première couche convolutionnelle après une couche pooling avec $\mathbf{N_2}$ filtres de taille $\mathbf{k_2 \times k_2}$; et un pas (stride) de (**1,1**).
- La deuxième couche de pooling vient après la deuxième couche de convolution et elle dispose de mêmes paramètres que la couche de pooling précédente.
- La troisième couche occupe la cinquième position du modèle avec N_3 filtres de taille $\mathbf{k_3 \times k_3}$ avec un stride (**1,1**) en utilisant un padding avec l'application d'une fonction d'activation ReLU.
- Notre modèle se termine par une couche entièrement connectée (FC) avec F_1 unités, nous avons utilisé dans cette couche une fonction d'activation ReLU suivi par une couche dropout pour réduire le problème de overfitting.

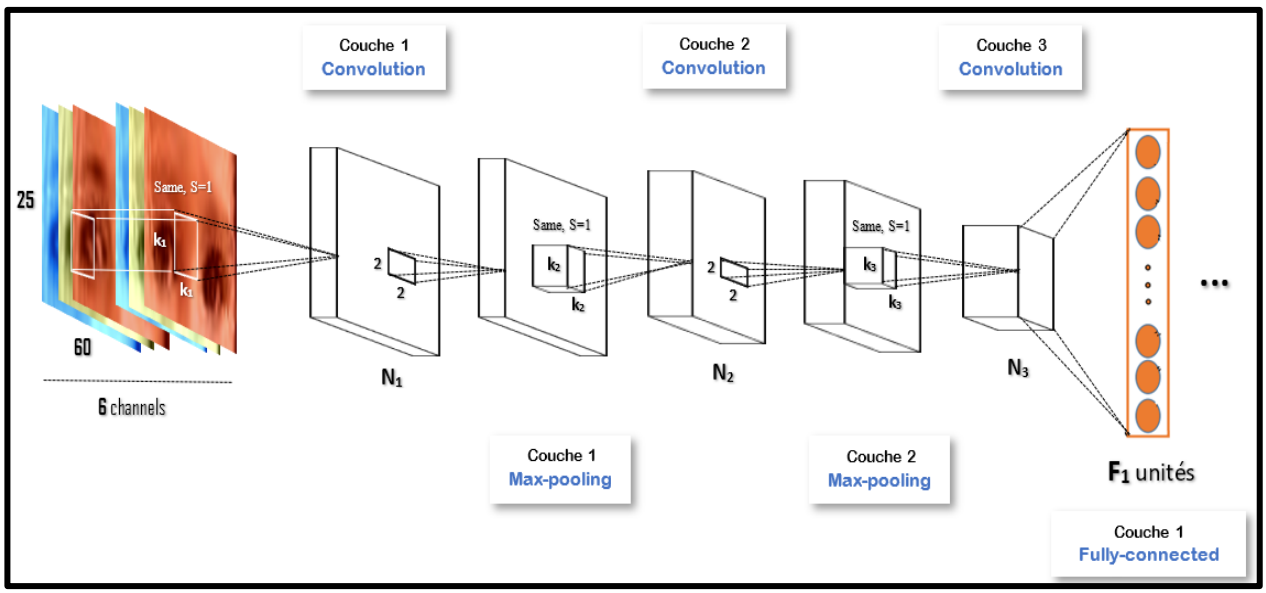


Figure 3.6 : Architecture proposée du modèle **Eyes_CNN**.

B) Modèle proposé pour la région de nez (Nose_CNN)

L'architecture proposée dans ce modèle, illustrée dans la **figure 3.7**, l'entrée du modèle sera les deux régions des nez du père et de l'enfant générés par l'algorithme de détection (**landmarks**), ce modèle est composé de 5 couches réparties : 2 couches convolutions, 2 couches pooling et une couche entièrement connectée. La seule différence entre cette architecture et la précédente est que dans ce modèle nous n'avons pas de troisième couche de convolution.

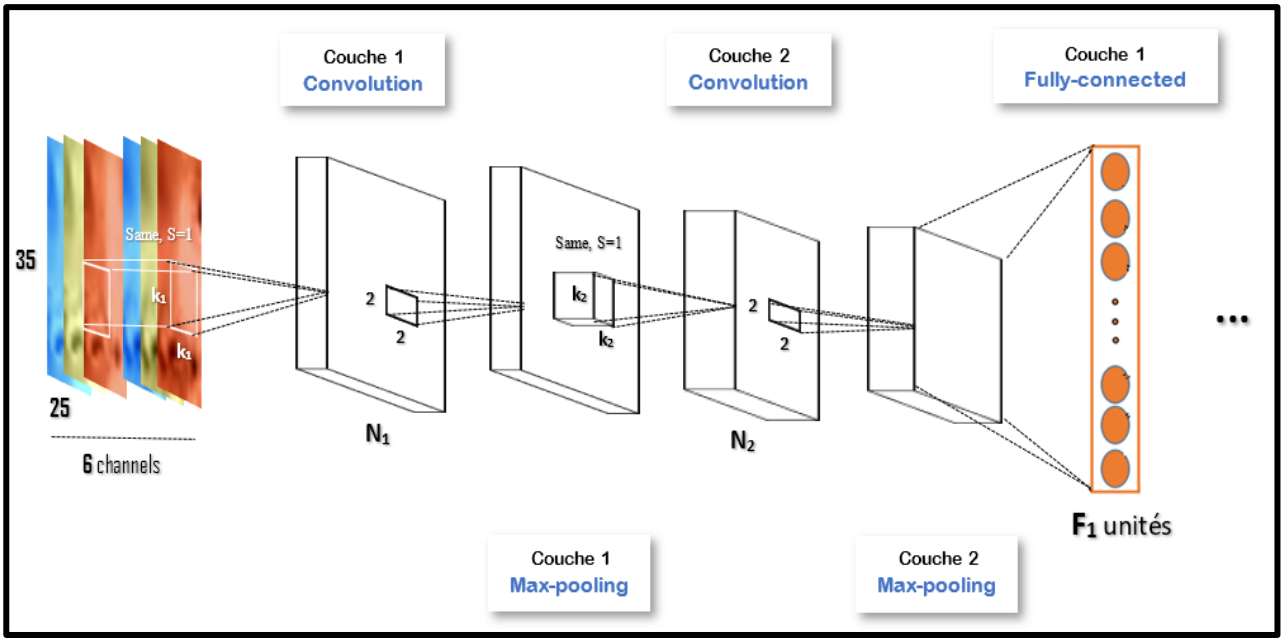


Figure 3.7 : Architecture proposée du modèle **Nose_CNN**.

C) Modèle proposé pour la région de bouche (Mouth_CNN)

L'architecture proposée dans ce modèle, illustrée dans la **figure** ci-dessous, l'entrée du modèle sera les deux régions des bouches du père et de l'enfant générés par l'algorithme de détection (**landmarks**), ce modèle est composé de 6 couches réparties : 3 couches convolutions, 2 couches pooling et une couche entièrement connectée. Ce modèle a la même architecture que celui choisi pour les régions des yeux.

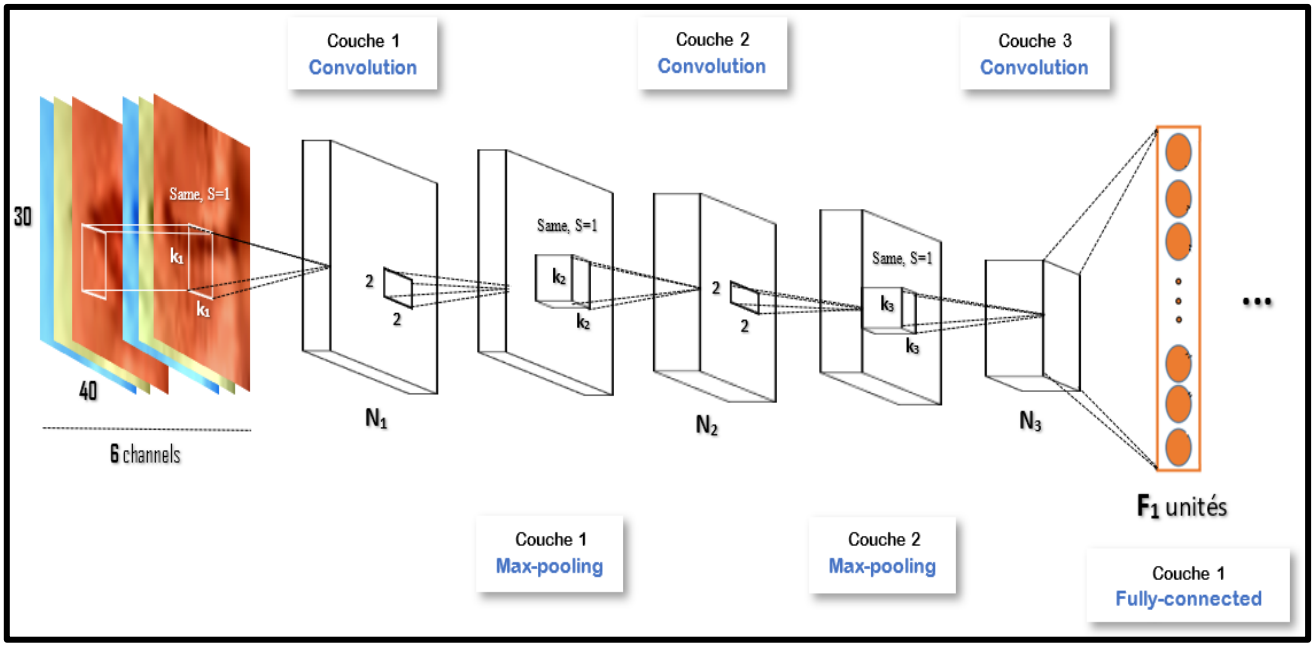


Figure 3.8 : Architecture proposée du modèle **Mouth_CNN**.

3. Expérimentations et résultats

Afin de montrer les performances des méthodes proposées, Une série d'expérimentation a été menée. Dans chaque expérimentation, nous avons évalué et comparé les performances des méthodes d'extraction de caractéristiques sur les deux bases de données KinFaceW-I et KinFaceW-II.

3.1. Bases de données

Pour évaluer les méthodes proposées, nous avons utilisés deux bases de données de parenté : KinFaceW-I et KinFaceW-II (Kin Face in the Wild). Les images de visage sont collectées sur Internet, y compris des images de visage de personnages publics ainsi que des images de visage de parents et d'enfants. Les images de visage sont capturées dans des environnements non contrôlés dans deux ensembles de données sans restriction en termes de pose, d'éclairage, d'arrière-plan, d'expression, d'âge, d'ethnicité et d'occlusion partielle. La différence de KinFaceW-I et KinFaceW-II est que les images de visage avec une relation de parenté ont été acquises à partir de différentes photos dans KinFaceW-I et de la même photo dans KinFaceW-II dans la plupart des cas. La **figure 3.9** montre des échantillons des deux bases de visage.



(a) (b)

Figure 3.9 : Exemples d'image du visage : (a) KinFaceW-I et (b) KinFaceW-II.

Dans ces bases d'images, il existe quatre relations de parenté dans deux ensembles de données : père-fils (**F-S**), père-fille (**F-D**), mère-fils (**M-S**) et mère-fille (**M-D**). Dans l'ensemble de données KinFaceW-I, il existe 156, 134, 116 et 127 paires d'images de parenté pour ces quatre relations. Pour l'ensemble de données KinFaceW-II, chaque relation contient 250 paires d'images de parenté. Pour faciliter l'utilisation, les coordonnées de la position des yeux de chaque image de visage ont été étiquetées manuellement, et la région du visage a été alignée et recadré en **64 x 64** pour supprimer l'arrière-plan. La **figure 3.9** montre des images de visage recadrées à partir des ensembles de données KinFaceW-I et KinFaceW-II, respectivement (de haut en bas se trouvent les relations de parenté F-S, F-D, M-S et M-D, et les deux images voisines dans chaque rangée sont avec la relation de parenté, respectivement). Ces deux bases de données de visages sont disponibles publiquement à l'adresse : https://www.kinfacew.com/datasets.html

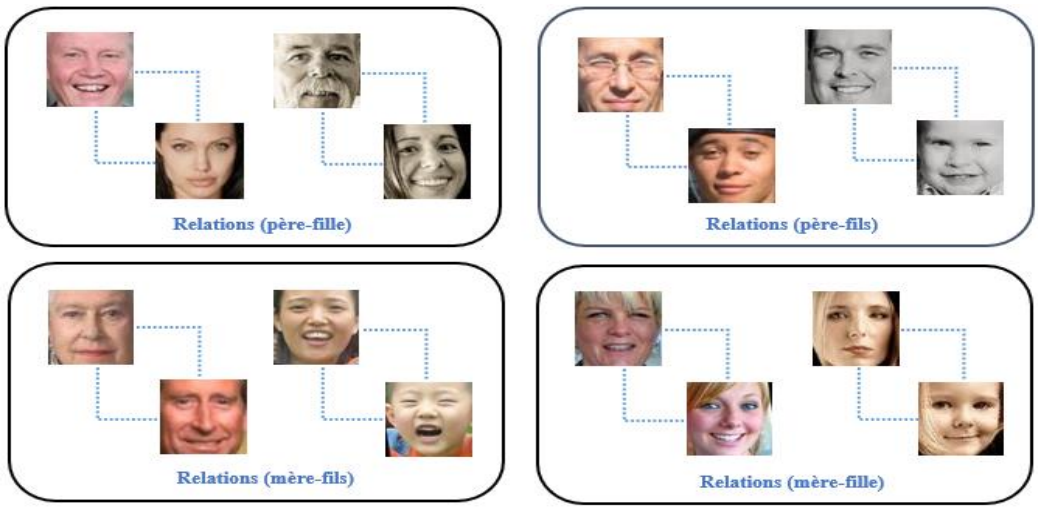


Figure 3.10 : Les quatre types de relations : père-fille, père-fils, mère-fils et mère-fille, qui définit les deux bases des images utilisées.

❖ **Format d'images**

Les images dans les deux bases de données utilisées dans notre travail sont stockées en tant que : (**fd** / **fs** / **md** / **ms**) _ **XXX**_ (**1** / **2**) .jpg

- **(fs / fd / ms / md) :** le type de relation, de « **fs** » désigne la relation père-fils, « **fd** » désigne relation père-fille, « **ms** » désigne relation mère-fils et « **md** » désigne relation mère-fille.
- **XXX :** l'identifiant unique des paires de images (même identifiant pour le parent et l'enfant), dans la base KinFaceW-II, varie de 001 à 250 pour chacune des quatre relations, et pour la base KinFaceW-I, le nombre de paires varie d'une relation à l'autre : père-fils (varie de 001 à 156), père-fille (de 001 à 134), mère-fils (de 001 à 116) et mère-fille (de 001 à 127).
- **(1 / 2) :** l'indice de l'identité de la personne dans la paire, avec « **1** » désigne le parent et « **2** » désigne l'enfant.

3.2. Protocole de tests

Afin d'avoir une comparaison crédible entre les méthodes proposées de vérification de parenté dans la littérature, le protocole de test basé sur la validation croisée K-fold (**k=5**) est fourni avec les deux bases des images KinFaceW-I et KinFaceW-II.

Dans ce protocole de test, chaque sous-ensemble (par exemple, relation père-fils (**F-S**)) de ces ensembles de données KinFaceW-I et KinFaceW-II était également divisé en cinq folds de sorte que chaque fold contienne à peu près le même nombre de paires de visages avec une relation de parenté. Pour évaluer les performances des modèles proposés, chacun des sous-ensembles ou relations et parmi les cinq folds, un seul fold est conservé comme données de test, et les quatre autres folds restantes sont utilisés comme données d'apprentissage. Le processus de validation croisée est répété cinq fois, jusqu'à ce que chaque fold des cinq folds ait été utilisé comme ensemble de test, avec le calcul de la moyenne les cinq résultats comme une estimation de la performance prédictive finale des modèles. Le **tableau 3.1** répertorie les plages d'index des numéros de visages de différents fold de ces deux ensembles de données.

Fold	KinFaceW-I				KinFaceW-II
	F-D	F-S	M-D	M-S	tout sous-ensemble (relations)
1	[1, 27]	[1, 31]	[1, 25]	[1, 23]	[1, 50]
2	[28, 54]	[32, 64]	[26, 50]	[24, 46]	[51, 100]
3	[55, 81]	[65, 96]	[51, 75]	[47, 69]	[101, 150]
4	[82, 108]	[97, 124]	[76, 101]	[70, 102]	[151, 200]
5	[109, 134]	[125, 156]	[102, 127]	[93, 116]	[201, 250]

Tableau 3.1 : Plages d'index des numéros de visages pour 5-folds des ensembles d'images de KinFaceW-I et KinFaceW-II [83].

Plus précisément, pour les images de visage dans chaque fold de ces ensembles de données, toutes les paires d'images de visage avec relation de parenté ont été utilisées comme échantillons positifs, et celles sans relation de parenté comme échantillons négatifs. Par conséquent, les échantillons positifs sont les vraies paires d'images de visage (l'une du parent et l'autre de l'enfant), et les échantillons négatifs sont de fausses paires d'images du visage (l'une du parent et l'autre de l'image de l'enfant qui n'est pas son véritable enfant du parent). En général, le nombre d'échantillons positifs est beaucoup plus faible que celui des échantillons négatifs. Dans notre étude, chaque image de visage de parent a été combinée au hasard avec une image d'enfant qui n'est pas son véritable enfant du parent pour construire une paire négative. De plus, chaque paire d'images de parents et d'enfants est apparue une fois dans les échantillons négatifs [83].

3.3. Evaluation des performances

Les résultats expérimentaux que nous allons présenter dans cette section sont divisés dans deux grandes parties. La première partie contient les résultats expérimentaux des méthodes basées sur les techniques d'extraction classiques. Tandis que la deuxième partie regroupe les méthodes de vérification de parenté basée sur le **CNN**. Dans la fin de ces expériences une comparaison entre les meilleurs systèmes est faite afin de sélectionner le meilleur système.

3.3.1. Performances des méthodes basées sur les techniques d'extraction de caractéristiques classiques

Dans cette section, les performances des méthodes de vérification de parenté basées sur les méthodes d'extraction de caractéristiques classiques proposées sont évaluées. Comme nous l'avons déjà présenté dans la section précédente (**Figure. 3.1**), nous allons tester à chaque fois les résultats de l'un de divers descripteurs de texture récents : **LBP**, **HOG**, **TPLBP** et **FPLBP** sur les bases de données KinFaceW-I et KinFaceW-II. Ensuite, une étape de classification basée la métrique de ***similarité Cosinus*** est utilisé. Finalement, les descripteurs vont être comparés afin de mettre en évidence l'efficacité et la performance des caractéristiques.

A) Résultats obtenus utilisant le descripteur LBP

Nous avons mentionné dans le chapitre précédent que la technique **LBP** dépend de deux paramètres importants. Pour avoir des meilleurs paramètres (paramètres optimaux), nous avons testé le descripteur LBP avec différents nombres d'échantillons voisins et rayons (**N, R**), par exemple (4, 1), (4, 2), (8, 1), (8, 2), (12, 1), (16, 2). L'objectif de la variation dans ces paramètres est d'avoir des meilleurs résultats en fonction de la précision.

Paramètres du **LBP**		Relations de parenté				**Moyennes [%]**
		F-S	**F-D**	**M-S**	**M-D**	
Nombre d'échantillons voisins et rayon (N, R)	**(4, 1)**	67,3	62,6	68,1	71,8	67,5
	(4, 2)	67,0	63,4	68,1	70,8	67,3
	(8, 1)	70,2	69,0	70,7	75,2	71,3
	(8, 2)	69,5	66,8	69,4	71,9	69,4
	(12, 1)	70,2	69,8	**71,5**	75,2	**71,7**
	(12, 2)	**72,1**	68,6	70,7	72,3	70,9
	(16, 1)	70,2	**70,2**	70,3	**75,6**	71,5
	(16, 2)	70,8	66,0	68,9	73,9	69,9

Tableau 3.2 : Performance de la vérification de parenté utilisant le descripteur local **LBP** sur la base KinFaceW-I.

Paramètres du **LBP**		Relations de parenté				**Moyennes [%]**
		F-S	**F-D**	**M-S**	**M-D**	
Nombre d'échantillons voisins et rayon (N, R)	**(4, 1)**	71,2	71,0	70,0	69,0	70,3
	(4, 2)	70,4	69,4	67,8	69,8	69,4
	(8, 1)	77,0	**74,4**	**73,4**	73,4	**74,6**
	(8, 2)	74,6	73,8	72,2	72,2	73,2
	(12, 1)	77,8	72,8	73,2	**74,0**	74,5
	(12, 2)	75,0	73,8	**73,4**	72,2	73,6
	(16, 1)	**78,6**	72,0	73,2	73,6	74,4
	(16, 2)	76,2	73,6	73,0	72,4	73,8

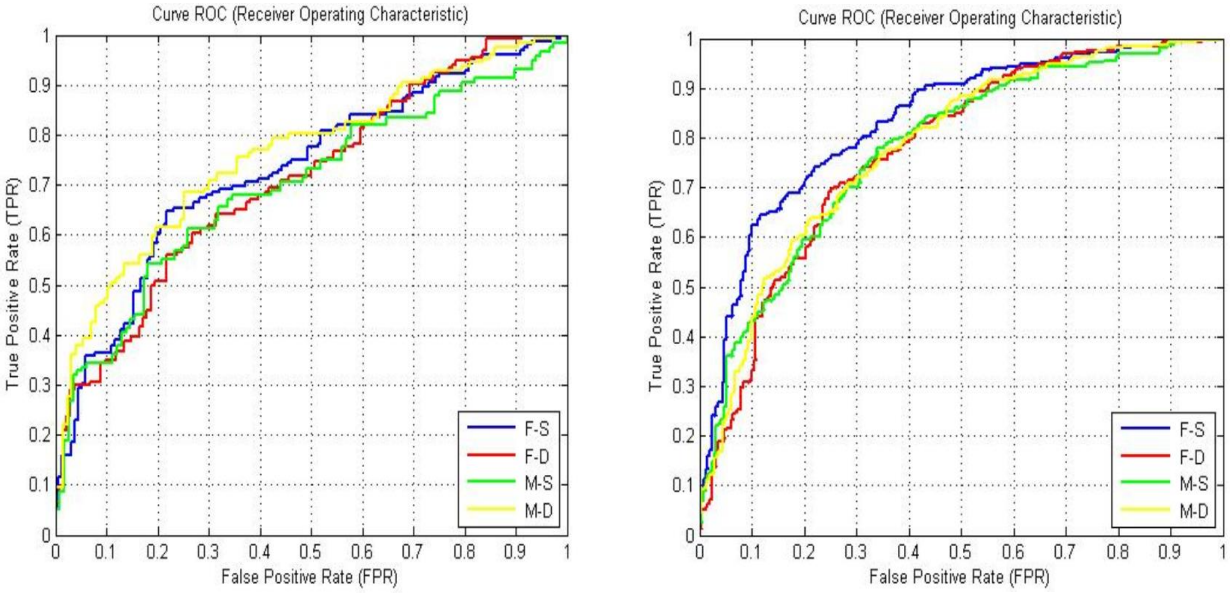
Tableau 3.3 : Performance de la vérification de parenté utilisant le descripteur local **LBP** sur la base KinFaceW-II.

Les **tableaux 3.2** et **3.3** représentent les différents résultats obtenus en fonction de plusieurs nombres d'échantillons et rayon (N, R) et la métrique (***Cosinus similarité)***

Pour la base des images KinFaceW-I, le **tableau 3.2** montre que la méthode proposée offre des résultats acceptables en terme de précision, où nous notons la plage des taux de précision est petite avec environ 5% de différence entre la valeur la plus élevée qui est 75,6% prise dans la relation M-D, pour un nombre d'échantillon égale à 16 et pour rayon égale à 1, et la plus basse 70,2% en la relation F-D avec les mêmes paramètres. Pour la relation F-S, le taux de vérification 70.2% a été obtenu avec un nombre d'échantillon égale à 12 et pour rayon qui égale à 2, le taux de vérification de parenté le plus proche était avec la relation M-S, avec une valeur de 71,5%, ce taux a été atteint avec un nombre d'échantillon égale à 12 et pour rayon égale à 1.

Pour la deuxième base des images KinFaceW-II, le **tableau 3.3** montre que les taux obtenus pour la vérification de parenté sont légèrement plus élevés que les résultats obtenus avec la base des images KinFaceW-I. La précision la plus élevée de 78,6% a été obtenue dans le F-S, par le pair de paramètres suivant (N = 16 et R = 1), le reste des relations était très proche en valeurs

de précision, F-D a atteint une précision de 74,4% avec 8 échantillons et un rayon de longueur 1, le taux de précision de M-D, était de 74,0%, il faut noter que la relation M-S atteignait la précision de 73,4 % avec deux pairs différents de paramètres (N = 8, R = 1) et (N = 12, R = 2). La **figure** suivante nous montre l'évolution de la méthode **LBP** avec des courbes ROC qui démontrent les taux de précision dans chaque relation avec les meilleurs paramètres, à la fois dans les bases de données W-I et II.



(a) **(b)**

Figure 3.11 : Courbes ROC représentant les précisions de la vérification de parenté utilisant le descripteur **LBP** dans les bases (a) KinFaceW-I et (b) KinFaceW-II.

B) Résultats obtenus utilisant le descripteur HOG

Avant d'évaluer les performances de la méthode proposée, une phase de sélection de paramètre concernant le descripteur **HOG** est exécutée. Ça dépend de deux paramètres importants qui sont la taille des blocs et le pourcentage de chevauchement entre les blocs adjacents. Ce que nous intéressent c'est bien la taille de bloc (plusieurs travaux montrent qu'un pourcentage de chevauchement égale à 50% est suffisant pour que l'algorithme fonctionne efficacement). Pour avoir des meilleurs paramètres (paramètres optimaux), nous avons testé le descripteur **HOG** avec des blocs de différents tailles, par exemple 10×10, 12×12,…,32×32 et avec le même pourcentage de chevauchement. L'objectif de la variation dans la taille des blocs est d'avoir des meilleurs résultats en fonction de la précision. Les tableaux suivants représentent les différents résultats obtenus avec les deux bases KinFaceW-I et KinFaceW-II, respectivement en utilisant plusieurs tailles des blocs et la métrique de similarité ***Cosinus***. Dans notre travail, la

sélection des paramètres de **HOG** est basée essentiellement sur le taux de vérification (c'est le paramètre d'évaluation le plus important dans les méthodes de vérification de la parenté).

Paramètres du HOG		Relations de parenté				Moyennes [%]
		F-S	F-D	M-S	M-D	
Taille du block (w×w)	**10*10**	82,4	72,8	73,2	77,2	76,4
	12*12	82,1	72,4	71,9	78,3	76,2
	14*14	82,1	71,7	71,5	75,6	75,2
	16*16	**84,0**	71,7	72,4	77,6	76,4
	18*18	82,4	73,2	71,9	77,9	76,3
	20*20	82,1	71,3	69,8	75,9	74,8
	22*22	83,7	72,8	71,1	75,9	75,9
	24*24	82,4	72,4	72,8	77,9	76,4
	26*26	82,7	73,2	74,2	**78,4**	**77,1**
	28*28	81,4	**73,5**	**73,3**	76,8	76,3
	30*30	81,1	72,8	71,1	75,9	75,2
	32*32	80,8	71,7	68,9	72,8	73,6

Tableau 3.4 : Performance de la vérification de parenté utilisant le descripteur local **HOG** sur la base KinFaceW-I.

Paramètres du HOG		Relations de parenté				Moyennes [%]
		F-S	F-D	M-S	M-D	
Taille du block (w×w)	**10*10**	79,4	71,8	76,2	71,2	74,7
	12*12	79,2	72,8	**76,4**	70,4	74,7
	14*14	79,0	70,6	74,6	71,2	73,9
	16*16	80,4	**73,8**	75,8	71,8	**75,5**
	18*18	**81,0**	72,6	74,8	**73,0**	75,4
	20*20	79,8	70,8	74,2	71,0	74,0
	22*22	77,8	72,6	73,8	69,4	73,4
	24*24	79,4	72,8	74,6	69,8	74,2
	26*26	80,2	73,0	74,4	71,2	74,7
	28*28	80,0	72,6	74,6	71,6	74,7
	30*30	79,2	70,8	74,2	70,2	73,6
	32*32	75,0	70,8	71,0	66,2	70,8

Tableau 3.5 : Performance de la vérification de parenté utilisant le descripteur local **HOG** sur la base KinFaceW-II.

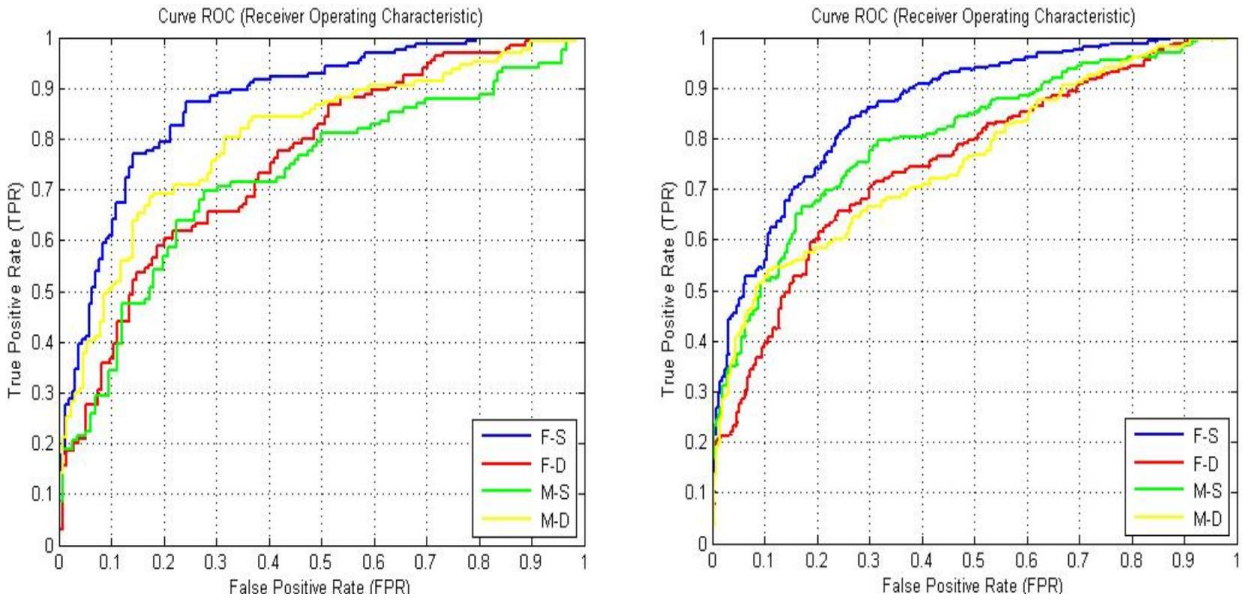
Les **tableaux 3.4** et **3.5** présentent les différents résultats obtenus avec les deux bases KinFaceW-I et KinFaceW-II, en utilisant plusieurs tailles des blocs.

Pour la base des images KinFaceW-I, le **tableau 3.4** montre que la méthode proposée offre de bons résultats en termes de précision et en particulier pour la relation F-S, avec la meilleure précision égale à 84.0% avec un bloc de taille 16×16. Les taux de vérification les plus petits

73.3%, 73.5% ont été obtenus en appliquant 28x28 comme taille de bloc dans les relations F-D et M-S, respectivement. Alors pour la dernière relation M-D, le taux de vérification de parenté était moyen avec 78.4% et une taille de bloc égale à 26x26.

Pour la deuxième base des images KinFaceW-II, les taux de vérification de parenté obtenus sont presque les mêmes que les résultats obtenus avec la base des images KinFaceW-I. Une précision égale à 81.0% est obtenu pour la relation F-S avec un bloc de taille 18×18. Avec une taille de bloc 12x12, la vérification de la relation M-S est acceptable avec une précision égale à 76.4%. Pour les taux de vérification des relations F-D et M-D étaient les moins précis avec des valeurs de 73.8% et 73.0% avec des tailles de bloc 16x16 et 18x18, respectivement.

Pour mieux illustrer l'efficacité de la méthode proposée pour la vérification de la parenté pour différentes relations, la **figure 3.8** montre les courbes ROC, qui clarifient l'évolution de la performance de classification du système avec le meilleur paramétrage pour chaque relation.



(a) **(b)**

Figure 3.12 : Courbes ROC représentant les précisions de la vérification de parenté utilisant le descripteur **HOG** dans les bases (a) KinFaceW-I et (b) KinFaceW-II.

C) Résultats obtenus utilisant le descripteur TPLBP

Afin d'extraire les caractéristiques de texture avec le descripteur **TPLBP**, quatre paramètres doivent être définis à savoir la taille de blocs, la taille des patchs w, le nombre de patchs distribués autour de l'anneau S et le rayon r. Dans nos expérimentations, la taille des blocs est fixée à 16 x 16. Notons que ce choix est maintenu après plusieurs tests empiriques. Nous devons aussi noter que les images sont de taille 64 x 64 et cela génère 16 blocs de taille 16 x 16.

Pour avoir des meilleurs paramètres, nous avons testé la technique **TPLBP** avec des patchs de différentes tailles, différents nombres de patchs voisins, et avec différents rayons. Par exemple (4, 3), (4, 5), (6, 3), (6, 5), (8, 3), (8, 5) et avec différents rayons de 1...6. L'objectif de la variation ces paramètres est d'avoir des meilleurs résultats en termes de précision.

Les tableaux suivants présentent les différents résultats obtenus avec les modifications apportées aux valeurs des paramètres par les tests empiriques sur les deux bases KinFaceW-I et KinFaceW-II.

Types de relations	Paramètres du **TPLBP**		*Nombre de patchs voisins et taille de patchs (**S, w**)*					
			(4, 3)	**(4, 5)**	**(6, 3)**	**(6, 5)**	**(8, 3)**	**(8, 5)**
F-S	*Rayon (**r**)*	**1**	66,7	64,8	73,4	74,1	73,7	74,0
		2	64,8	64,4	75,6	75,9	74,4	74,1
		3	64,8	64,7	74,9	75,9	77,9	**78,9**
		4	63,2	62,2	76,3	75,9	77,6	75,6
		5	64,1	63,8	74,9	74,7	74,9	76,3
		6	64,4	67,3	73,7	73,0	73,7	74,4
F-D	*Rayon (**r**)*	**1**	59,4	63,8	67,2	70,5	66,5	69,8
		2	59,3	63,0	66,8	67,2	67,6	67,9
		3	62,3	62,3	69,0	70,2	**70,6**	67,6
		4	62,7	59,7	66,8	66,5	68,7	67,9
		5	60,8	58,9	66,8	67,6	69,8	70,2
		6	61,9	62,3	68,3	66,4	70,2	70,2
M-S	*Rayon (**r**)*	**1**	61,6	66,4	68,9	68,1	66,8	67,7
		2	57,7	64,2	69,8	69,8	67,2	66,8
		3	61,2	66,3	69,4	72,4	67,2	66,8
		4	64,6	67,2	71,1	**73,6**	69,8	71,5
		5	67,2	65,1	71,5	73,2	72,3	**73,6**
		6	66,8	63,4	70,6	71,5	73,2	72,6
M-D	*Rayon (**r**)*	**1**	64,1	63,8	70,1	69,4	70,4	70,9
		2	60,6	63,8	71,3	71,3	70,9	71,3
		3	64,2	62,3	73,2	72,0	70,9	72,4
		4	65,3	63,1	72,4	73,5	75,9	73,9
		5	61,8	63,0	73,6	75,2	**76,4**	74,4
		6	65,3	65,4	73,9	73,2	74,4	74,0

Tableau 3.6 : Performance de la vérification de parenté utilisant le descripteur local **TPLBP** sur la base KinFaceW-I.

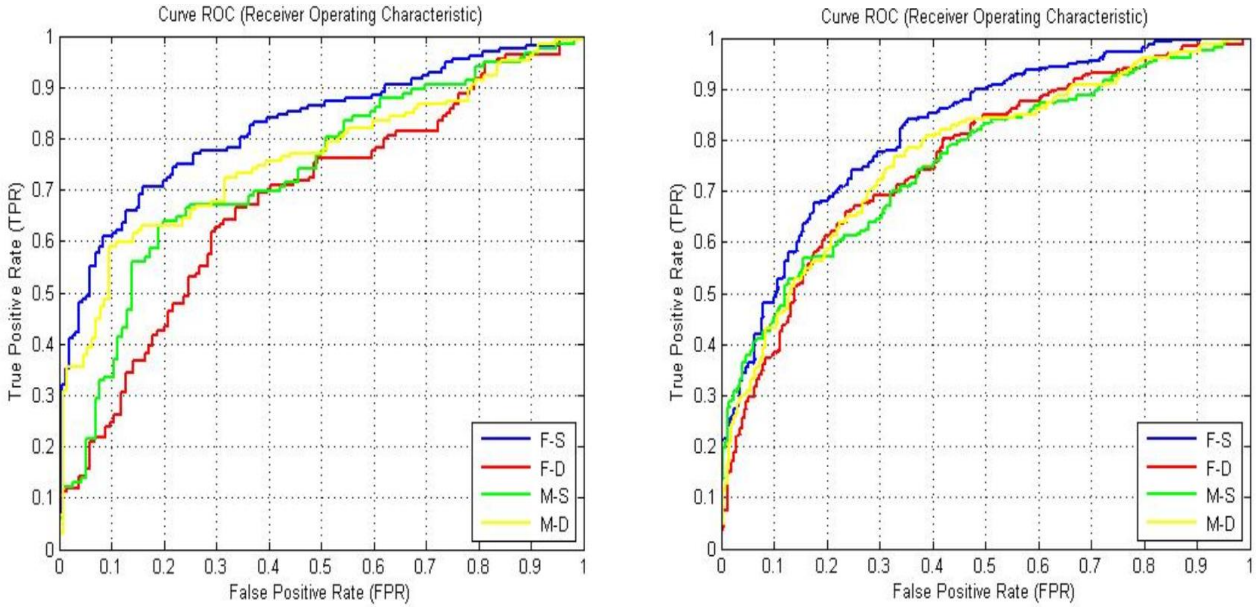
Types de relations	Paramètres du TPLBP		*Nombre de patchs voisins et taille de patchs (S, w)*					
			(4, 3)	**(4, 5)**	**(6, 3)**	**(6, 5)**	**(8, 3)**	**(8, 5)**
F-S	*Rayon (r)*	**1**	67,2	61,6	73,2	74,6	74,6	76,2
		2	61,6	63,2	77,6	77,2	76,0	75,8
		3	61,6	62,4	76,0	75,8	76,8	77,6
		4	61,6	62,0	76,0	76,6	76,4	**78,0**
		5	63,2	61,4	74,6	75,2	76,4	76,6
		6	61,4	62,6	74,8	73,4	77,4	77,8
F-D	*Rayon (r)*	**1**	61,6	61,2	62,4	65,2	64,0	65,4
		2	57,8	60,2	65,4	65,6	66,4	67,2
		3	57,0	59,2	68,0	69,0	65,4	68,0
		4	59,0	62,4	66,8	70,2	67,6	69,2
		5	59,6	61,6	67,8	69,8	70,0	72,0
		6	59,4	60,4	67,6	71,6	72,2	**73,6**
M-S	*Rayon (r)*	**1**	66,0	62,0	69,6	69,0	69,2	68,4
		2	62,0	58,4	70,4	70,8	70,0	69,2
		3	62,4	59,8	71,0	73,0	70,8	70,6
		4	58,6	61,0	71,0	71,0	71,8	72,0
		5	59,6	62,0	71,4	73,4	72,0	72,0
		6	61,6	62,8	72,8	73,0	71,6	**73,8**
M-D	*Rayon (r)*	**1**	65,4	61,2	67,4	68,0	65,8	67,8
		2	59,8	62,6	70,8	70,2	67,2	67,8
		3	63,0	61,8	70,0	71,2	68,0	70,6
		4	63,2	60,4	70,0	71,2	69,8	71,6
		5	63,6	59,0	71,0	72,2	71,2	72,8
		6	59,4	59,0	71,2	73,2	72,0	**73,6**

Tableau 3.7 : Performance de la vérification de parenté utilisant le descripteur local **TPLBP** sur la base KinFaceW-II.

Les **tableaux 3.6** et **3.7** montrent les résultats des tests empiriques du la méthode proposée basée sur la technique **TPLBP** avec les changements des valeurs des différents paramètres. Nous observons que pour la base KinFaceW-I, une meilleure précision égale à 78.9%, est obtenus pour la relation F-S avec un rayon de 3, un nombre de patchs voisins égale à 8 et une taille de patchs égale à 5. Un taux moyen égale à 70.6 % a été obtenu en utilisant les valeurs r = 3, S = 8, w = 3 pour la relation F-D, tandis que pour la relation M-S, un taux égal à 73.6% a été obtenu avec deux paramétrages différents (r=4, S=6, w=5) et (r=6, S=8, w=5). Pour la dernière relation M-D, la précision était avec un taux de 76.4%, avec r=5, S=8 et w=3.

Pour la base des images KinFaceW-II, les taux de précision étaient dans la même plage que la première base de données, mais il y a eu quelques changements mineurs qui peuvent être remarqués, la meilleure précision égale à 78.0%, elle a été obtenu pour la relation F-S avec un rayon d'une longueur de 4, 8 patchs voisins et avec taille de patchs 5. Nous remarquons que les

trois relations restantes possèdent le même paramétrage (r=6, S=8, w=5), deux parmi eux (F-D et M-D) ont la même précision qui est 73.6%, cependant la relation M-S était plus précise avec 73.8%.



(a) **(b)**

Figure 3.13 : Courbes ROC représentant les précisions de la vérification de parenté utilisant le descripteur **TPLBP** dans les bases (a) KinFaceW-I et (b) KinFaceW-II.

La **figure 3.13** montre les courbes ROC représentant les taux de précisions pour la vérification de parenté pour chaque relation avec les valeurs optimales pour chaque paramètre dans les deux bases KinFaceW-I et II.

D) Résultats obtenus utilisant le descripteur FPLBP

D'une façon similaire que la méthode **TPLBP**, avant d'évaluer les performances de la méthode proposée, une phase de sélection de paramètre concernant l'algorithme **FPLBP** est exécutée. Ça dépend de cinq paramètres importants qui sont la taille des blocs, la taille des patchs w, le nombre de patchs S et deux rayons au lieu d'un (r1 et r2) formant les deux anneaux autour du pixel (extérieur et intérieur). Pour avoir des meilleurs paramètres (paramètres optimaux), nous avons testé l'algorithme **FPLBP** avec des patchs de différents nombres et tailles, par exemple (4, 3), (4, 5), (6, 3), (6, 5), (8, 3), (8, 5), (10, 3), (10, 5), (12, 3), (12, 5) et avec différents rayons (1, 2), (1, 4), (1, 6), (1, 8), (3, 4), (3, 6), (3, 8), (5, 6), (5, 8), (7, 8).
En raison du grand nombre de résultats obtenus, le **tableau 3.8** présente les meilleurs taux de précision obtenus pour chaque relation dans les deux bases d'images KinFaceW-I et II.

Types de relations	Paramètres du **FPLBP**		Précision [%]
	Taille et nombre de patchs ***(s,w)***	*Rayons des anneaux* ***(r_1,r_2)***	
	Bases d'images KinFaceW-I		
F-S	(12, 3)	(3, 4)	**76.3**
F-D	(12, 3)	(7, 8)	**73.2**
M-S	(12, 5)	(3, 6)	**75.5**
M-D	(10, 5)	(7, 8)	**78.8**
	Bases d'images KinFaceW-II		
F-S	(12, 5)	(1, 2)	**78.2**
F-D	(12, 5)	(1, 8)	**72.6**
M-S	(10, 5)	(1, 8)	**73.4**
M-D	(10, 5)	(7, 8)	**72.2**

Tableau 3.8 : Meilleurs performances de la vérification de parenté utilisant le descripteur local **FPLBP** sur les deux bases KinFaceW-I et II.

Le **tableau 3.8** montre que les résultats obtenus avec la méthode **FPLBP** sont meilleurs que ceux obtenus avec la méthode **TPLBP**, et c'était à propos de la base d'images KinFaceW-I. Cependant, dans la base d'images KinFaceW-II, les résultats étaient presque les mêmes. Les taux de précisions créés par les tests empiriques sur les modifications apportées aux valeurs des paramètres sont indiqués.

Afin de montrer les performances de la méthode **FPLBP** pour la vérification de parenté, nous avons présenté, dans la **figure 3.14**, une comparaison entre les quatre relations avec toujours les valeurs optimales pour le paramétrage dans les deux bases KinFaceW-I et II.

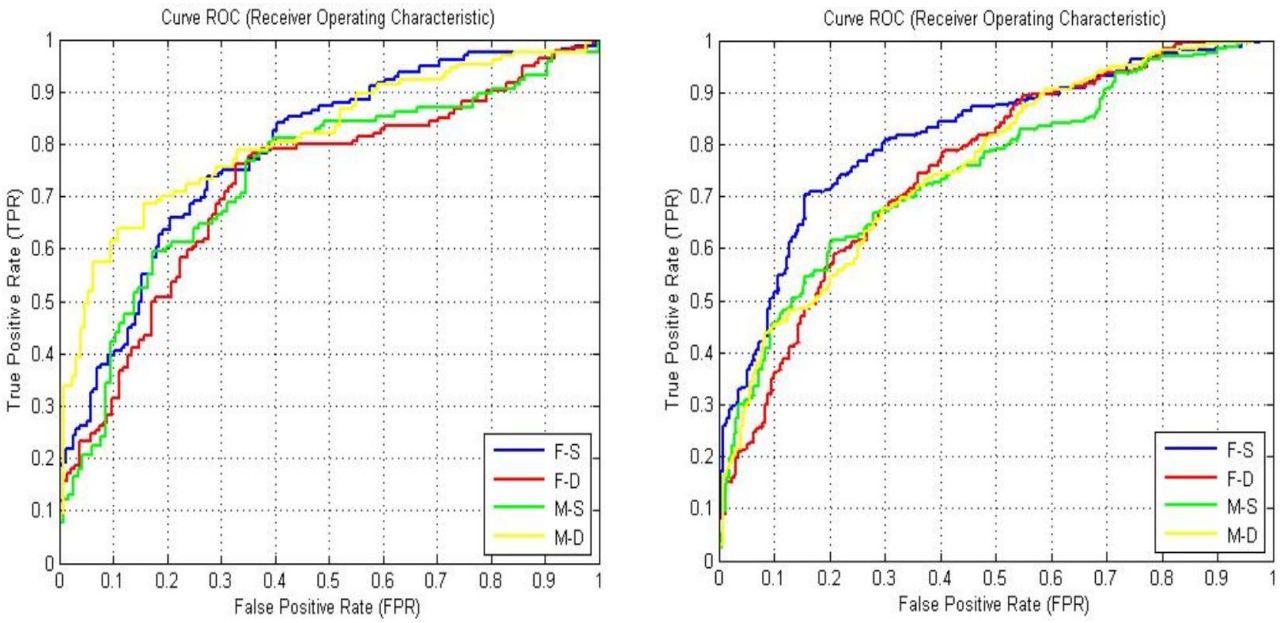


Figure 3.14 : Courbes ROC représentant les précisions de la vérification de parenté utilisant l'algorithme **FPLBP** dans les bases (a) KinFaceW-I et (b) KinFaceW-II.

3.3.2. Performances des méthodes basées sur l'apprentissage en profondeur

Dans cette section, nous présentons les performances du deuxième schéma de vérification de parenté basée sur les réseaux de neurones conventionnels en l'occurrence le **CNN** basique et le **KP-CNN** proposé.

A) Vérification de parenté basée sur le CNN basique

L'architecture de tous **CNN** est basée généralement de plusieurs paramètres dont les principaux : la taille des filtres et le nombre de filtres dans chaque couche de convolution ainsi que le nombre d'unités dans la couche entièrement connectée (FC). Dans notre travail, une série d'expérimentations a été réalisé pour chaque relation de parenté afin de sélectionner les meilleurs paramètres offrant les meilleurs taux de vérification de parenté exprimé en termes de précision. Dans nos expérimentations, pour chaque relation de parenté nous avons utilisé différents filtres de taille 3×3, 5×5, 7×7, et 11×11 et différents nombres de filtres de 2, 4, 8, 16, 32, 64, 128, et 256 et une couche entièrement connectée avec un nombre d'unités différentes dans le premier stage afin de sélectionner les meilleurs paramètres (taille et nombre de filtre). Ensuite, ce processus est répété dans le deuxième et troisième stage.

Notons que, dans une première étape, nous avons considéré une architecture de trois stages. Ensuite, nous avons remarqué après les tests effectués que le taux de la vérification de parenté pour les différentes relations ait été diminué dans le troisième stage. C'est la raison pour laquelle une architecture de **CNN** de deux stages a été maintenue.

Base KinFaceW-I

Type de relation	Stage 1 **(Conv1)**		Stage 2 **(Conv2)**		Nombre des unités **(F)**	**Accuracy** [%]
	Nombre de Filtre **(N_1)**	Taille de Filtre **(k_1)**	Nombre de Filtre **(N_2)**	Taille de Filtre **(k_2)**		
F-S	64	7×7	32	7×7	**640**	**79.33**
F-D	64	7×7	16	5×5	**320**	**71.26**
M-S	4	5×5	64	11×11	**640**	**65.28**
M-D	16	7×7	8	5×5	**320**	**82.95**

Tableau 3.9 : Performance de la vérification de parenté basée **CNN** basique dans la base KinFaceW-I.

D'une façon similaire les quatre relations sont évaluées pour la base KinFaceW-II et les résultats sont représentés dans le **tableau 3.10**. Les taux de vérification des relations sont meilleurs par rapport à ceux obtenus dans la base KinFaceW-I.

Base KinFaceW-II						
Type de relation	Stage 1 **(Conv1)**		Stage 2 **(Conv2)**		Nombre des unités **(F)**	**Accuracy** [%]
	Nombre de Filtre **(N_1)**	Taille de Filtre **(k_1)**	Nombre de Filtre **(N_2)**	Taille de Filtre **(k_2)**		
F-S	32	11×11	4	7×7	**160**	**87.99**
F-D	16	7×7	32	3×3	**640**	**83.19**
M-S	32	5×5	64	11×11	**640**	**85.69**
M-D	32	7×7	16	7×7	**320**	**90.40**

Tableau 3.10 : Performance de la vérification de parenté basée **CNN** basique dans la base KinFaceW-II.

B) Vérification de parenté basée sur le KP-CNN

L'architecture du **KP-CNN** proposée pour la vérification de parenté est basée sur trois **CNN** basique utilisés pour chaque région de l'image (yeux, nez, et la bouche). D'une façon similaire, une série d'expérimentations a été réalisé pour chaque modèle et pour chaque relation de parenté afin de sélectionner les meilleurs paramètres offrant les meilleurs taux de vérification de parenté exprimé en termes de précision. Dans nos expérimentations, dans chaque modèle et pour chaque relation de parenté nous avons utilisé différents filtres de taille 3×3, 5×5, 7×7, et 11×11 et différents nombres de filtres de 2, 4, 8, 16, 32, 64, 128, et 256 et différents nombres d'unités à la fois dans la première et la deuxième couche entièrement connectée dans le premier stage afin de sélectionner les meilleurs paramètres (taille et nombre de filtre). Ensuite, ce processus est répété dans le deuxième et troisième stage. En raison de grand nombre de résultats (4×3×3=36 tableaux), les **tableaux 3.11** et **3.12** présentes un récapitulatif des meilleurs précisons obtenues dans les deux bases d'images KinFaceW-I et KinFaceW-II en fonction de la taille et le nombre des filtres sélectionnés.

Le **tableau 3.11** indique que la méthode **KP-CNN** améliore la précédente méthode avec des taux d'amélioration de **0.9%** pour la relation **F-S**, **1.23%**, pour la relation **F-D**, **2.42%**, pour la relation **M-S**, et **0.75%** pour la relation **M-D** sur l'ensemble des images de KinFaceW-I.

Base KinFaceW-I									
Type de relation	Régions	Stage 1 (Conv1)		Stage 2 (Conv2)		Stage 3 (Conv3)		Fully-connected		Accuracy [%]
		Taille de Filtre (k_1)	Nombre de Filtre (N_1)	Taille de Filtre (k_2)	Nombre de Filtre (N_2)	Taille de Filtre (k_3)	Nombre de Filtre (N_3)	Nombres des unités (F_1)	Nombres des unités (F_2)	
F-S	Yeux	5×5	16	7×7	8	3×3	32	**320**	**1280**	**80.23**
	Nez	3×3	32	5×5	16			**160**		
	Bouche	5×5	32	3×3	16	5×5	64	**640**		
F-D	Yeux	5×5	64	7×7	16	3×3	64	**640**	**1280**	**71.60**
	Nez	3×3	16	5×5	32			**320**		
	Bouche	5×5	32	3×3	16	5×5	128	**640**		
M-S	Yeux	5×5	16	7×7	64	3×3	64	**640**	**1280**	**67.70**
	Nez	3×3	64	5×5	32			**320**		
	Bouche	5×5	32	3×3	16	5×5	64	**640**		
M-D	Yeux	5×5	16	7×7	8	3×3	16	**160**	**1280**	**83.70**
	Nez	3×3	32	5×5	16			**160**		
	Bouche	5×5	32	3×3	32	5×5	64	**640**		

Tableau 3.11 : Performance de la vérification de parenté basée **KP-CNN** dans la base KinFaceW-I.

D'une façon similaire, les résultats présentés dans le **tableau 3.12** indiquent que la méthode KP-CNN améliore les taux obtenus dans la méthode précédente avec des taux d'amélioration de **0.4%** pour la relation **F-S**, **0.21%**, pour la relation **F-D, 1.6%**, pour la relation **M-S**, et **3.21%** pour la relation **M-D** sur l'ensemble des images de KinFaceW-II.

Base KinFaceW-II									
Type de relation	Régions	Stage 1 (Conv1)		Stage 2 (Conv2)		Stage 3 (Conv3)		Fully-conncted		Accuracy [%]
		Taille de Filtre (k_1)	Nombre de Filtre (N_1)	Taille de Filtre (k_2)	Nombre de Filtre (N_2)	Taille de Filtre (k_3)	Nombre de Filtre (N_3)	Nombres des unités (F_1)	Nombres des unités (F_2)	
F-S	Yeux	5×5	32	7×7	8	3×3	128	**640**	**1280**	**88.0**
	Nez	3×3	32	5×5	16			**160**		
	Bouche	5×5	32	3×3	16	3×3	64	**640**		
F-D	Yeux	5×5	64	7×7	16	3×3	128	**640**	**1280**	**84.60**
	Nez	3×3	16	5×5	32			**320**		
	Bouche	5×5	32	3×3	16	5×5	128	**640**		
M-S	Yeux	5×5	16	7×7	64	3×3	64	**640**	**1280**	**87.70**
	Nez	3×3	64	5×5	32			**320**		
	Bouche	5×5	32	3×3	16	5×5	64	**640**		
M-D	Yeux	5×5	16	7×7	8	3×3	128	**640**	**1280**	**93.70**
	Nez	3×3	32	5×5	16			**160**		
	Bouche	5×5	32	3×3	32	5×5	32	**320**		

Tableau 3.12 : Performance de la vérification de parenté basée **KP-CNN** dans la base KinFaceW-II.

4. Etude comparative

Dans cette section, nous avons effectué une étude comparative concernant les différentes méthodes proposées. Le taux de vérification exprimé en précision (%) est le taux de comparaison.

Le but est de sélectionner la meilleure méthode pour réaliser un système de vérification de la parenté.

D'après le **tableau 3.13**, on peut dire que les résultats obtenus par les différentes méthodes sont acceptables et compétitifs. On remarque que les meilleurs résultats pour la vérification des quatre relations (F-S, F-D, M-S, et M-D) sont obtenus par la méthode proposée basée **KP-CNN** avec des taux **88.00 %**,**84.60 %**, **87.70 %**, et **93.97 %** sur l'ensemble des images de KinFaceW-II.

Tandis sur l'ensemble d'images de KinFaceW-I, les meilleurs taux de vérification des relations F-S et F-D sont donnés par la méthode basée **HOG** avec des taux de 84.00% et 73.50%, respectivement. Alors que la méthode proposée basée sur **FPLBP** offre le meilleur taux pour la relation M-S avec un taux de 75.50%. Une autre fois la méthode proposée basée **KP-CNN** offre le meilleur taux pour la relation M-D avec un taux de 83.70%.

Relation de parenté	Méthodes proposées basées sur					
	LBP	HOG	TPLBP	FPLBP	CNN	KP-CNN
	Bases d'images KinFaceW-I					
F-S	72.10	84.00	78.90	76.3	79.33	80.23
F-D	70.20	73.50	70.60	73.2	71.26	71.60
M-S	71.50	73.30	73.60	75.5	65.28	67.70
M-D	75.60	78.40	76.40	78.8	82.95	83.70
	Bases d'images KinFaceW-II					
F-S	78.60	81.00	78.00	78.2	87.99	88.00
F-D	74.40	73.80	73.60	72.6	83.19	84.60
M-S	73.40	76.40	73.80	73.4	85.69	87.70
M-D	74.00	77.00	73.60	72.2	90.40	93.70

Tableau 3.13 : Comparaison des méthodes proposées.

5. Conclusion

Dans ce chapitre, nous avons proposé plusieurs méthodes de vérification de parenté basées sur des méthodes d'extraction classique (Handcraft) et d'autre basées sur l'apprentissage en profondeur. Ces différentes méthodes sont testées dans le but d'améliorer le taux de vérification de parenté pour chaque relation de parenté (F-S, F-D, M-S, et M-D). En validant ces méthodes sur les deux bases de données KinFaceW-I et KinFaceW-II, nous avons dégagé une amélioration considérable des taux de vérification de parenté avec la méthode **KP-CNN** avec l'ensemble des images de la base KinFaceW-II. Cette méthode a montré une amélioration remarquable par rapport aux d'autres méthodes proposées (**LBP**, **HOG**, **TPLBP**, **FPLBP**, et **CNN**). Tandis pour la base de données KinFaceW-I, les méthodes **LBP**, **HOG**, **TPLBP**, **FPLBP**, **CNN** et **KP-CNN** sont compétitives pour la vérification de différentes relations.

Conclusion Générale

Les performances des systèmes de reconnaissance de formes sont toujours liées à la méthode d'extraction des caractéristiques. En fait, un système de vérification de parenté représente l'un de systèmes les plus importants dans le domaine de la reconnaissance des formes, dont l'efficacité est jugée par la précision de la vérification ou l'identification de parenté. Les tendances de recherche sur les systèmes de vérification de parenté sont basées sur la manière d'extraction de différentes relations de parenté rapidement et facilement à partir des images faciales. Pour un ordinateur, cette compétence est très difficile vue des changements dans les traits du visage, l'angle de vue, les conditions d'éclairage, la pose, les occlusions, modifications de l'expression faciale et de la coiffure, ou effets du vieillissement.

Afin d'analyser les différentes relations de parenté, dans ce travail, nous avons proposé quatre schémas de vérification de parenté basés sur les méthodes d'extraction de caractéristiques LBP, HOG, TPLBP et FPLBP. Une série d'expérimentions est réalisée avec les deux bases de données KinFaceW-I et KinFaceW-II et un classifieur KPPV basé sur la métrique de similarité cosinus. Les résultats expérimentaux ont montré que les méthodes d'extraction de caractéristiques HOG et FPLBP sont compétitives et offres une précision de vérification de parenté améliorer par rapport aux travaux existants. Nous avons constaté aussi que les méthodes classiques d'extraction de caractéristiques ne peuvent pas bien représenter les informations sous-jacentes de parenté impliquée dans les images faciales, et donc la performance de la vérification de parenté reste toujours non satisfaisante.

C'est la raison pour laquelle dans la seconde partie de nos travaux, nous avons proposé d'utiliser les réseaux de neurones convolutifs basée comme le CNN. Afin d'exploiter les performances du CNN dans l'analyse des relations de parenté, dans une première étape, nous utilisé pour chaque relation de parenté un CNN appliqué sur l'image faciale globale. Alors, dans une deuxième étape, nous avons utilisé un groupe de CNN pour chaque relation sur des régions du visage obtenues par un prédicteur des points clés dans l'image faciale.

En perspective pour les travaux futurs de cette étude nous proposons de :

- Augmenter les bases de données.

- La fusion des vecteurs de caractéristiques obtenus par les deux techniques HOG et FPLBP.
- L'utilisation d'autres techniques d'apprentissage en profondeur comme ICANet et DCTNet.
- Tester les modèles CNN pré-entraînés comme ResNet, VGG16 et AlexNet.

Bibliographie

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

[2] *International Encyclopedia of the Social Sciences (2008) Kinship.*

[3] *Blaustein AR, O'Hara RK (1981) Genetic control for sibling recognition? Nature 290(5803), pp.246–248.*

[4] *Lieberman D, Tooby J, Cosmides L (2007) The architecture of human kin detection. Nature 445(7129), pp.727–731*

[5] *International Encyclopedia of the Social Sciences (1968) Kinship.*

[6] *T. B. Underpinning and A. H. Relationships, "Kinship: Definition in the Study of Sociology," pp. 1–3, 2020.*

[7] *Jha M (1994) Introduction to Indian Anthropology. Vikas Publishing House Pvt Ltd.*

[9] *Jiwen L, Junlin H, Zhou X, Shang Y, Tan Y-P, Wang G (2012) Neighborhood repulsed metric learning for kinship verification. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 2594–2601.*

[10] *Fang R, Tang KD, Snavely N, Chen T (2010) Towards computational models of kinship verification. In: 2010 17th IEEE international conference on image processing (ICIP). IEEE, pp 1577–1580.*

[11] *Lu J, Zhou X, Tan YP, Shang Y, Zhou J (2014) Neighborhood repulsed metric learning for kinship verification. IEEE Trans Pattern Anal Mach Intell 36(2), p 331.*

[12] *Xia S, Shao M, Yun F (2011) Kinship verification through transfer learning. In: IJCAI Proceedings- international joint conference on artificial intelligence, vol 22, p 2539.*

[13] *Guo Y, Dibeklioglu H, van der Maaten L (2014) Graph-based kinship recognition. In: 22nd international conference on pattern recognition (ICPR 2014).*

[14] *Jafri R, Arabnia HR (2009) A survey of face recognition techniques. JIPS 5(2), pp 41–68.*

[15] *Somanath G, Rohith MV, Kambhamettu C (2011) Vadana: a dense dataset for facial image analysis. In: 2011 IEEE international conference on computer vision workshops (ICCV Workshops). IEEE, pp 2175– 2182.*

[16] *Somanath G, Kambhamettu C (2012) Can faces verify blood-relations? In: 2012 IEEE Fifth international conference on biometrics: theory, applications and systems (BTAS). IEEE, pp 105–112.*

[17] *Taigman Y, Yang M, Ranzato M, Wolf L (2014) Deepface: closing the gap to human-level performance in face verification. In: 2014 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 1701–1708.*

[18] *Zhang, K., Huang, Y., Song, C., Wu, H. & Wang, L. Kinship Verification with Deep Convolutional Neural Networks. 148.1-148.12 (2015). doi:10.5244/c.29.148*

[19] *Biswas S, Bowyer KW, Flynn PJ (2011) A study of face recognition of identical twins by humans. In: 2011 IEEE international workshop on information forensics and security (WIFS). IEEE, pp 1–6.*

[20] *Lamba H, Sarkar A, Vatsa M, Singh R, Noore A (2011) Face recognition for look-alikes: a preliminary study. In: 2011 international joint conference on biometrics (IJCB). IEEE, pp 1–6.*

[21] *Lin S-H (2000) An introduction to face recognition technology. Informing Science 3(1):1–8.*

[22] *Senior AW, Bolle RM (2002) Face recognition and its application. In: Biometric Solutions. Springer, pp 83–97.*

[23] *D. Xu, S. Yan, L. Zhang, S. Lin, H. Zhang, and T. S. Huang, "Reconstruction and Recognition of Tensor-Based Objects With Concurrent Subspaces Analysis," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, pp. 36-47, 2008.*

[24] *D. Xu, S. Yan, L. Zhang, S. Lin, H. Zhang, and T. S. Huang, " Reconstruction and Recognition of Tensor-Based Objects With Concurrent Subspaces Analysis," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, pp. 36-47, 2008.*

[25] *M BESSAOUDI, Thèse de Doctorat LMD en Génie Electrique, Filière : Electronique, Option : Signaux et Communication, « Reconnaissance de Visage basée sur l'Analyse Multidimensionnelle », Université Mohamed Khider Biskra, le 27 Juin 2019.*

[26] *Maloney LT, Dal Martello MF (2006) Kin recognition and the perceived facial similarity of children. J Vis 6(10):4*

[27] *Qin, X., Liu, D. & Wang, D. A literature survey on kinship verification through facial images. Neurocomputing **377**, 213–224 (2020).*

[28] *S. Hamida. " L'extraction de caractéristiques pour l'analyse biométrique tridimensionnelle d'un visage ". Thèse de Mastère 2, Département d'informatique, Université de Mohamed Khider Biskra, Juin 2010.*

[29] *Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 721-732, 1997.*

[30] *S. G. Ababsa. "Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D ". Thèse de doctorat, Université Evry Val d'Essonne. 03 octobre 2008.*

[31] *D. Blackburn, M. Bone, P. J Phillips. "Face recognition vendor test 2000". Tech. Rep. http://www.frvt.org, 2001.*

[32] *P. J. Phillips, P. J. Grother, R. J. Micheals, D. M. Blackbrun, E. Tabassi, and J. M. Bone. "Face recognition vendor test 2002: Evaluation report". NISTIR 6965. Available online at http://www.frvt.org, 2003.*

[33] *C. Beumier. "Authentification d'identité par analyse tridimensionnelle du visage ". Thèse de doctorat, Université de Paris, Décembre 2003.*

[34] *Xia S, Shao M, Luo J, Fu Y (2012) Understanding kin relationships in a photo. IEEE Trans Multimedia 14(4), pp 1046–1056.*

[35] *T. Ojala, M. Pietikäinen, and D. Harwood: A comparative study of texture measures with classifiction based on featured distribution. Pattern Recognition (Elsevier). Vol.29, No.1, pp.51-59, 1996.*

[36] *Huang, D., Shan, C., Ardabilian, M., Wang, Y. & Chen, L. Local binary patterns and its applications on facial image: A survey. IEEE Trans. Syst. Man, Cybern. C Appl. Rev.* ***41****, 765–781 (2011).*

[37] *T. Ojala, M. Pietikäinen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Analys is and Machine Intelligence, vol. 24, no. 7, pp. 971–987, 2002.*

[38] *K.Saadallah « Système de reconnaissance des empreintes palmaires », ". Mémoire de Master, Spécialité Système de Télécommunications, Département d'électronique, Université Badji Mokhtar. Annaba, 2019.*

[39] *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.*

[40] *Heikkil¨a, M., Pietik¨ainen, M., Schmid, C.: Description of interest regions with center-symmetric local binary patterns. In: Computer Vision, Graphics and Image Processing, 5th Indian Conference. (2006) 58–69.*

[41] *Dalal, N.; and Triggs, B. Histograms of oriented gradients for human detection. Computer Vision and Pattern Recognition, pp.886-893, 2005.*

[42] *CARCAGNÌ, Pierluigi, DEL COCO, Marco, LEO, Marco, et al.Facial expression recognition and histograms of oriented gradients: a comprehensive study. SpringerPlus, 2015, vol. 4, no 1, p. 645.*

[43] *Bio-Inspired Credit Risk Analysis: Computational Intelligence with Support Vector Machine By Lean Yu, Shouyang Wang, Kin Keung Lai, Ligang Zhou.*

[44] *Belalia, M. (2016). Classification des images selon la sémantique, mémoire master , université de Mostaganem.*

[45] *Taibaoui, M., & Debbar, D. La découverte des concepts sémantiques cachés avec plusieurs niveaux d'abstraction pour la recherche d'images (Doctoral dissertation).*

[46] *Navlani, A., "KNN Classification using Scikit-learn", August 2, 2018.*

[47] *Dargan, S., Kumar, M., Ayyagari, M. R. & Kumar, G. A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning. Arch. Comput. Methods Eng. (2019). doi:10.1007/s11831-019-09344-w*

[48] *Lecun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature **521**, 436–444 (2015).*

[49] *Du, M., Liu, N. & Hu, X. Techniques for interpretable machine learning. Commun. ACM 63, 68–77 (2020).*

[50] *Bastien, "Comprendre le deep learning : Deep learning vs Machine learning", 13 Nov, 2019. Available : https://blog.clevy.io/nlp-et-ia/comprendre-le-deep-learning-3-3-deep-learning-vs-machine-learning/*

[51] *Bergach, M., "Expanding defense capabilities by applying deep learning techniques, Kontron america", jan 29th, 2018. Available : https://www.militaryaerospace.com/sponsored/kontronamerica/article/16728328/expanding-defense-capabilities-by-applying-deep-learning-techniques*

[52] *Les applications du Deep Learning, Antoine Krajnc, 21 Aug 2018*

[54] *Pierre-Alain, J. & Julien, K., Classification d'images : les réseaux de neurones convolutifs en toute simplicité, 25 Oct 2016. Available : https://blog.octo.com/classification-dimages-les-reseaux-de-neurones-convolutifs-en-toute-simplicite/.*

[55] *Vineethmohan, M., Hemanth, P., Mounica, M. & Lakshmi Prasanna, P. Image classification using deep learning. J. Adv. Res. Dyn. Control Syst.* ***12****, 1–10 (2020).*

[57] *Benjamin, P., "Traitement et analyse d'images, convolution", 2017. Available : https://perso.esiee.fr/~perretb/I5FM/TAI/convolution/index.html*

[58] *Benlahmar, Deeplearning,les réseaux de neurones convolutifs, 12 Oct 2018.*

[59] *A. Krizhevsky, I. Sutskever et G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks », Advances in neural Processing Systems de traitement. 2012.*

[60] *Lambert, R., Focus : Le Réseau de Neurones Convolutifs, 11 Jan 2019.*

[61] *Prabhu, R.," Understanding of Convolutional Neural Network (CNN) — Deep Learning", Mar 4, 2018. Available : https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148*

[62] *Beale, M. H., Hagan, M. T. & Demuth, H. B. Neural Network Toolbox ™ User ' s Guide.*

[63] *Wu, H. & Gu, X. Max-pooling dropout for regularization of convolutional neural networks. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* ***9489****, 46–54 (2015).*

[64] *Niranjan, K., Batch Normalization and Dropout in Neural Networks with Pytorch, Oct 21, 2019.*

[65] *TensorFlow for Deep Learning. Reza Bosagh Zadeh, Bharath Ramsundar. 2017.*

[66] *Gal, Y., Islam, R., Ghahramani, Z. "Deep Bayesian Active Learning with Image Data." ICML (2017)*

[67] *Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. (1998). "Gradient-based learning applied to document recognition". Proceedings of the IEEE. 86 (11): 2278–2324. doi:10.1109/5.726791 (https://doi.org/10.1109%2F5.726791).*

[68] *Golovko, V., Egor, M., Brich, A. & Sachenko, A. A shallow convolutional neural network for accurate handwritten digits classification. Commun. Comput. Inf. Sci.* ***673****, 77–85 (2017).*

[69] *Han, X., Zhong, Y., Cao, L. & Zhang, L. Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification. Remote Sens.* ***9****, (2017).*

[70] *Farag, M. S., Mohie El Din, M. M. & El Shenbary, H. A. Deep learning versus traditional methods for parking lots occupancy classification. Indones. J. Electr. Eng. Comput. Sci.* ***19****, 964–973 (2020).*

[71] *Muhammad, U., Wang, W., Chattha, S. P. & Ali, S. Pre-trained VGGNet Architecture for Remote-Sensing Image Scene Classification. Proc. - Int. Conf. Pattern Recognit. 2018-August, 1622–1627 (2018).*

[72] *Pandiyan, V. et al. In-process virtual verification of weld seam removal in robotic abrasive belt grinding process using deep learning. Robot. Comput. Integr. Manuf.* ***57****, 477–487 (2019).*

[73] *Sangeetha, V. & Prasad, K. J. R. Syntheses of novel derivatives of 2-acetylfuro[2,3-a]carbazoles, benzo[1,2-b]-1,4-thiazepino[2,3-a]carbazoles and 1-acetyloxycarbazole-2- carbaldehydes. Indian J. Chem. - Sect. B Org. Med. Chem.* ***45****, 1951–1954 (2006).*

[74] *Deng, L., Chu, H. H., Shi, P., Wang, W. & Kong, X. Region-based CNN method with deformable modules for visually classifying concrete cr. Appl. Sci.* ***10****, (2020).*

[75] *R.R., "Tanagra Spv Learning Curves, Didacticiel - Études de cas, Objectif. ", 1–23 (2011).*

[76] *Tarayoun, T. Méthodes de classification supervisée appliquées aux terres agricoles - Résumé. Journées Méthodologie Stat. (2018).*

[77] *Yan, H., Lu, J. & Zhou, X. Prototype-Based Discriminative Feature Learning for Kinship Verification. IEEE Trans. Cybern.* ***45****, 2535–2545 (2015).*

[78] *Liu, Q., Puthenputhussery, A. & Liu, C. Inheritable Fisher vector feature for kinship verification. 2015 IEEE 7th Int. Conf. Biometrics Theory, Appl. Syst. BTAS 2015 (2015). doi:10.1109/BTAS.2015.7358768*

[79] *A FEATURE SUBTRACTION METHOD FOR IMAGE BASED KINSHIP VERIFICATION UNDER UNCONTROLLED ENVIRONMENTS Xiaodong Duan , Zheng-Hua Tan. Int. Conf. Image Process. 1573–1577 (2015).*

[80] *Xu, M. & Shang, Y. Kinship verification using facial images by robust similarity learning. Math. Probl. Eng.* ***2016****, (2016).*

[81] *Alirezazadeh, P., Fathi, A. & Abdali-Mohammadi, F. Effect of Purposeful Feature Extraction in High-dimensional Kinship Verification Problem. J. Comput. Secur.* ***3****, 183–191 (2016).*

[82] *Lu, J., Hu, J. & Tan, Y. P. Discriminative Deep Metric Learning for Face and Kinship Verification. IEEE Trans. Image Process.* ***26****, 4269–4282 (2017).*

[83] *Yan, H., Zhou, X. & Ge, Y. Neighborhood repulsed correlation metric learning for kinship verification. 2015 Vis. Commun. Image Process. VCIP 2015* ***36****, 331–345 (2016).*

[84] *Yan, H., Lu, J., Deng, W. & Zhou, X. Discriminative multimetric learning for kinship verification. IEEE Trans. Inf. Forensics Secur.* ***9****, 1169–1178 (2014).*

[85] *Kou, L., Zhou, X., Xu, M. & Shang, Y. Learning a genetic measure for kinship verification using facial images. Math. Probl. Eng.* ***2015****, (2015).*

[86] *Zhou, X., Shang, Y., Yan, H. & Guo, G. Ensemble similarity learning for kinship verification from facial images in the wild. Inf. Fusion* ***32****, 40–48 (2016).*

[87] *Hu, J., Lu, J., Tan, Y. P., Yuan, J. & Zhou, J. Local Large-Margin Multi-Metric Learning for Face and Kinship Verification. IEEE Trans. Circuits Syst. Video Technol.* ***28****, 1875–1891 (2018).*

[88] *Zhou, X., Jin, K., Xu, M. & Guo, G. Learning deep compact similarity metric for kinship verification from face images. Inf. Fusion* ***48****, 84–94 (2019).*

[90] *H.Ouamane et M.Benatia « Identification de reconnaissance faciale avec des expressions » , Mémoire de Fin d'Etudes ,Université Mohamed Khider Biskra , le 07 Juin 2012.*

[91] *H. V. Nguyen and L. Bai, "Cosine similarity metric learning for face verification," in Asian conference on computer vision, 2010, pp. 709-720.*

[92] *Wu, Y. & Ji, Q. Facial Landmark Detection: A Literature Survey. Int. J. Comput. Vis.* ***127****, 115–142 (2019).*

❖ *Quelques sites Web:*

[8] *http://www.guide-genealogie.com/guide/parente.html*

[53] *https://medium.com/breathe-publication/top-15-deep-learning-applications-that-will-rule-the-world-in-2018-and-beyond-7c6130c43b01*

[56] *https://www.natural-solutions.eu/blog/la-reconnaissance-dimage-avec-les-rseaux-de-neurones-convolutifs*

[89] *http://solutions.journaldunet.com/0410/041005_biometriques.shtml*