

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE MINISTÈRE DE
L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ A. MIRA DE BÉJAÏA



Faculté de Technologies
Département d'Automatique, Télécommunication et d'Électronique (ATE)

Mémoire de Fin d'Étude

EN VUE DE L'OBTENTION DU DIPLÔME DE MASTER EN TÉLÉCOMMUNICATIONS

Option : Réseaux et Télécommunications

Thème

Vérification de la Parenté Faciale à l'Aide de Descripteurs Profonds et de Caractéristiques Texturales

Réalisé par :

M. ZENATI Yacine

M. ABDELLI Badri

Soutenu le 19/ 06/ 2025, Devant le jury composé de :

Président :	<i>Dr.</i> MEKHOUKH Abdenour	M.C.A	Université de Béjaïa.
Examinateur :	<i>Pr.</i> KASMI Reda	Professeur	Université de Béjaïa.
Encadrant :	<i>Dr.</i> TOUNSI Mohamed	M.C.B	Université de Béjaïa.
Co-Encadrant :	<i>Dr.</i> BELABBACI El Ouanas	Docteur	Université de Béjaïa.

Promotion : 2024/2025

Dédicace

Je dédie ce travail, fruit de plusieurs mois d'efforts, à toutes les personnes qui ont cru en moi et m'ont soutenu sans relâche.

À mes chers parents, véritables piliers de ma vie, premières sources de mon courage et de ma persévérance.

À mon père bien-aimé, à qui je dois la vie, mes valeurs, mes accomplissements, et tout mon respect.

À ma mère adorée, dont les sacrifices silencieux ont adouci mon chemin, qui a toujours répondu présente sans jamais se plaindre, et dont l'amour inépuisable a été ma plus grande force.

À mon frère et mes sœurs (**Djamel, Liticia, Liza**), merci pour votre présence, vos encouragements, votre affection sincère, et cette complicité qui m'a tant porté, surtout à ma chère sœur **Liza** qui à été d'une grande source de motivation et de force durant mon parcours

À ma chère tante **Naima**, dont l'aide précieuse, la générosité et le soutien m'ont profondément touché. Que ce travail soit aussi le tien. ainsi qu'à toute ma famille.

Et enfin, à tous mes amis fidèles, pour votre soutien, votre écoute, vos encouragements, et tous ces instants partagés qui ont illuminé ce parcours.

À mon binôme **ABDELLI Badri**, plus qu'un collègue, un véritable frère de projet. Merci pour ta présence, ton humour et ta détermination sans faille. Travailler à tes côtés a été un vrai plaisir.

À notre co-encadrant, **Dr. BELABACI El Ouanas**, nous vous remercions sincèrement pour votre aide, vos conseils, et tout ce que vous avez fait pour nous.

Que ce travail soit le reflet de ma reconnaissance et de mon attachement à chacun d'entre vous.

ZENATI Yacine

Dédicace

Je dédie ce travail à mes parents, dont l'amour, le soutien indéfectible et les valeurs transmises ont été les fondations de ma détermination. À mon père, pour sa sagesse, sa force, et l'exemple qu'il représente chaque jour. À ma mère, pour sa tendresse, ses sacrifices et son soutien constant, qui ont rendu chaque étape plus douce et chaque épreuve plus surmontable.

À mon frère Naïm, dont la présence, les encouragements et la bienveillance ont été un véritable moteur dans les moments de doute.

Merci d'avoir toujours été là.

À mes deux tantes, Aïda Manser et Samira Manser , que je considère comme mes deuxièmes mamans. Leur affection, leur générosité et leur présence réconfortante m'ont beaucoup apporté, bien au-delà des mots. À mon oncle Samir Abdelli, pour sa gentillesse, son attention sincère et son soutien constant.

À la mémoire de mon défunt oncle Nabil Abdelli, que je considère comme un grand frère. Son souvenir m'accompagne et continue d'inspirer mes pas.

À mes amis proches, pour leur écoute, leur confiance, et les instants de complicité qui ont apporté équilibre et lumière à ce chemin parfois exigeant.

À Yacine Zenati, mon binôme tout au long de ce projet, merci pour ta détermination, ton engagement, et l'ambiance fraternelle que tu as su instaurer.

Enfin, à notre co-encadrant, Dr. Belabaci El Ouanas, pour votre accompagnement, vos conseils avisés et votre disponibilité précieuse.

Votre encadrement a grandement contribué à la réalisation de ce travail.

Que ces quelques lignes témoignent de ma profonde gratitude envers chacun d'entre vous

ABDELLI Badri

Remerciements

Tout d'abord, nous remercions **ALLAH** du plus profond de notre coeur de nous avoir éclairés sur le droit chemin.

Nous tenons à remercier très sincèrement nos encadrants, **Dr. TOUNSI Mohamed** et **Dr. BELABBACI El Ouanas**, pour le temps précieux qu'ils nous ont consacré, leur aide et leurs conseils avisés. Leur disponibilité et leur patience tout au long de ce travail nous ont été d'une grande utilité.

Nous sommes profondément reconnaissants envers tous les membres du jury pour avoir accepté de consacrer leur temps à la lecture de ce travail et à son évaluation.

Nous tenons à remercier nos familles pour leur soutien affectif et moral inconditionnel ainsi que leurs encouragements constants. Enfin, on tiens à remercier chaleureusement nos amis, pour leur aide, leur écoute et les échanges enrichissants que nous avons partagés tout au long de cette aventure.

Résumé

La vérification de lien de parenté à partir d'images faciales suscite un intérêt croissant parmi les chercheurs et s'impose progressivement comme un enjeu majeur en vision par ordinateur. Grâce à l'analyse automatique des traits faciaux, il est possible de vérifier si deux individus sont apparentés. Ce type de vérification offre un large éventail d'applications, allant de la reconstruction d'arbres familiaux à l'organisation intelligente d'albums photos, l'annotation d'images, la localisation d'enfants disparus, ou encore l'appui aux enquêtes judiciaires.

Cet article présente un système performant de vérification de la parenté, qui utilise la méthode MTCNN à l'étape de prétraitement pour localiser les visages dans les images, en assurant une détection précise. En outre, nous proposons deux descripteurs texturaux (LPQ, BSIF). Nous explorons également la complémentarité avec les caractéristiques profondes (InceptionV3), en les fusionnant au niveau des scores via la méthode de régression logistique.

Des expérimentations approfondies menées sur deux bases de données de parenté ont permis d'atteindre des précisions de vérification de 93.75 % et 89.46 % respectivement sur les ensembles de données Cornell KinFace et UBKinFace.

Mots clés : Vérification de la Parenté, MTCNN, Caractéristiques Profondes, Caractéristiques Texturales, TXQEDA+WCCN, Fusion LR.

Table des matières

Liste des Figures	iii
Liste des Tableaux	iv
Liste des acronymes	v
Introduction Générale	1
I Vérification de la parenté faciale : état de l'art	4
I.1 Introduction	5
I.2 Vérification de la parenté en vision par ordinateur	5
I.3 Catégories de relations de parenté	5
I.4 Application	6
I.5 Défis	6
I.5.1 Défis unique à la vérification de la parenté	7
I.5.2 Défis commun avec la reconnaissance faciale	7
I.6 Vérification faciale et vérification de la parenté	10
I.7 Système général de la vérification de la parenté	10
I.8 Approches classiques et modernes pour la vérification de la parenté faciale	11
I.8.1 Méthodes d'apprentissage automatique (Machine Learning) . . .	11
I.8.2 Méthodes d'apprentissage profond (Deep Learning)	12
I.9 Conclusion	13
II Méthode Pour La Vérification De La Parenté	14
II.1 Introduction	15
II.2 Travaux Connexes	15
II.3 Bases de Données de vérification de la parenté	16
II.4 Approches pour la vérification de la parenté	19
II.4.1 Prétraitement du visage	19
II.4.1.1 Détection du visage (avec MTCNN)	19
II.4.2 Extraction de Caractéristiques	20
II.4.2.1 Caractéristiques profondes basées sur un réseau de neurones convolutifs (CNN)	20
II.4.2.1.1 Architecture CNN	21
II.4.2.1.2 Modèle pré-entraîné	26
II.4.2.2 Descripteurs texturaux	27

Table des matières

II.4.3 Classification	30
II.4.3.1 Réduction de dimension par TXQEDA	30
II.4.3.2 Normalisation par la Covariance Intra-classe (WCCN) . .	30
II.4.3.3 Similarité cosinus	31
II.4.4 Décision	32
II.5 Architecture générale du système	33
II.6 Conclusion	33
III Éxpérimentations et résultats	34
III.1 Introduction	35
III.2 Environnement de travail	35
III.3 Protocole de travail	35
III.4 Expérimentations et Résultats	36
III.4.1 Expérimentations sur la base de données Cornell kinFace	36
III.4.2 Expérimentations sur la base de données de UB KinFace	38
III.5 Synthèse des performances obtenues	39
III.6 Comparaison avec les travaux connexes	40
III.7 Conclusion	40
Conclusion Générale	41
Perspectives	43
Bibliographie	44
Résumer	47

Table des figures

I.1	Éxemples des défis uniques : a) À travers l'âge Père-Fils, b) À travers le genre Frère-Sœur, c) À travers l'âge et le sexe Père-Fille.	7
I.2	Éxemples des variations de pose [1]	8
I.3	Éxemples d'éléments d'occlusions [2]	8
I.4	Éxemples changement d'expression facial [3]	8
I.5	Éxemple du vieillissement	9
I.6	Éxemples des conditions d'éclairage variables [4]	9
I.7	Éxemples des différentes résolutions [3]	9
I.8	Architecture générale du système de vérification de parenté	11
I.9	Architecture d'un réseau de neurones.	13
II.1	Éxemples d'images extraites de bases de données utilisées, comprenant pour chaque base deux échantillons visuels : un représentant un cas positif et un représentant un cas négatif [5]	18
II.2	Vue d'ensemble des résultats de MTCNN [6]	20
II.3	Éxemple de l'architecture CNN [7]	21
II.4	Éxemple des calculs fondamentaux réalisés à chaque passage dans une couche convolutionnelle.	22
II.5	Éxemple des calculs réalisés après un pas Stride ($S = 1$)	23
II.6	Éxemple du pas (Stride) pour $S = 2$	23
II.7	Éxemple du zero-padding	23
II.8	Éxemple de la Couche d'activation ReLu	24
II.9	Éxemple des deux variant max pooling et average pooling	25
II.10	Éxemple de la couche d'aplatissement (Flatten layer)	25
II.11	Éxemple de la couche entièrement connecté (fully connected layer)	26
II.12	Exemple de l'architecture du modèle InceptionV3	27
II.13	Exemple de la Quantification de phases locales (LPQ) avec différents rayons $R = 1$ à 10	28
II.14	Éxemple des étapes nécessaires à la génération du vecteur de caractéristiques par la méthode LPQ (Local Phase Quantization),	28
II.15	Exemple de Caractéristiques BSIF avec différents filtres de 3×3 à 15×15	29
II.16	Exemple de WCCN	31
II.17	Éxemple de Similarité cosinus	32
II.18	Schéma du système proposé de vérification de la parenté faciale	33
III.1	Éxemple de la validation croisée K-fold, avec $K=5$	36

Liste des tableaux

I.1	Types de relations de parenté utilisées dans les systèmes de vérification faciale	6
I.2	Comparaison entre la vérification de parenté et la vérification faciale	10
II.1	Caractéristiques des bases de données de visages utilisées	18
III.1	Précision moyenne (%) des descripteurs texturaux LPQ et BSIF sur la base de données Cornell kinFace	37
III.2	La précision moyenne (%) des meilleures scores de chaque descripteur sur la base de données Cornell kinFace	37
III.3	Résultat final après la fusion des meilleurs résultats de chaque descripteur (Texturaux et profond) avec la régression logistique (LR)	38
III.4	Précision moyenne (%) des descripteurs texturaux LPQ et BSIF sur la base de données UB KinFace	38
III.5	La précision moyenne (%) des meilleures scores de chaque descripteur sur la base de données UB KinFace	39
III.6	Résultat final après la fusion des meilleurs scores de chaque descripteur (Textural et profond) avec la régression logistique (LR)	39
III.7	Comparaison des performances des méthodes de vérification de parenté sur les bases de données Cornell et UB	40
tocchapter	Liste des acronymes	

Liste des acronymes

2DSTFT	<i>Two-Dimensional Short-Time Fourier Transform</i>
2DWFT	<i>Two-Dimensional Wavelet Fourier Transform</i>
ADN	<i>Acide DésoxyriboNucléique</i>
ANN	<i>Artificial Neural Network</i>
AUC	<i>Area Under the Curve</i>
BB	<i>Brother Brother</i>
BS	<i>Brother Sister</i>
BSIF	<i>Binarized Statistical Image Features</i>
CNN	<i>Convolutional Neural Network</i>
FD	<i>Father Daughter</i>
FIW	<i>Families In the Wild</i>
FS	<i>Father Son</i>
GAP	<i>Global Average Pooling</i>
GFGD	<i>Grand Father Grand Daughter</i>
GFGS	<i>Grand Father Grand Son</i>
GMGD	<i>Grand Mother Grand Daughter</i>
GMGS	<i>Grand Mother Grand Son</i>
ICA	<i>Independant Component Analysis</i>
IR	<i>Infrared</i>
KIN	<i>Kinship</i>
KNN	<i>K-Nearest Neighbors</i>
LBP	<i>Local Binary Pattern</i>

Liste des acronymes

LPQ	<i>Local Phase Quantization</i>
LR	<i>Logistic Regression</i>
MD	<i>Mother Daughter</i>
MS	<i>Mother Son</i>
MTCNN	<i>Multi-task Cascaded Convolutional Neural Network</i>
NMS	<i>Non-Maximum Suppression</i>
NON-KIN	<i>Non-Kinship</i>
O-NET	<i>Output Network</i>
P-NET	<i>Proposal Network</i>
R-NET	<i>Refine Network</i>
ReLU	<i>Rectified Linear Unit</i>
ROC	<i>Receiver Operating Characteristic</i>
SIFT	<i>Scale-Invariant Feature Transform</i>
SS	<i>Sister Sister</i>
SVM	<i>Support Vector Machine</i>
TXQEDA	<i>Tensor Cross-view Quadratic Exponential Discriminant Analysis</i>
VGG	<i>Visual Geometry Group</i>
WCCN	<i>Within-Class Covariance Normalization</i>
XQDA	<i>Cross-view Quadratic Discriminant Analysis</i>
XQEDA	<i>Cross-view Quadratic Exponential Discriminant Analysis</i>

Introduction Générale

Contexte et Motivation

La vérification de la parenté par analyse d'images faciales est un domaine de recherche en pleine expansion dans le champ de la vision par ordinateur. Ce champ s'inscrit dans la continuité des avancées réalisées en reconnaissance faciale, en particulier avec l'émergence des réseaux neuronaux profonds, qui ont radicalement transformé la capacité des machines à interpréter des données visuelles complexes. Introduite de manière formelle à partir de 2010[8], l'analyse de parenté faciale a depuis évolué pour devenir un sujet central, mobilisant à la fois des concepts issus de l'apprentissage profond, de l'algèbre multilinéaire (notamment l'utilisation de tenseurs), ainsi que des techniques avancées de traitement et d'analyse d'images[9].

Cette technologie vise à déterminer automatiquement l'existence de liens familiaux entre deux individus à partir de leurs traits faciaux. L'intérêt grandissant pour cette approche réside dans son potentiel à offrir une alternative rapide, économique et non-invasive aux méthodes traditionnelles d'identification génétique, comme l'analyse ADN.

Les cas d'usage sont nombreux et variés. Dans un cadre personnel, elle peut faciliter la gestion intelligente de photos ou la recherche généalogique. Sur le plan institutionnel, ses applications s'étendent à des domaines sensibles tels que la sécurité nationale, les enquêtes criminelles, le contrôle aux frontières ou encore l'identification de victimes de catastrophes. Elle peut également jouer un rôle crucial dans le cadre de procédures humanitaires, telles que la réunification familiale pour les enfants réfugiés ou disparus. Dans toutes ces situations, la capacité à effectuer des vérifications fiables en temps réel est un enjeu déterminant.

Problématique

Malgré les avancées technologiques récentes, la mise en oeuvre de systèmes automatisés de vérification de parenté faciale se heurte encore à des défis majeurs. La complexité du problème réside dans la nature hautement variable et multidimensionnelle de l'apparence humaine. Là où l'oeil humain est capable de percevoir intuitivement des ressemblances familiales, les systèmes d'intelligence artificielle doivent apprendre à extraire et interpréter des indices subtils, souvent masqués par d'importantes variations visuelles.

Parmi les difficultés les plus notables, on distingue :

- **Les variations intra-sujets** : changements d'angle de prise de vue, expressions faciales, conditions d'éclairage, ou effets liés au vieillissement.

- **Les différences inter-sujets** : écarts d'âge importants (comme entre parents et enfants), diversité ethnique, ou encore variations morphologiques naturelles.

Dans divers contextes sensibles tels que l'organisation d'albums photo, la recherche de liens familiaux dans des bases de données publiques, l'analyse de demandes d'asile, ou encore les enquêtes judiciaires, il est souvent nécessaire d'établir des relations de parenté entre individus. Actuellement, la méthode la plus fiable pour vérifier une relation biologique demeure l'analyse ADN. Toutefois, cette solution, bien que précise, présente des inconvénients majeurs : elle est coûteuse, chronophage et inadaptée aux situations nécessitant une réponse rapide, notamment lorsque les individus concernés ne sont pas coopératifs.

Ces facteurs introduisent une forte variabilité dans les données, rendant l'extraction de caractéristiques discriminantes à la fois stables et pertinentes particulièrement complexe. La problématique centrale de cette recherche est donc de concevoir des algorithmes robustes, capables de modéliser efficacement les traits héréditaires tout en résistant aux perturbations visuelles. L'enjeu est d'atteindre un haut niveau de fiabilité, indispensable pour que ces systèmes puissent être intégrés dans des contextes d'utilisation concrets et critiques.

Contribution

Ce travail de fin d'études s'inscrit dans la problématique de la vérification de la parenté basée sur des images faciales. Il vise à proposer une approche innovante permettant d'améliorer la précision et la robustesse des systèmes existants dans ce domaine en combinant des techniques d'apprentissage automatique et des outils mathématiques avancés.

La contribution principale de ce projet consiste à concevoir, implémenter et évaluer un modèle de vérification de parenté capable de capturer efficacement les traits héréditaires communs, tout en étant robuste face aux variations intra- et inter-individuelles (pose, âge, expressions, etc.).

Les contributions spécifiques de ce travail sont les suivantes :

- **Prétraitement des données** : Nous avons utilisé l'algorithme MTCNN (Multi-task Cascaded Convolutional Networks) pour la détection et le recadrage automatique des visages, assurant ainsi une normalisation efficace des images d'entrée.
- **Extraction de caractéristiques** : Deux descripteurs d'images à texture locale ont été évalués : Local Phase Quantization (LPQ) et Binarized Statistical Image Features (BSIF). Plusieurs échelles ont été testées pour chacun des descripteurs, et la meilleure configuration a été sélectionnée selon les performances obtenues.
- **Méthode proposée** : Nous introduisons une nouvelle méthode de classification pour la vérification de parenté basée sur une approche discriminante tensorielle nommée Tensor Cross-view Quadratic Exponential Discriminant Analysis (TXQEDA). Cette méthode permet à la fois la réduction de dimensionnalité et l'optimisation des performances de classification.
- **Fusion de scores** : Pour améliorer davantage la précision, nous avons mis en oeuvre une fusion au niveau des scores en utilisant la régression logistique (Logistic Regression). Cette fusion combine les caractéristiques dites "shallow" (LPQ et BSIF) avec des descripteurs profonds extraits à l'aide du modèle pré-entraîné InceptionV3. Cette complémentarité entre descripteurs permet d'obtenir de meilleures performances globales.

- **Évaluation expérimentale :** Le modèle a été entraîné et testé sur des bases de données publiques de visages familiaux. Les performances ont été évaluées à l'aide de métrique standard la précision moyenne, afin de valider son efficacité dans des scénarios applicatifs concrets.

Organisation du mémoire

Ce mémoire présente une introduction générale dans laquelle nous abordons la problématique de notre étude sur la vérification de la parenté faciale à l'aide de descripteurs profonds et de caractéristiques texturales.

Notre mémoire est strucuré en trois chapitres comme suit :

- **Chapitre 1:** Ce premier chapitre est consacré à l'état de l'art. Il présente une vue d'ensemble sur la vérification de la parenté faciale, en introduisant le concept, ses applications, les différents types de relations de parenté, ainsi que les défis rencontrés. Il expose également les approches classiques et modernes en la matière, notamment les techniques d'apprentissage automatique et d'apprentissage profond.
- **Chapitre 2:** Ce deuxième chapitre décrit la méthodologie proposée. Il détaille les différentes étapes du système conçu, à savoir le prétraitement des images, l'extraction des caractéristiques à l'aide de descripteurs texturaux (comme LPQ et BSIF) et de modèles profonds (comme InceptionV3), la réduction de dimension (TXQEDA), la normalisation (WCCN), ainsi que la mesure de similarité et la prise de décision.
- **Chapitre 3:** Ce troisième chapitre est dédié aux expérimentations. Il présente l'environnement de travail, les bases de données utilisées, le protocole expérimental ainsi que les résultats obtenus. Une discussion comparative avec les travaux existants est également menée afin d'évaluer les performances de la méthode proposée.

Chapitre

I

Vérification de la parenté faciale : état de l'art

I.1 Introduction

La vérification de la parenté faciale est une tâche spécifique de la vision par ordinateur, visant à déterminer si deux individus présentent un lien biologique à partir de leurs images faciales [10]. C'est une thématique de recherche en constante évolution vu son intérêt croissant dans diverses applications.

Dans ce chapitre, nous présentons un état de l'art détaillé sur cette problématique. Nous commençons par introduire les concepts fondamentaux liés à la vérification de parenté dans le cadre plus large de la reconnaissance faciale. Nous donnerons une classification des différents types de relations de parenté ainsi qu'une analyse des principaux défis liés à cette tâche, notamment les variations morphologiques et les contraintes contextuelles.

Nous y soulignerons également les différences fondamentales entre la vérification d'identité faciale classique et la vérification de parenté, souvent confondues à tort. Enfin, nous proposerons un panorama des approches existantes, qu'elles soient traditionnelles ou basées sur des méthodes récentes d'intelligence artificielle, en vue de situer notre propre contribution.

I.2 Vérification de la parenté en vision par ordinateur

La vérification de la parenté est un domaine émergent de la vision par ordinateur, qui vise à déterminer si deux individus ont un lien familial à partir de leurs traits faciaux uniquement [11]. Contrairement à d'autres modalités biométriques comme la voix, l'iris ou les empreintes digitales, cette approche repose sur l'exploitation des caractéristiques physiologiques héritées, visibles dans le visage. Inspirée par l'idée que les traits du visage, tout comme l'ADN, se transmettent génétiquement, cette méthode permet une vérification non invasive, réalisable à distance et sans contact physique. Elle est ainsi bien adaptée aux applications en sécurité, en identification ou en regroupement familial, tout en étant rapide, fiable et peu coûteuse. Malgré ses avantages, ce domaine reste encore en développement. Il doit faire face à des défis majeurs liés aux variations d'âge, de pose, de luminosité et d'expression, qui compliquent la reconnaissance précise des similarités.

En pratique, un tel système prend en entrée deux images de visages et retourne un verdict : apparentés (KIN) ou non apparentés (NON-KIN), grâce à une combinaison de prétraitement, d'extraction de caractéristiques et de classification.

I.3 Catégories de relations de parenté

Les liens de parenté entre deux individus varient en fonction de leur degré de relation biologique. Dans le domaine de la vérification de la parenté faciale, ces liens sont souvent regroupés en trois catégories principales, couvrant un total de 11 types distincts de relations, comme présenté dans le tableau I.1 ci-dessous.

Type de relation	Exemples
Parent - Enfant	Père - Fille (F-D) Père - Fils (F-S) Mère - Fille (M-D) Mère - Fils (M-S)
Frères - Sœurs	Frère - Frère (B-B) Sœur - Sœur (S-S) Frère - Sœur (B-S)
Grands-parents - Petits-enfants	Grand-père - Petit-fils (GF-GS) Grand-père - Petite-fille (GF-GD) Grand-mère - Petit-fils (GM-GS) Grand-mère - Petite-fille (GM-GD)

TAB. I.1 – Types de relations de parenté utilisées dans les systèmes de vérification faciale

I.4 Application

La reconnaissance automatique des liens de parenté à partir d’images faciales suscite un fort engouement dans la communauté scientifique. Elle mobilise des expertises variées issues de disciplines telles que la psychologie, la biométrie, l’intelligence artificielle ou encore la vision par ordinateur. Ensemble, ces chercheurs cherchent à concevoir des systèmes capables de déterminer si deux visages appartiennent à des individus biologiquement liés. Mais les usages d’une telle technologie dépassent largement ce seul objectif. Elle pourrait en effet répondre à de nombreux besoins concrets dans différents secteurs [12] , parmi lesquels on trouve :

- Le soutien aux opérations de recherche d’enfants portés disparus, en facilitant leur identification via des bases de données visuelles.
- Le renforcement des contrôles dans les aéroports afin de lutter contre le trafic d’enfants.
- L’assistance aux autorités de l’immigration et aux services de sécurité frontalière pour vérifier certains liens familiaux revendiqués.
- L’exploitation dans des recherches historiques ou en généalogie pour établir des filiations à partir d’archives photographiques.
- Le tri et l’analyse intelligente des milliers de photos publiées quotidiennement sur les plateformes sociales.
- L’optimisation de la gestion des albums photo numériques grâce à la détection automatique des membres d’une même famille.

I.5 Défis

La vérification de la parenté faciale présente des défis particuliers qui la distinguent des tâches classiques de reconnaissance faciale. Ces défis peuvent être regroupés en deux grandes catégories [13] :

I.5.1 Défis unique à la vérification de la parenté

La vérification de parenté faciale est une tâche complexe qui diffère de la reconnaissance faciale traditionnelle, car elle ne cherche pas à identifier une même personne, mais plutôt à détecter des similitudes faciales entre deux individus liés biologiquement. Ces ressemblances proviennent de l'héritage génétique, chaque individu recevant des traits distincts de ses ancêtres. Contrairement à la reconnaissance faciale, où les caractéristiques sont propres à une seule personne, la vérification de parenté nécessite l'extraction de traits communs partagés entre différentes personnes. Cette tâche soulève trois défis majeurs spécifiques, dont l'analyse permet de mieux cerner la complexité du problème et d'orienter la conception de systèmes de vérification automatisés [14].

- **Influence de l'âge et du sexe :** Les traits familiaux sont mesurés entre des individus de sexes et d'âges différents (par exemple, père-fils, père-fille, frère-soeur), ce qui engendre une grande variabilité morphologique.
- **Héritage différencié selon les paires :** Chaque paire familiale présente des propriétés spécifiques, car les enfants héritent différemment des traits de leurs parents, même lorsqu'ils sont du même sexe.
- **Caractère stochastique de la parenté :** La parenté est fondée sur une combinaison aléatoire (stochastique) de traits familiaux, ce qui nécessite une modélisation probabiliste plutôt que déterministe ; ces traits servent de base à la mesure de la parenté, généralement abordée comme une tâche de vérification faciale automatique.



FIG. I.1 – Exemples des défis uniques : a) À travers l'âge Père-Fils, b) À travers le genre Frère-Sœur, c) À travers l'âge et le sexe Père-Fille.

I.5.2 Défis commun avec la reconnaissance faciale

La vérification de la parenté faciale, bien qu'étant un sous-ensemble de la vérification faciale automatique, présente des défis uniques liés à la sensibilité des visages apparentés aux variations d'expressions faciales, d'occlusions et de poses. De plus, des facteurs externes tels que l'éclairage, le flou et la faible résolution des images ont également un impact considérable sur l'apparence des visages. Ces éléments peuvent altérer les traits familiaux de manière différente, rendant la tâche de vérification encore plus complexe, et soulignant ainsi les défis communs rencontrés dans les deux domaines.

- **Variations de pose** : Les changements d'orientation du visage, que ce soit sous différents angles ou inclinaisons, peuvent perturber l'identification et la comparaison des traits faciaux [15].



FIG. I.2 – Exemples des variations de pose [1]

- **Présence/absence d'éléments structurants ou occlusions** : La présence de lunettes, de coiffures, de mains ou d'autres objets qui couvrent partiellement le visage peut gêner l'extraction complète des caractéristiques faciales et compliquer la vérification de la parenté.

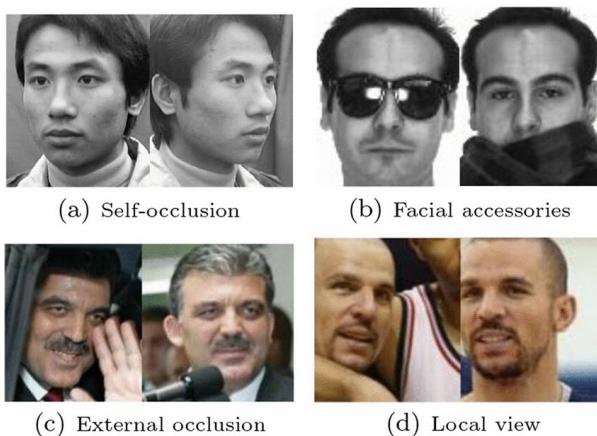


FIG. I.3 – Exemples d'éléments d'occlusions [2]

- **Changements d'expressions faciales** : Les expressions émotionnelles (sourire, froncement de sourcils, etc.) modifient la configuration des traits faciaux, ce qui rend plus difficile la comparaison précise entre deux visages [3].



FIG. I.4 – Exemples changement d'expression facial [3]

- **Vieillissement du visage** : L'évolution des traits au fil du temps (rides, modification de la peau, etc.) change l'apparence des individus, ce qui peut rendre plus difficile la détection de similitudes entre des membres de la même famille à différents âges .



FIG. I.5 – Exemple du vieillissement

- **Conditions d'éclairage variables** : Les variations d'éclairage peuvent provoquer des ombres et des réflexions sur le visage, ce qui peut altérer l'apparence des traits et rendre les comparaisons plus complexes.

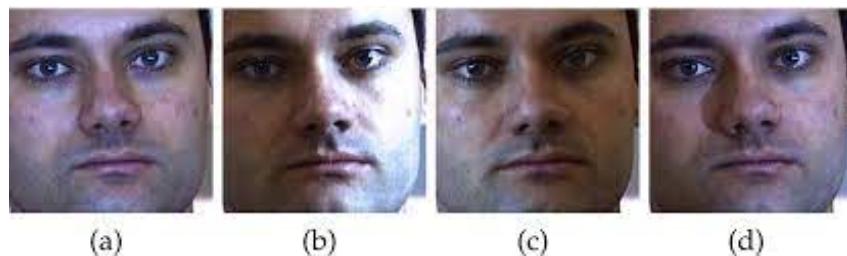


FIG. I.6 – Exemples des conditions d'éclairage variables [4]

- **Résolution d'image et modalité** : Les images de faible résolution ou provenant de différentes modalités (par exemple, IR, visible, etc.) peuvent perdre des détails cruciaux des traits du visage, affectant la précision des mesures de similarité [3].



FIG. I.7 – Exemples des différentes résolutions [3]

I.6 Vérification faciale et vérification de la parenté

La vérification de parenté faciale peut, à première vue, être confondue avec la vérification faciale classique. Bien que ces deux systèmes présentent certaines similarités, ils sont en réalité distincts. On pourrait même considérer que la vérification de parenté représente une forme avancée de la vérification faciale classique [16].

Les deux approches partagent une structure de base similaire, mais diffèrent sur plusieurs aspects, comme le montre le tableau I.2 .

Vérification de parenté	Vérification faciale
<ul style="list-style-type: none"> • Extraction des caractéristiques de deux personnes différentes • Vérification du lien de parenté • Traits différents entre l'image requête 1 et l'image requête 2 • Système de plus haut niveau • À l'étape de décision : Apparenté ou non apparenté • Précision autour de 90 % 	<ul style="list-style-type: none"> • Extraction des caractéristiques de la même personne • Vérification ou identification • Même trait entre l'image requête 1 et l'image requête 2 • Système de bas niveau • À l'étape de décision : Correspondant ou non correspondant • La performance de la machine est globalement aussi bonne que celle de l'humain

TAB. I.2 – Comparaison entre la vérification de parenté et la vérification faciale

I.7 Système général de la vérification de la parenté

En raison de la complexité et des défis liés à la vérification de parenté, il est nécessaire de mettre en place un cadre structuré pour organiser les différentes méthodes existantes. Ce cadre se compose de quatre étapes principales : le prétraitement, l'extraction des caractéristiques, classification et la phase de décision comme le montre la figure I.8

- **Prétraitement** : Cette étape consiste à détecter les visages dans les images, les recadrer, améliorer leur qualité (par exemple, en appliquant des techniques d'amélioration de contraste ou de normalisation) et les mettre à une même taille et alignement [17]. L'objectif est de standardiser les images pour faciliter les étapes suivantes.
- **Extraction des caractristique** : Ici, on extrait des descripteurs visuels qui représentent les traits du visage. Ces descripteurs peuvent être texturaux (comme LPQ ou BSIF) ou profonds (issus de réseaux de neurones convolutifs comme InceptionV3). Cette représentation numérique permet de comparer les visages à un niveau plus abstrait [14].
- **Classification** : Une fois les caractéristiques extraites, on calcule un score de similité entre deux représentations faciales. Des métriques comme la distance euclidienne ou la distance cosinus sont utilisées pour quantifier la ressemblance [18].

L'idée est que les visages de personnes apparentées devraient avoir un score de similarité plus élevé.

- **Décision :** Enfin, une décision est prise à partir du score de similarité : si celui-ci dépasse un certain seuil, la relation de parenté est confirmée, sinon elle est rejetée. Des classificateurs comme la régression logistique [19] peuvent être utilisés pour renforcer cette étape décisionnelle.

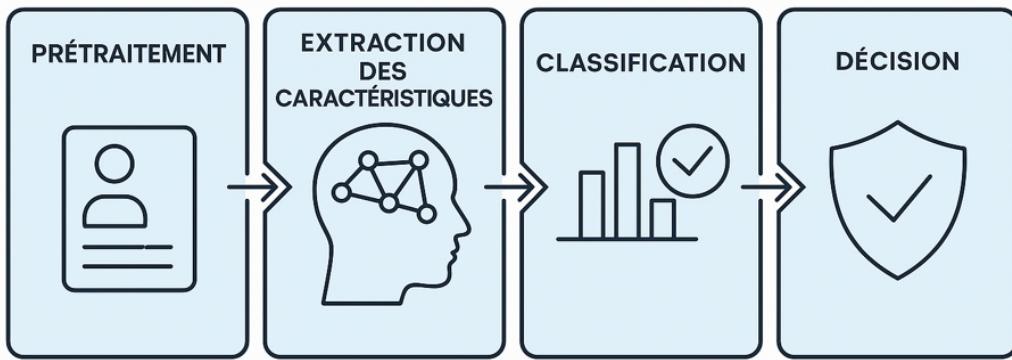


FIG. I.8 – Architecture générale du système de vérification de parenté

I.8 Approches classiques et modernes pour la vérification de la parenté faciale

Dans un monde où la reconnaissance faciale est devenue un outil omniprésent, l'idée d'aller plus loin – jusqu'à détecter des liens de parenté à partir des traits du visage – pose un défi à la fois passionnant et complexe. Contrairement à l'identification classique d'un individu, la vérification de parenté cherche à mettre en lumière des ressemblances héritées, souvent subtiles, parfois floues, entre deux personnes. Cette tâche mobilise des approches variées qui s'étendent du traitement d'images traditionnel aux modèles profonds inspirés de la biologie humaine[16], [5], [17].

Au fil des années, la littérature scientifique s'est enrichie de multiples stratégies visant à formaliser cette tâche. Certaines s'appuient sur des caractéristiques visuelles explicites, extraites de manière manuelle, tandis que d'autres misent sur la puissance des réseaux neuronaux pour apprendre des représentations latentes, plus abstraites. Ce chapitre passe en revue les approches les plus marquantes, classées en deux grandes familles : les méthodes d'apprentissage automatique classiques, et les approches modernes fondées sur l'apprentissage profond.

I.8.1 Méthodes d'apprentissage automatique (Machine Learning)

Les premières approches de la vérification de parenté ont exploité des descripteurs texturals, tels que LPQ (local phase quantization), LBP (Local Binary Patterns), BSIF (Binarized Statisitcal Image Features), pour extraire des caractéristiques texturales du visage.

Une fois les descripteurs extraits, des classificateurs supervisés comme SVM (Support Vector Machines), Random Forests, ou K-NN (K-Nearest Neighbors) sont utilisés pour apprendre un modèle de similarité entre les paires d'images. Certains travaux ont aussi recours à des apprentissages métriques (metric learning) pour estimer la distance dans un espace projeté, où les visages apparentés sont plus proches les uns des autres que ceux non apparentés.

Ces méthodes présentent toutefois des limitations : elles sont souvent sensibles à la pose, à l'illumination et aux expressions faciales, et leur capacité à généraliser des relations parentales complexes (parent-enfant, frère-soeur) reste limitée.

I.8.2 Méthodes d'apprentissage profond (Deep Learning)

L'apprentissage profond (Deep Learning), introduit par Hinton et Salakhutdinov (2006), désigne une famille de méthodes d'apprentissage automatique s'appuyant sur des réseaux de neurones profonds (Deep Neural Networks). Ces réseaux sont constitués de multiples couches de transformations non linéaires permettant de capturer des représentations complexes des données.

Les réseaux de neurones artificiels (ANN), inspirés du cerveau humain, traitent des données d'entrée (comme des images faciales) et fournissent des sorties pertinentes, telles que des scores de similarité ou des classifications. Leur efficacité découle du théorème d'approximation universelle démontré par Cybenko (1989) et Hornik (1991), ainsi que de l'algorithme de rétropropagation du gradient, proposé par LeCun (1986), permettant l'apprentissage efficace des poids du réseau.

Dans le cadre de la vérification de la parenté faciale, ces méthodes ont permis de générer des représentations profondes, appelées descripteurs profonds, capables de capturer les similarités génétiques subtiles entre visages. Ils offrent une robustesse accrue face aux variations d'âge, de pose ou d'éclairage.

L'entraînement de ces réseaux nécessite des bases de données annotées. Il consiste à ajuster progressivement les paramètres internes du réseau afin de minimiser une fonction de perte mesurant l'écart entre la prédiction du modèle et la vérité terrain.

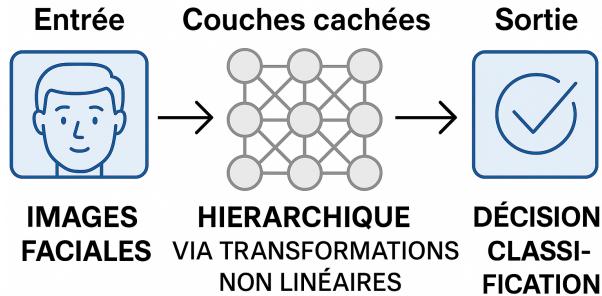


FIG. I.9 – Architecture d'un réseau de neurones.

I.9 Conclusion

Dans ce chapitre, nous avons présenté les concepts fondamentaux liés à la vérification de la parenté faciale. Nous avons abordé les différentes catégories de relations de parenté, leurs domaines d'application ainsi que les défis spécifiques et communs avec la reconnaissance faciale. Nous avons également exploré les approches classiques et modernes utilisées dans ce domaine.

Dans le chapitre suivant, nous détaillerons la méthode proposée dans ce travail, en présentant les étapes de prétraitement, d'extraction de caractéristiques, de réduction de dimension et de classification.

Chapitre **II**

Méthode Pour La Vérification De La Parenté

II.1 Introduction

La vérification de la parenté est un domaine de recherche émergent qui a suscité un vif intérêt ces dernières années. Dans ce chapitre, nous allons présenter un résumé des bases de données les plus connues utilisées dans ce domaine. Nous allons également exposer les travaux connexes et fournir un aperçu global des différentes approches de vérification de la parenté. Ces méthodes ont été regroupées en deux grandes catégories : celles reposant sur des caractéristiques texturales (*Shallow feature*) et celles basées sur des descripteurs profonds (*deep features*), afin de mieux cerner l'état de l'art dans ce domaine.

II.2 Travaux Connexes

La vérification de la parenté faciale a émergé comme un champ de recherche spécifique au début des années 2010 [8], bien après l'émergence de la reconnaissance faciale classique. Les premiers travaux notables dans ce domaine ont été motivés par des défis uniques, tels que la difficulté de détecter des ressemblances entre des visages de personnes apparentées malgré des écarts d'âge, de sexe, d'expression ou d'angle de vue. La publication de bases de données dédiées comme KinFaceW-I et KinFaceW-II en 2011, a marqué un tournant important, en fournissant une référence commune pour l'évaluation des algorithmes.

Les premières approches ont utilisé des descripteurs façonnés à la main tels que LBP (Local Binary Patterns), HOG ou SIFT, associés à des classificateurs classiques comme les SVM. Ces méthodes ont permis de poser les premières bases méthodologiques pour la tâche de parenté faciale [20].

Avec le succès croissant des modèles d'apprentissage profond dans la reconnaissance faciale, les chercheurs ont commencé à transférer ces architectures (comme VGG-Face, FaceNet ou InceptionV3) vers la vérification de la parenté. Cette transition a permis d'extraire automatiquement des représentations profondes et discriminantes, adaptées à la reconnaissance de liens familiaux dans des contextes complexes [21], [22].

Par ailleurs, des architectures spécifiques telles que les réseaux siamois et les réseaux triplet, ont été introduites. Elles permettent d'optimiser la capacité des modèles à apprendre des fonctions de similarité entre paires d'images, en s'appuyant sur des fonctions de perte comme la "contrastive loss" ou la "triplet loss".

Afin d'améliorer davantage la robustesse des systèmes, des études ont proposé de combiner les descripteurs profonds avec des caractéristiques texturales ou morphologiques[10]. Cette approche hybride a permis de tirer parti des avantages des méthodes profondes et des caractéristiques façonnées à la main, notamment dans des cas de faible quantité de données.

La mise à disposition de nouvelles bases de données plus complètes, comme FIM (*Families In the Wild*)[23], a significativement contribué au développement du domaine. Cette base offre une grande variété de relations (parent-enfant, fratrie, etc.) et une diversité d'images en conditions réelles, facilitant la généralisation des approches.

Enfin, face au défi persistant du manque de données annotées, des approches récentes s'orientent vers l'apprentissage faiblement supervisé ou auto-supervisé, permettant l'entraînement de modèles performants à partir de données partiellement étiquetées, voire non annotées.

II.3 Bases de Données de vérification de la parenté

La reconnaissance des liens de parenté à partir d'images faciales suscite un intérêt croissant, notamment en raison de ses nombreuses applications abordées précédemment. Cette dynamique a conduit à la création de plusieurs bases de données dédiées, reflétant différentes relations familiales. Le visage humain est un objet d'analyse complexe, dont l'apparence peut être influencée par divers facteurs tels que l'orientation, l'éclairage, l'âge, les expressions, les accessoires, ou encore les caractéristiques propres à chaque individu.

Pour permettre le développement d'algorithmes capables de faire face à ces variabilités, il est crucial de disposer des bases de données suffisamment vastes et diversifiées. Ces ressources permettent non seulement de concevoir des modèles robustes, mais aussi d'évaluer et de comparer les performances des approches existantes dans des conditions contrôlées. De plus, l'évolution du domaine a favorisé l'introduction de bases couvrant des types de relations plus variés, allant au-delà des simples liens parents-enfants.

Dans ce qui suit, nous présentons les bases de données publiques les plus utilisées pour la vérification de la parenté faciale, tout en mettant en lumière celle retenue dans le cadre de notre projet.

- **KinFaceW-I et KinFaceW-II :**

La base de données KinFaceW-I et KinFaceW-II comprend deux sous-ensembles, contenant des images de personnalités publiques et de leurs proches, collectées à partir d'Internet avec une résolution de 64x64 pixels [24]. Les images ont été capturées dans des conditions non contrôlées, ce qui signifie qu'il n'y a pas de restrictions sur la pose, l'éclairage, le fond, l'expression, l'âge, l'ethnie, etc. La principale différence entre KinFaceW-I et KinFaceW-II réside dans la provenance des images : dans KinFaceW-I, les images des membres d'une même famille proviennent de photos différentes, tandis que dans KinFaceW-II, les images proviennent majoritairement de la même photo. La base couvre quatre relations de parenté : père-fils (F-S), père-fille (F-D), mère-fils (M-S) et mère-fille (M-D). KinFaceW-I contient respectivement 156, 134, 116 et 127 paires, tandis que pour KinFaceW-II, chaque relation comporte 250 paires d'images. Ces bases ont été parmi les premières à permettre une vérification de liens de parenté dans des conditions réalistes avant l'introduction de la base FIW [23].

- **TSKinFace :**

La base de données TSKinFace est constituée de triplets familiaux, comprenant des images de pères, mères et enfants, collectées à partir de scènes de séries télévisées turques. Les images sont obtenues dans des contextes plus naturels, ce qui ajoute de la variabilité en termes d'éclairage, de pose et d'expressions faciales. Cette base de données est composée de 513 triplets, où chaque triplet représente une relation père-mère-enfant. Elle se distingue des autres bases par son approche triangulaire,

qui permet de traiter les relations familiales sous un angle différent, en exploitant la dynamique familiale entre trois membres au lieu de deux. TSKinFace est particulièrement adaptée pour les méthodes qui évaluent la relation triangulaire et la similarité entre plusieurs membres d'une même famille [25].

- **UB Kinface :**

UB Kinface est collectée par des chercheurs de l'université de Northeastern aux Etats-Unis. L'ensemble de données contient 600 images de visages (64x64 pixels) correspondant à 400 personnes différentes réparties en 200 groupes (famille) ; Chaque groupe se compose d'images d'enfants, images de parents lorsqu'ils étaient jeunes et images de parents lorsqu'ils sont vieux. Les images de la base de données sont des collections du monde réel de personnages publics (célébrités et hommes politiques) provenant d'Internet. Les images des visages ont été collectées sans restriction en termes de pose, d'expression, d'éclairage, d'arrière-plan, d'âge, de sexe, d'éthnie, de saturation des couleurs, de qualité d'image, etc... UB Kinface peut être principalement divisé en deux parties selon la race (asiatique et non asiatique), chacune d'entre elles étant composée de 100 groupes, 200 personnes et 300 images. La répartition des relations est de : 93 paires père-fils (F-S) , 77 paires père-fille (F-D) , 12 paires mère-fils (M-S) et 18 paires mère-fille (M-D)[26].

- **FIW (Families In the Wild) :**

La base de données FIW est la plus grande et la plus complète en matière de vérification de parenté faciale [24]. Composée de plus de 10 000 images provenant de plus de 1000 familles issues de sources variées (réseaux sociaux, médias publics, etc.), FIW est conçue pour refléter la diversité des familles réelles. Elle couvre 11 types de relations familiales différentes, allant de parent-enfant à frères et sœurs, et comprend également des arbres généalogiques permettant une étude plus approfondie de la structure familiale. En raison de sa taille et de sa richesse, FIW est idéale pour entraîner des modèles de deep learning capables de gérer une grande variété de situations réelles, notamment les variations dues à l'âge, à l'éthnie, au genre et aux conditions d'éclairage.

- **Family101 :**

La base de données Family101 a été créée pour pallier certaines limitations des bases précédentes, notamment l'absence de diversité et l'impossibilité d'entraîner des modèles robustes pour des familles non occidentales. Elle contient 101 familles, avec des images couvrant une large gamme d'âges, de genres et d'éthnies. Ces images sont soigneusement sélectionnées pour fournir un ensemble de données équilibré, ce qui est essentiel pour l'entraînement de modèles capables de généraliser à des populations variées. Family101 est particulièrement utilisée pour valider des modèles développés sur d'autres bases plus grandes comme FIW, tout en permettant des expériences plus ciblées sur des contextes familiaux diversifiés [23].

- **Cornell KinFace :**

La base de données Cornell KinFace a été l'une des premières bases spécifiquement conçues pour la vérification de la parenté faciale. Elle contient des paires d'images de parent-enfant, obtenues dans des conditions relativement contrôlées, bien que moins rigoureuses que celles des bases plus récentes. Bien qu'elle soit moins utilisée aujourd'hui, elle reste utile pour des tests préliminaires et pour comparer des approches de détection de similarités faciales sur un ensemble plus restreint. Le nombre d'images est limité, mais l'ancienne popularité de cette base la rend toujours pertinente pour des expériences comparatives simples dans le domaine de la parenté faciale.[13]

Base de données	Nombre de familles	Nombre de personnes	Nombre de visages	Résolution	Variation d'âge	Arbre généalogique
CornellKin	150	300	300	100×100	Non	Non
UB Kinface	200	400	600	89×96	Oui	Non
KinfaceW-I	-	533	1066	64×64	Non	Non
KinfaceW-II	-	1000	2000	64×64	Non	Non
TS kinface	787	2589	-	64×64	Oui	Oui
Family101	101	607	14816	100×100	Oui	Oui

TAB. II.1 – Caractéristiques des bases de données de visages utilisées

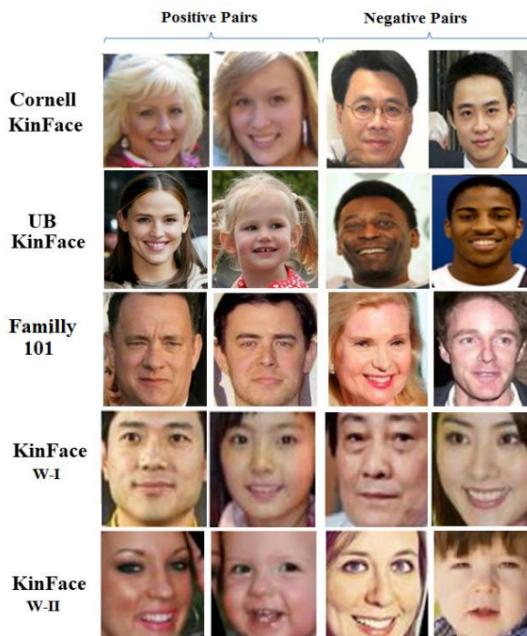


FIG. II.1 – Exemples d'images extraites de bases de données utilisées, comprenant pour chaque base deux échantillons visuels : un représentant un cas positif et un représentant un cas négatif [5]

II.4 Approches pour la vérification de la parenté

Au cours de la dernière décennie, de nombreuses études en physiologie humaine ont exploré les mécanismes cognitifs impliqués dans la reconnaissance et la vérification faciale. Ces avancées ont largement influencé les domaines de la vision par ordinateur et de l'apprentissage automatique, conduisant au développement de diverses méthodes automatiques de reconnaissance faciale, notamment dans le cadre de la vérification de parenté. Bien que ces approches aient montré des performances prometteuses, leur comparaison reste complexe en raison de l'hétérogénéité des bases de données utilisées, rendant difficile l'évaluation objective de leur efficacité respective.

II.4.1 Prétraitement du visage

Le prétraitement facial désigne l'ensemble des techniques utilisées pour préparer les images de visages en vue d'une analyse ultérieure. Parmi les techniques les plus couramment utilisées, on trouve notamment :

II.4.1.1 Détection du visage (avec MTCNN)

Le réseau neuronal convolutif en cascade multitâches (MTCNN) constitue une méthode efficace et robuste pour la détection et l'alignement automatique de visages dans les images [6]. L'architecture MTCNN se distingue par sa capacité à accomplir simultanément plusieurs tâches liées au visage, à savoir : la détection précise des visages, l'estimation de leur position ainsi que la localisation des points caractéristiques faciaux (appelés landmarks : yeux, nez, bouche). Elle repose sur une architecture composée de trois réseaux neuronaux convolutifs distincts, chacun effectuant une tâche spécifique :

Ces trois étapes successives, clairement illustrées par la figure II.2, assurent une détection robuste et efficace des visages, adaptée à une utilisation ultérieure dans des tâches telles que la vérification de parenté.

À partir d'une image en entrée, celle-ci est d'abord redimensionnée à différentes échelles afin de construire une pyramide d'images, qui servira d'entrée au cadre en cascade à trois étapes décrit ci-dessous [6].

- Étape 1 : Le processus débute par l'utilisation du Proposal Network (P-Net), un réseau convolutif léger chargé de parcourir l'image à différentes échelles pour identifier des zones candidates susceptibles de contenir un visage. Chaque région proposée est accompagnée d'un vecteur de régression ajustant sa boîte englobante. Pour éliminer les doublons ou détections trop proches, une opération de suppression non maximale (NMS) est appliquée.
- Étape 2 : Les régions conservées sont ensuite transmises au Refine Network (R-Net), un second réseau plus profond, qui permet de filtrer davantage les fausses détections. Ce réseau affine les prédictions des boîtes englobantes et applique une seconde NMS pour conserver uniquement les propositions les plus fiables.
- Étape 3 : Enfin, les candidats validés sont analysés par le Output Network (O-Net), qui effectue une vérification plus fine et précise. En plus de confirmer la présence d'un visage, ce réseau fournit la localisation de cinq points clés du visage (facial landmarks) : les deux yeux, le nez, et les deux coins de la bouche, essentiels pour les tâches d'alignement facial.

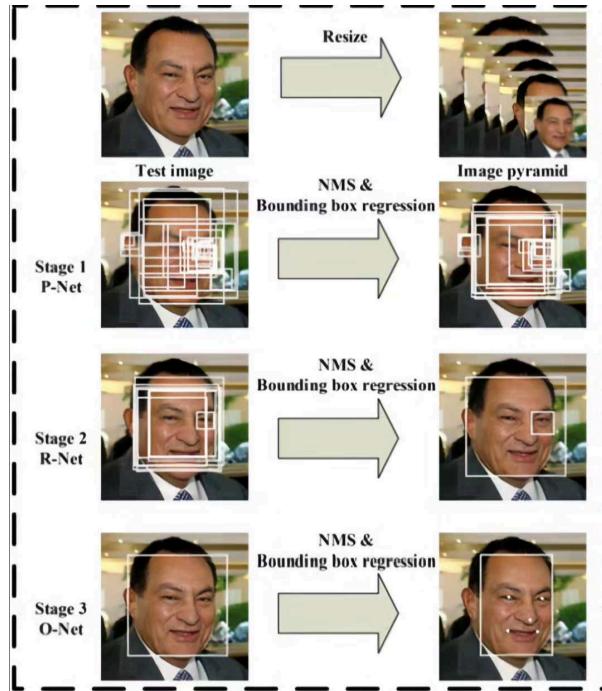


FIG. II.2 – Vue d’ensemble des résultats de MTCNN [6]

II.4.2 Extraction de Caractéristiques

Les méthodes d’extraction de caractéristiques se divisent en deux grandes catégories : celles basées sur des descripteurs texturaux, souvent associées à des techniques classiques de machine learning (shallow features), et celles reposant sur l’apprentissage profond (deep learning). L’utilisation de modèles profonds dans la vérification de parenté a longtemps été limitée, principalement en raison du manque de données suffisantes pour entraîner efficacement ces architectures.

II.4.2.1 Caractéristiques profondes basées sur un réseau de neurones convolutifs (CNN)

Avec les progrès récents en apprentissage profond, les réseaux de neurones convolutifs (CNNs) se sont imposés comme des outils fondamentaux dans le domaine de la vision par ordinateur. Conçus pour exploiter la structure spatiale des images, ces réseaux ont transformé la manière dont les systèmes automatisés perçoivent et interprètent le contenu visuel. Depuis leur introduction par LeCun [27], les CNNs ont démontré une efficacité remarquable dans un large éventail d’applications, allant de la classification d’images à la détection d’objets, en passant par la reconnaissance faciale et la segmentation d’images. Leur capacité à extraire automatiquement des représentations hiérarchiques des données visuelles en fait aujourd’hui une technologie incontournable pour le traitement et l’analyse d’images.

II.4.2.1.1 Architecture CNN

De nombreuses architectures de réseaux de neurones convolutifs (CNNs) ont été proposées [28]. L'architecture du modèle joue un rôle essentiel dans l'amélioration des performances selon les tâches visées. Depuis leur apparition en 1989, les CNNs ont connu de nombreuses évolutions [29], incluant des ajustements structurels, des techniques de régularisation, et des méthodes avancées d'optimisation des paramètres. Toutefois, l'amélioration significative des performances des CNNs résulte en grande partie de la réorganisation des unités de traitement et de l'introduction de nouveaux blocs d'opérations.

L'étude approfondie des caractéristiques de ces architectures telles que la taille d'entrée, la profondeur du réseau ou encore sa robustesse est cruciale pour orienter le choix du modèle selon l'application ciblée. Classiquement, une architecture CNN se compose de plusieurs couches : convolutionnelles, de sous-échantillonnage (pooling)[30], d'activation non linéaire (ReLU)[31], et de couches entièrement connectées. Ces réseaux assurent à la fois l'extraction de caractéristiques discriminantes et la classification comme illustré dans la figure II.3[7].

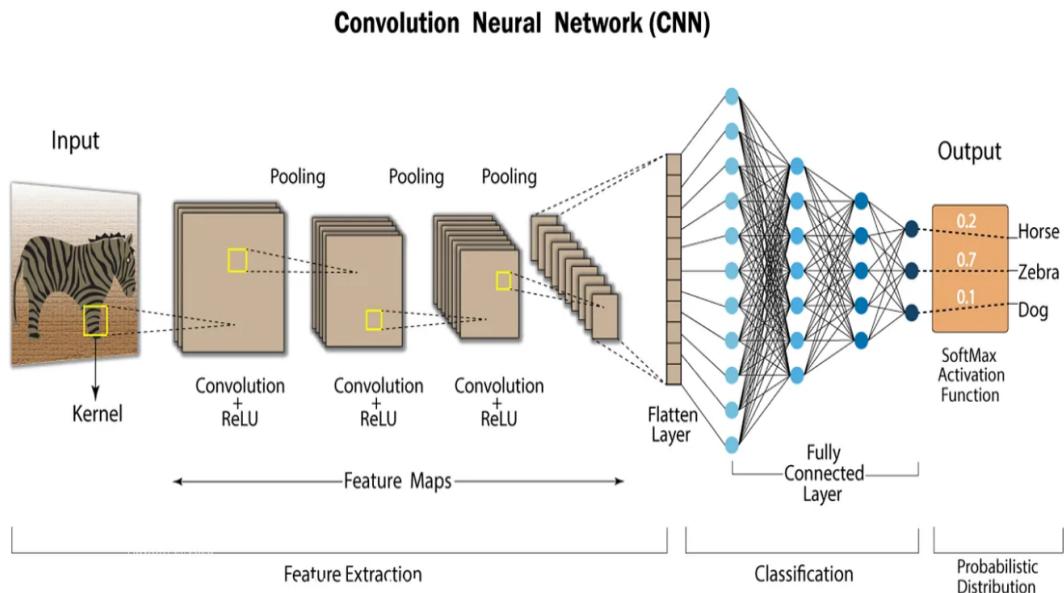


FIG. II.3 – Exemple de l'architecture CNN [7]

- **La couche d'entrée (Input)** : Chaque image est de dimension $[W_i, H_i, C_i]$, où W_i est sa largeur en pixels, H_i sa hauteur en pixels et C_i le nombre de canaux (1 pour une image niveau de gris, 3 pour une image en couleurs).
- **La couche de convolution** : Elle constitue l'élément central des réseaux de neurones convolutifs (CNN), car elle agit comme un extracteur automatique de caractéristiques. Son principe repose sur le balayage de l'image à l'aide de noyaux de convolution (appelés filtres), capables de détecter des motifs locaux simples tels que des bords, des textures ou des courbes. Contrairement aux approches classiques où ces filtres sont définis manuellement, dans un CNN, les poids des filtres sont appris directement par le réseau. Ils sont initialement attribués de manière aléatoire, puis ajustés progressivement durant l'entraînement à l'aide de l'algorithme de rétro-propagation du gradient, permettant ainsi au modèle d'optimiser sa capacité

à capturer des structures visuelles pertinentes.

La couche de convolution possède quatre hyperparamètres

- **Le nombre de filtres K** : définit la profondeur du volume de sortie.
- **La taille F des filtres** : chaque filtre est de dimensions F. F. D pixels.
- **Le pas (Stride) S** : Il désigne le nombre de pixels avec lequel le filtre se déplace à chaque étape lors de la convolution, il contrôle le chevauchement entre les fenêtres.
- **Le zero-padding P** : consiste à ajouter un contour noir (valeurs nulles) d'épaisseur P pixels autour de l'image d'entrée, afin de préserver ses dimensions après la convolution et éviter la perte d'information en bordure.

Produit une matrice de dimension [W0, H0, C0] : C0 correspond au nombre de filtres K.

$$W0 = \frac{Wi - F + 2P}{S} + 1 \quad (\text{II.1})$$

$$H0 = \frac{Hi - F + 2P}{S} + 1 \quad (\text{II.2})$$

► Opération de convolution de base :

On applique le filtre sur l'entrée, réaliser un produit élément par élément, puis sommer le résultat.

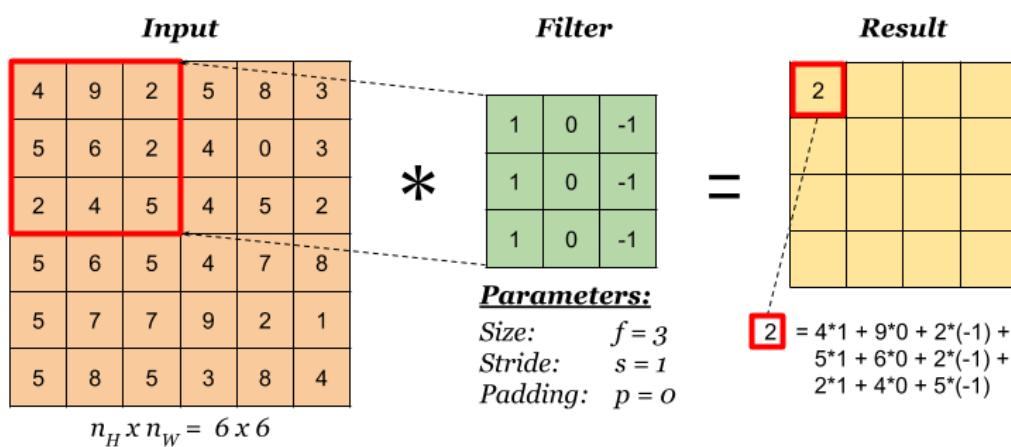


FIG. II.4 – Exemple des calculs fondamentaux réalisés à chaque passage dans une couche convolutionnelle.

On déplace la superposition d'une position vers la droite (ou selon la valeur du stride), puis répéter le calcul précédent pour obtenir le résultat suivant, et ainsi de suite.

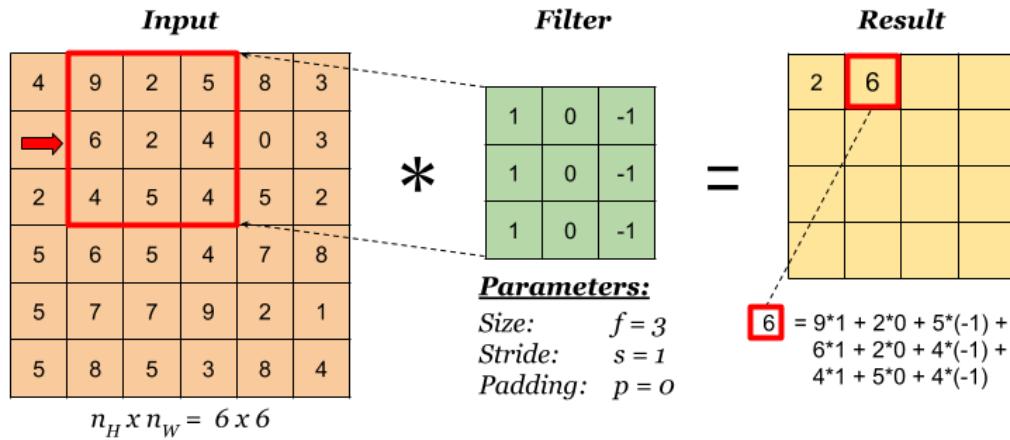


FIG. II.5 – Éxemple des calculs réalisés après un pas Stride (S = 1)

Le stride détermine de combien de positions le filtre se déplace dans l'entrée pour calculer la cellule suivante du résultat.

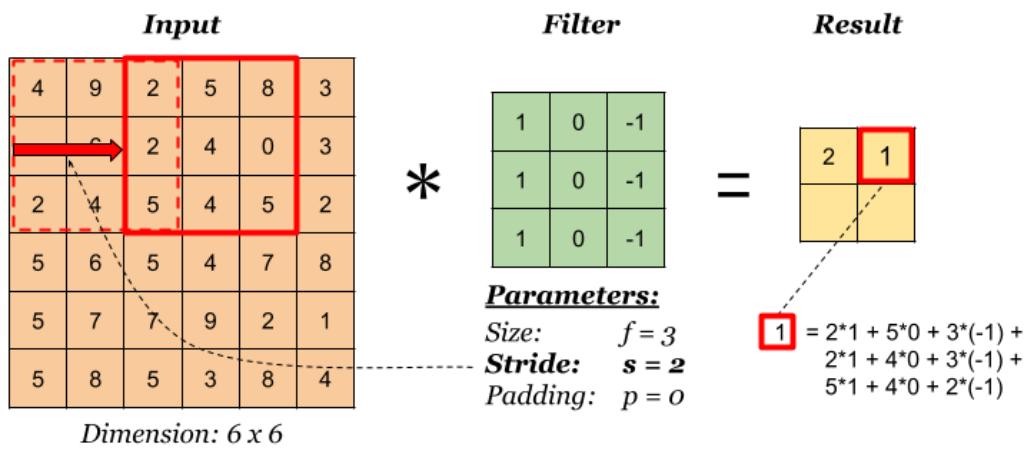


FIG. II.6 – Éxemple du pas (Stride) pour S = 2

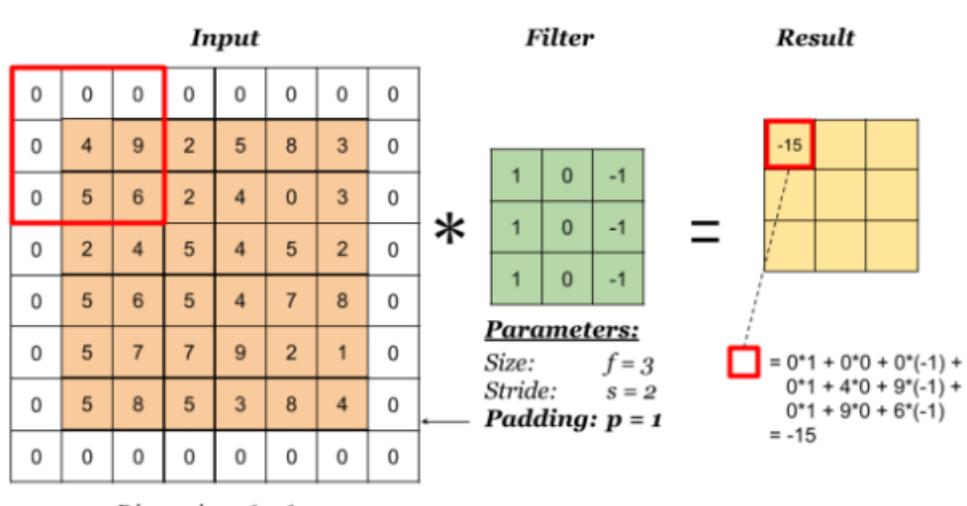


FIG. II.7 – Éxemple du zero-padding

- **La couche d’activation ReLU :** C’est la fonction la plus utilisée dans les réseaux de neurones convolutifs [31]. Elle transforme toutes les valeurs négatives de l’entrée en zéro, tout en conservant les valeurs positives inchangées. Son principal avantage réside dans sa simplicité de calcul, ce qui la rend plus rapide et moins coûteuse en ressources que d’autres fonctions d’activation. Sa définition mathématique est donnée par la formule suivante :

$$F_{ReLU}(x) = \max(0, x) \quad (\text{II.3})$$

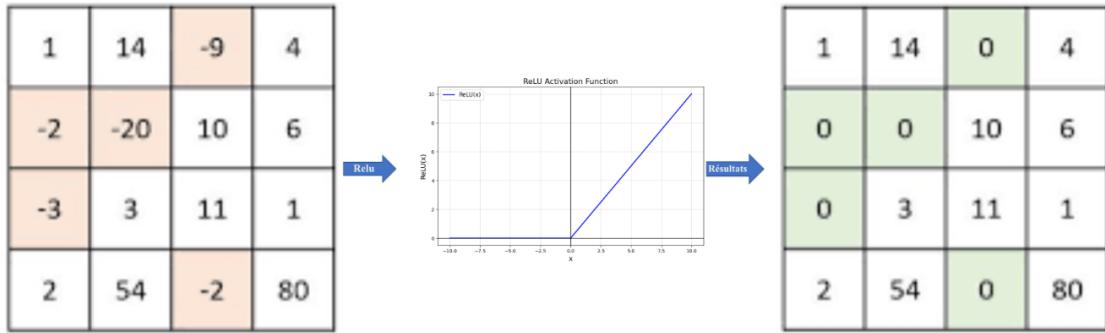


FIG. II.8 – Exemple de la Couche d’activation ReLu

- **La couche pooling :** Dans les réseaux de neurones convolutifs (CNN), la couche de Pooling est généralement placée entre deux couches de convolution. Elle joue un rôle essentiel en réduisant la dimension spatiale des cartes de caractéristiques générées par les opérations de convolution, tout en préservant les informations les plus importantes. Cela permet de limiter le nombre de paramètres, de réduire la charge computationnelle et de diminuer le risque de sur-apprentissage (overfitting). Concrètement, l’image est découpée en cellules régulières (patchs), et une opération statistique est appliquée sur chacune d’elles, comme le maximum (Max-Pooling) [30] ou la moyenne (Average-Pooling) [30]. Cette technique permet également d’obtenir des représentations plus robustes, invariantes aux translations, rotations et petites déformations des objets.

Il existe plusieurs variantes de pooling selon les besoins de la tâche, notamment le Global Average Pooling (GAP), le Global Max Pooling, le Min-Pooling, le Tree Pooling, ou encore le Gated Pooling [32]. Ces méthodes peuvent être choisies en fonction des objectifs du réseau et de la nature des données traitées.

Le choix le plus fréquent est l’utilisation du Max-Pooling avec un filtre de taille 2×2 et un stride de 2, ce qui évite le chevauchement entre les régions traitées. La couche de Pooling possède deux hyperparamètres :

- La taille des patchs F : l’image est découpée en patch carrées de taille $F \times F$ pixels.
- Le pas S : les patchs sont séparées les unes des autres de S pixels.

Accepte un volume de taille [L1,H1,C1] et produit matrice de dimensions [W2,H2,C2]
où :

$$W2 = \frac{W1 - F}{S} + 1 \quad (\text{II.4})$$

$$H2 = \frac{H1 - F}{S} + 1 \quad (\text{II.5})$$

$$C2 = C1 \quad (\text{II.6})$$

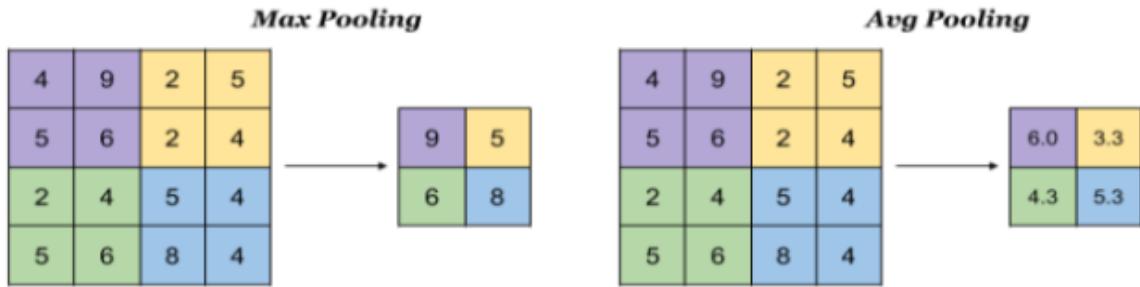


FIG. II.9 – Exemple des deux variant max pooling et average pooling

- **La couche d'aplatissement (Flatten layer) :** Elle joue un rôle de transition entre les couches convolutionnelles (et de pooling) et les couches entièrement connectées dans une architecture de réseau de neurones. Elle a pour fonction de transformer les cartes de caractéristiques (feature maps) multidimensionnelles (telles que les tenseurs 3D issus des convolutions) en un vecteur unidimensionnel. Cette opération est nécessaire, car les couches entièrement connectées attendent une entrée sous forme de vecteur. Ainsi, la couche Flatten Layer permet de "linéariser" les données tout en conservant les informations extraites par les couches précédentes. Bien qu'elle n'effectue aucun calcul, cette étape structurelle est essentielle pour préparer les données à la phase de classification finale.

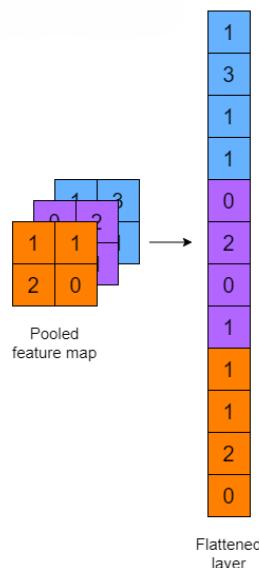


FIG. II.10 – Exemple de la couche d'aplatissement (Flatten layer)

- **La Couche entièrement connectées (Fully connected layer)** : La couche entièrement connectée, souvent située en fin de réseau convolutif (CNN), joue un rôle essentiel dans la phase de classification. Contrairement aux couches précédentes qui traitent des cartes de caractéristiques locales, elle relie chaque neurone à tous ceux de la couche précédente, permettant une combinaison globale des informations extraites.

Son objectif est de transformer les représentations spatiales issues des couches convolutionnelles et de pooling en un vecteur de décision, adapté à la tâche de classification. Elle agit ainsi comme un classificateur traditionnel.

Dans le cadre de la vérification de parenté faciale, cette couche se termine généralement par un seul neurone de sortie, activé par une fonction sigmoïde, qui renvoie une probabilité indiquant si les deux individus sont apparentés ou non.

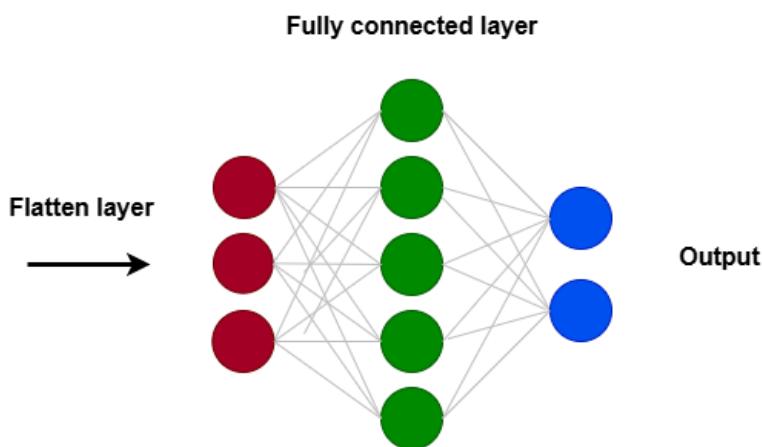


FIG. II.11 – Éxemple de la couche entièrement connecté (fully connected layer)

II.4.2.1.2 Modèle pré-entraîné

- **InceptionV3 :**

C'est un réseau de neurones convolutif développé par Google pour la classification d'images. Il a été entraîné sur la base de données ImageNet, contenant plus d'un million d'images réparties sur 1000 catégories, ce qui lui permet d'apprendre des représentations visuelles efficaces et généralisables.

Ce modèle fait partie de la famille des architectures Inception, connues pour leur bon équilibre entre précision et efficacité computationnelle. Il utilise des modules qui combinent plusieurs tailles de filtres (1x1, 3x3, 5x5), afin de capturer l'information à différentes échelles. Il intègre également des techniques modernes comme la batch normalization [33] et la réduction de dimensions.

Dans notre projet, nous utilisons une version pré-entraînée d'InceptionV3. Les couches finales de classification sont supprimées, et nous extrayons les caractéristiques profondes depuis la couche de moyenne globale (avg-pool). Le résultat est un vecteur de 1000 valeurs qui décrit le contenu visuel du visage de manière détaillée. Ce vecteur sert ensuite à comparer les visages entre eux, afin de vérifier s'il existe un lien de parenté.

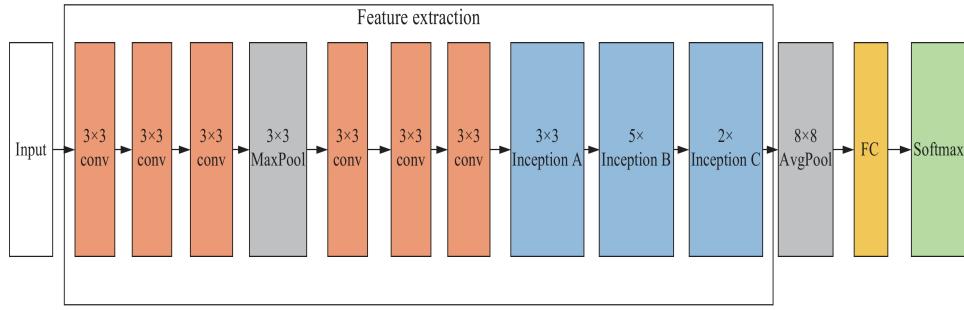


FIG. II.12 – Exemple de l’architecture du modèle InceptionV3

II.4.2.2 Descripteurs texturaux

Les descripteurs texturales désignent des caractéristiques extraites à l'aide de méthodes classiques, sans apprentissage automatique. Ils reposent sur des mesures statistiques ou texturales calculées directement à partir des pixels de l'image. Simples à mettre en œuvre et peu coûteux en calcul, ces descripteurs permettent de capturer des informations locales utiles, en complément des descripteurs profonds, dans des tâches telles que la vérification de parenté.

- **Quantification de phase locale (LPQ)**

En complément des descripteurs profonds extraits par des modèles CNN comme InceptionV3, l'utilisation de descripteurs texturaux (ou Shallow) permet d'exploiter des informations locales et texturales importantes, parfois négligées par les modèles profonds.

Le descripteur LPQ (Local Phase Quantization) a été initialement proposé par Ojansivu et Heikkilä comme opérateur de description de texture [34]. Il repose sur la propriété d'invariance au flou du spectre de phase de Fourier. Pour ce faire, LPQ extrait l'information de phase locale à l'aide de la transformée de Fourier à court terme bidimensionnelle (2D-STFT), appliquée sur un voisinage rectangulaire centré autour de chaque pixel de l'image, comme présenté dans l'équation suivante.

$$F_u(X) = \sum_{m \in N_x} h(m - x) f(m) \exp^{-2j\pi u^T m} = E_u^T f_x \quad (\text{II.7})$$

où E_u , de taille $1M^2$, est un vecteur de base de la transformée de Fourier discrète bidimensionnelle (2DWFT) associé à la fréquence u , et f_x , de taille $M^2 \times N$, est un vecteur contenant les valeurs des pixels de l'image dans N_x à chaque position x . La fonction fenêtre $h(x)$ est une fonction rectangulaire.

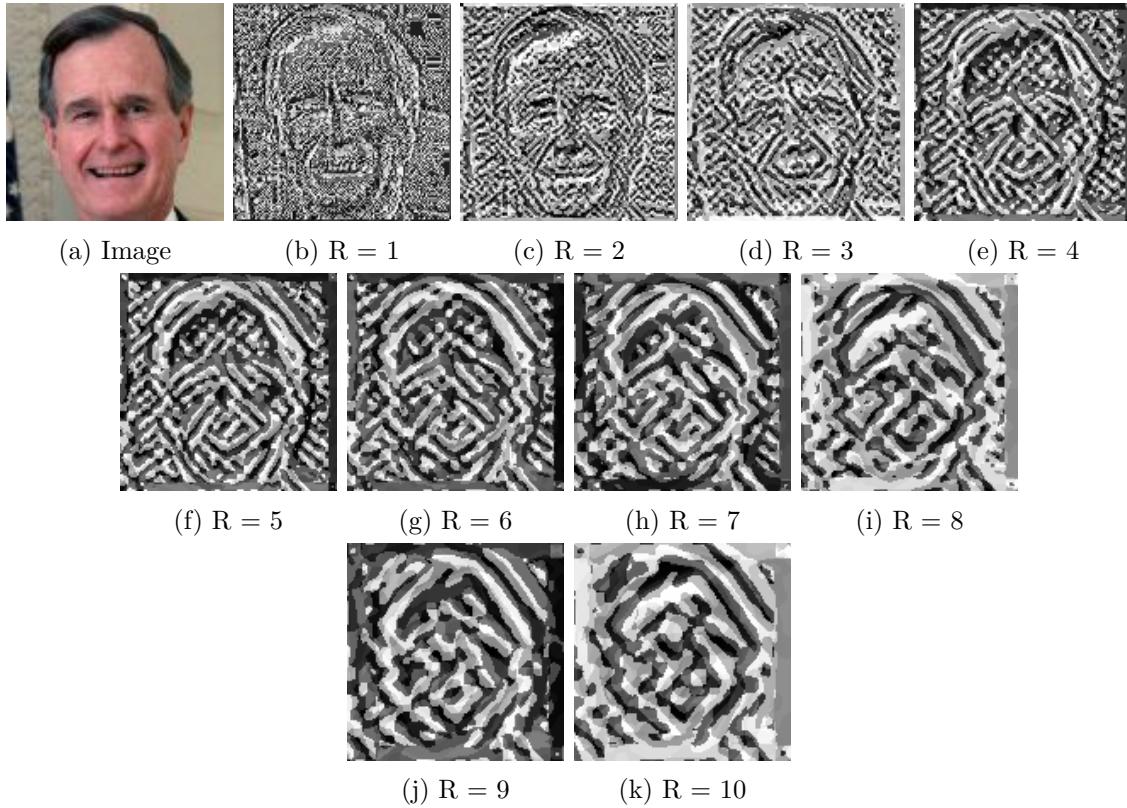


FIG. II.13 – Exemple de la Quantification de phases locales (LPQ) avec différents rayons $R= 1 \text{ à } 10$

Dans la méthode de Local Phase Quantization (LPQ), seuls quatre coefficients complexes correspondant à des fréquences bidimensionnelles spécifiques sont pris en compte. Dans nos expériences, nous avons utilisé la version originale du code fournie par les auteurs de LPQ. Cette méthode se déroule en quatre étapes principales. Tout d'abord, l'opérateur LPQ est appliqué à l'image d'entrée afin de produire une image étiquetée. Ensuite, cette image est divisée en petites régions locales. Pour chacune d'elles, un histogramme des valeurs d'étiquettes est calculé afin de constituer un vecteur de caractéristiques locales. Enfin, ces histogrammes sont concaténés pour former un vecteur global représentant l'ensemble de l'image.

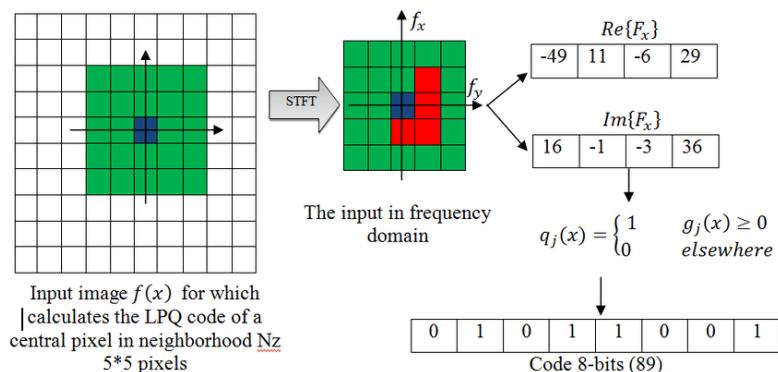


FIG. II.14 – Éxemple des étapes nécessaires à la génération du vecteur de caractéristiques par la méthode LPQ (Local Phase Quantization),

- Caractéristiques statistiques binarisées d'image (BSIF)

Le descripteur BSIF (Binarized Statistical Image Features), introduit par Kannala et Rahtu [35], repose sur un ensemble de filtres linéaires appris automatiquement à partir de patchs d'images naturelles à l'aide de l'Analyse en Composantes Indépendantes (ICA). Les réponses obtenues par convolution sont ensuite binarisées afin de produire une représentation compacte des textures locales. Contrairement à des méthodes comme LPQ, qui s'appuient sur le calcul de statistiques de labels dans les voisinages de pixels, BSIF encode directement les motifs locaux à partir de réponses filtrées statistiquement apprises.

Soit une image X de taille $l \times l$, et un ensemble de n filtres W_i . La réponse du filtre S_i est calculée par :

$$s_i = \sum_{u,v} W_i(u, v) X(u, v) = \mathbf{w}_i^T \mathbf{x} \quad (\text{II.8})$$

Soit un ensemble de n filtres linéaires. On peut empiler ces filtres dans une seule matrice W_i pour calculer toutes les réponses simultanément.

Dans la dernière étape, on introduit la notation vectorielle : les vecteurs w et x représentent respectivement les pixels du filtre W_i et de la région d'image X .

La fonction binarisée b_i est alors définie par :

$$b_i = \begin{cases} 1 & \text{si } s_i > 0 \\ 0 & \text{sinon} \end{cases} \quad (\text{II.9})$$

Chaque pixel est ainsi représenté par un code binaire de n bits. L'image peut ensuite être décrite globalement à l'aide d'un histogramme des codes binaires, facilitant l'analyse de texture pour des tâches telles que la reconnaissance faciale.

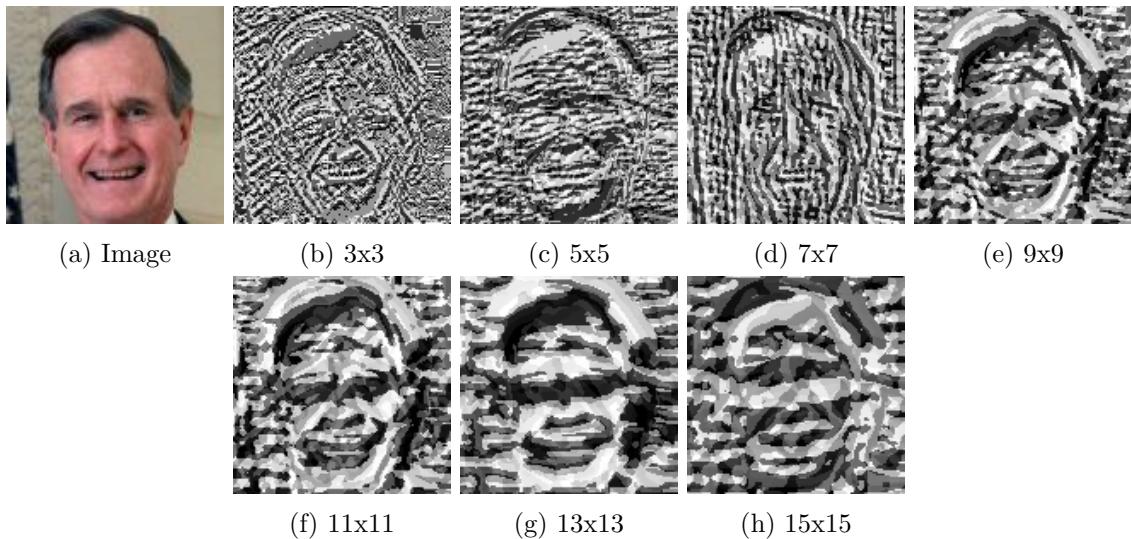


FIG. II.15 – Exemple de Caractéristiques BSIF avec différents filtres de 3x3 à 15x15.

II.4.3 Classification

Dans le cadre de la vérification de parenté faciale, la classification vise à déterminer si deux visages partagent un lien de parenté, à partir des caractéristiques extraites. Après fusion et réduction de dimension par TXQEDA, suivie d'une normalisation intra-classe via WCCN.

II.4.3.1 Réduction de dimension par TXQEDA

La méthode TXQEDA (Tensor Cross-view Quadratic Exponential Discriminant Analysis) est une extension tensorielle avancée de l'approche XQDA[36], intégrant les principes discriminatifs de XQEDA tout en exploitant la structure multi-dimensionnelle (tensorielle) des données issues de l'extraction de caractéristiques faciales. Dans le contexte de la vérification de parenté, cette méthode se révèle particulièrement pertinente, car elle permet de mieux modéliser la complexité des relations héréditaires tout en réduisant efficacement la dimensionnalité des descripteurs.

Contrairement aux méthodes vectorielles classiques, TXQEDA opère directement sur des tenseurs d'ordres supérieurs, préservant ainsi les corrélations entre les différentes dimensions des caractéristiques extraites. Son objectif est de maximiser la séparation entre les classes (parenté vs non-parenté) tout en minimisant la variance intra-classe dans un espace projeté de plus faible dimension, afin d'améliorer la capacité de discrimination du système.

TXQEDA utilise des tenseurs d'entraînement d'ordre 3 pour représenter les descripteurs extraits des images faciales. Ces tenseurs \mathcal{X} et $\mathcal{Y} \in R^{I_1 \times I_2 \times I_3}$, chaque mode du tenseur a une signification précise :

- I_1 correspond aux différentes échelles d'extraction
- I_2 contient les valeurs des histogrammes
- I_3 représente les échantillons de visages

Il apprend ensuite un espace réduit qui sépare mieux les paires apparentées des non-apparentées, tout en conservant les relations importantes entre les données. Cela permet d'améliorer la vérification de parenté.

II.4.3.2 Normalisation par la Covariance Intra-classe (WCCN)

La méthode WCCN (*Within-Class Covariance Normalization*) est utilisée comme technique de réduction de la variance intra-classe, dans le but d'améliorer la séparabilité entre les classes. Dans notre système, nous appliquons WCCN [37] avant le calcul de similarité afin de rendre l'espace des caractéristiques plus discriminant pour la vérification de parenté.

La matrice de normalisation WCCN, notée (B) , est obtenue à partir de la décomposition de Cholesky :

$$BB^T = W^{-1} \quad (\text{II.10})$$

La matrice de covariance intra-classe (W), est estimée à partir des vecteurs de caractéristiques w_i centrés autour de leur moyenne de classe, comme suit :

$$W = \frac{1}{S} \sum_{s=1}^S \sum_{i=1}^{n_s} (A^T (w_i^s - \bar{w}_s)) (A^T (w_i^s - \bar{w}_s))^T \quad (\text{II.11})$$

Enfin, chaque vecteur w_i est transformé selon :

$$\tilde{w}_i = B^{-1} w_i \quad (\text{II.12})$$

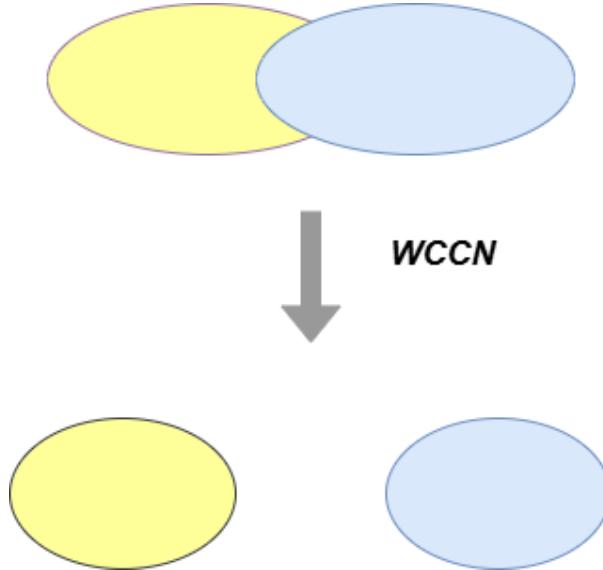


FIG. II.16 – Exemple de WCCN

II.4.3.3 Similarité cosinus

C'est une mesure utilisée pour évaluer la similarité entre deux vecteurs de caractéristiques, ici notés \mathbf{u} et \mathbf{v} . Elle s'appuie sur l'angle entre les deux vecteurs, indépendamment de leur longueur. Elle est particulièrement adaptée aux représentations issues de réseaux neuronaux profonds ou de descripteurs texturaux normalisés.

La formule de la similarité cosinus est donnée par :

$$\text{Similarité cosinus}(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} = \frac{\sum_{i=1}^n u_i v_i}{\sqrt{\sum_{i=1}^n u_i^2} \sqrt{\sum_{i=1}^n v_i^2}} \quad (\text{II.13})$$

où :

- $\mathbf{u} = (u_1, u_2, \dots, u_n)$ et $\mathbf{v} = (v_1, v_2, \dots, v_n)$ sont les vecteurs de caractéristiques de dimension n ,
- $\mathbf{u} \cdot \mathbf{v}$ représente le produit scalaire entre \mathbf{u} et \mathbf{v} ,
- $\|\mathbf{u}\| = \sqrt{\sum_{i=1}^n u_i^2}$ est la norme euclidienne du vecteur \mathbf{u} ,
- $\|\mathbf{v}\| = \sqrt{\sum_{i=1}^n v_i^2}$ est la norme euclidienne du vecteur \mathbf{v} .

La valeur de la similarité cosinus varie entre -1 et 1 , où 1 indique une forte similarité entre les visages, les deux individus sont probablement apparentés, 0 indique une faible similarité, les visages sont plutôt différents, et -1 indique une opposition, les visages sont très différents, donc probablement non apparentés.

En pratique, un seuil de décision est fixé : si la similarité est supérieure ou égale à ce seuil, la paire est classée comme apparenté, sinon non-apparenté

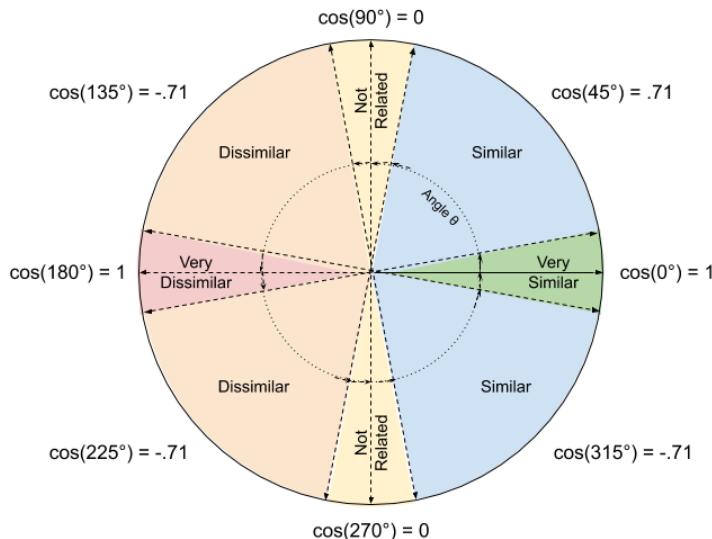


FIG. II.17 – Exemple de Similarité cosinus

II.4.4 Décision

La tâche de décision dans les systèmes de vérification consiste à déterminer si une paire d'images faciales sont apparentés ou non. Pour prendre cette décision, une mesure de similarité est utilisée afin d'évaluer la proximité entre les vecteurs de caractéristiques extraits. Dans notre système, nous utilisons la similarité cosinus. Si la valeur de cette similarité dépasse un seuil de décision prédéfini, la paire est considérée comme apparentée, sinon elle est jugée non apparentée.

Afin de garantir la fiabilité de l'évaluation, une validation croisée à 5 plis (5-fold cross-validation) est appliquée. Elle permet de tester le système sur différentes subdivisions des données.

Enfin, pour renforcer la décision, les différentes informations issues des descripteurs (profonds et texturaux) sont fusionnées au moyen d'un classifieur par régression logistique. Ce dernier apprend à discriminer les paires apparentées des non-apparentées en exploitant la complémentarité entre les descripteurs, et permet ainsi d'améliorer la précision finale du système.

II.5 Architecture générale du système

La reconnaissance des relations de parenté à partir d'images faciales repose sur l'extraction de caractéristiques robustes et discriminantes. Dans ce contexte, nous proposons une approche combinant un descripteur profond (InceptionV3), et des descripteurs texturaux (LPQ, BSIF), fusionnés à l'aide d'une technique puissante (fusion LR). Cette combinaison permet de capturer des informations complémentaires facilitant la vérification des liens biologiques entre individus.

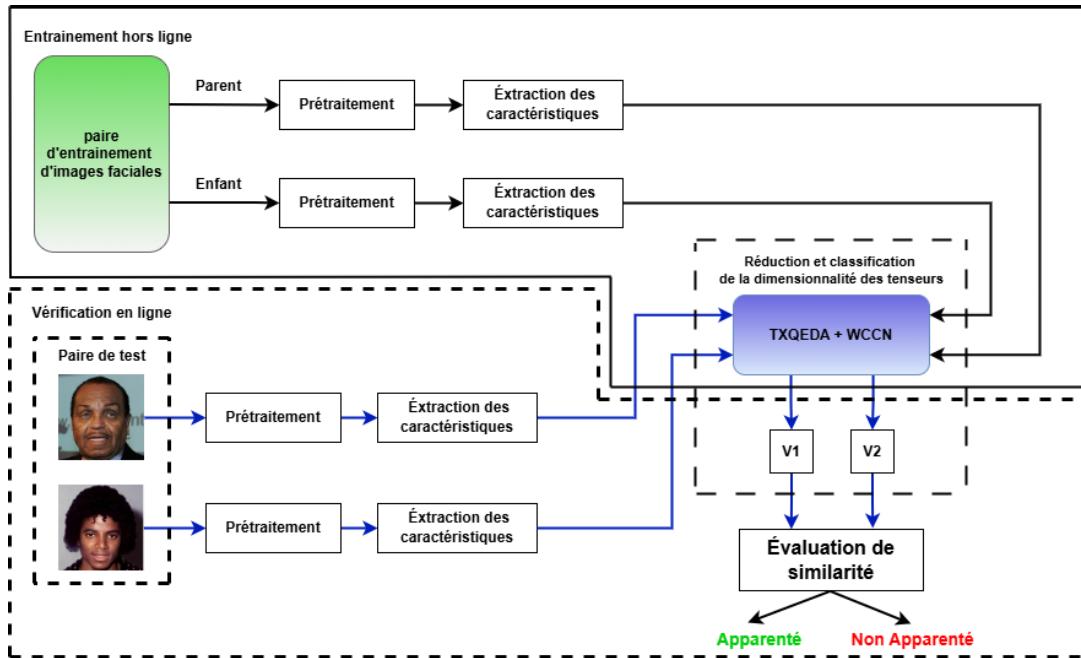


FIG. II.18 – Schéma du système proposé de vérification de la parenté faciale

II.6 Conclusion

Dans ce chapitre, nous avons détaillé la méthode proposée pour la vérification de la parenté faciale, en exposant les différentes étapes du processus : prétraitement des images, extraction de caractéristiques à l'aide de descripteurs profonds et texturaux, réduction de dimension, normalisation et classification. L'ensemble de ces modules constitue l'architecture complète de notre système.

Dans le chapitre suivant, nous procéderons à l'évaluation expérimentale de cette méthode, en analysant ses performances sur différentes bases de données et en la comparant aux approches existantes.

Chapitre **III**

Éxpérimentations et résultats

III.1 Introduction

Les chercheurs en vision par ordinateur à travers le monde s'efforcent continuellement d'optimiser les performances des systèmes de vérification de la parenté faciale, en explorant et en développant de nouvelles approches et techniques adaptées à cette tâche complexe.

Dans ce dernier chapitre, nous procédons à l'implémentation complète de notre système de vérification de parenté faciale en combinant deux types d'extraction de caractéristiques : les caractéristiques texturales (shallow feature) BSIF et LPQ, et le descripteur profond (deep feature) InceptionV3. Dans le but d'optimiser les performances de notre système, une série d'expérimentations est menée sur deux bases de données de parenté bien connues, à savoir Cornell KinFace et UB KinFace, en explorant plusieurs configurations et paramètres. Nous utilisons une méthode récente et performante de réduction de la dimension intitulée Tensor Cross-View Quadratic Exponential Discriminant Analysis (TXQEDA), une stratégie de fusion par régression logistique (Logistic Regression Fusion) est appliquée pour combiner les meilleurs scores issus de chaque descripteur.

III.2 Environnement de travail

Le langage de programmation utilisé dans ce travail est MATLAB, exécuté dans l'environnement de développement du même nom (dans notre cas MATLAB R2024a) et développé par The MathWorks. MATLAB permet une mise en oeuvre rapide et efficace des algorithmes, la réalisation de tâches nécessitant une puissance de calcul élevée, la manipulation et l'affichage de courbes, ainsi que la conception d'interfaces graphiques.

Les expériences ont été réalisées sur un ordinateur équipé d'un processeur Intel(R) Core(TM) i5-12400F cadencé à 2,5 GHz, d'une carte graphique NVIDIA GeForce RTX 3080, et de 16 Go de mémoire vive (RAM).

Nous avons utilisé l'environnement Deep Learning Toolbox Model pour la mise en œuvre des réseaux de neurones convolutifs (CNN).

III.3 Protocole de travail

Dans le cadre de notre travail, nous avons utilisé la technique de validation croisée à 5 plis (5-fold cross-validation), qui consiste à diviser aléatoirement la base de données en cinq sous-ensembles de taille égale. Cette méthode permet d'entraîner et de tester le modèle sur différentes partitions du base de données, dans le but d'assurer une bonne capacité de généralisation et de limiter le risque de sur-apprentissage. Le principe repose sur l'exécution de cinq itérations : à chaque itération, l'un des sous-ensembles est utilisé comme ensemble de test, tandis que les quatre autres servent à l'entraînement. Ce processus est répété cinq fois, en faisant varier à chaque fois le fold de test, de sorte que chaque sous-ensemble soit utilisé exactement une fois pour l'évaluation. Les performances finales sont obtenues en calculant la moyenne des scores obtenus sur les cinq itérations.

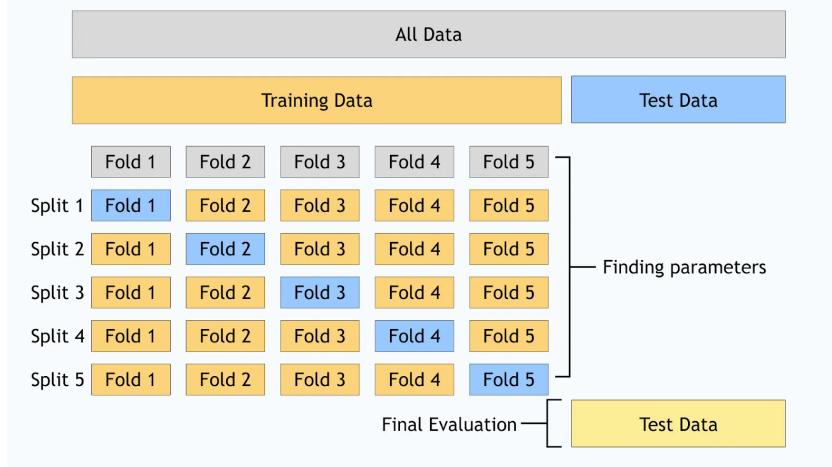


FIG. III.1 – Éxemple de la validation croisée K-fold, avec K=5

III.4 Expérimentations et Résultats

Nous menons plusieurs expériences pour évaluer la méthode utilisée :

- Pour les caractéristiques texturales, nous utilisons deux descripteurs texturaux (LPQ, BSIF)
 - a. Nous fixons le nombre de blocs à 12 pour LPQ et la valeur du rayon est modifiée de manière itérative avec les valeurs : 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.
 - b. Pareil avec BSIF nous fixons le nombre de blocs à 12 et la taille du filtre est modifiée de manière itérative avec les valeurs : 3x3, 5x5 7x7, 9x9, 11x11, 13x13, 15x15.
- Pour les caractéristiques profondes, nous utilisons le modèle pré-entraîné InceptionV3.
 - a. Dans InceptionV3, une seule échelle de caractéristiques est extraite à partir de la couche avg-pool, qui génère un vecteur de 2048 dimensions juste avant la couche de classification.
- Le meilleur paramétrage de BSIF et LPQ est déterminé, puis fusionné avec le score du modèle InceptionV3.

III.4.1 Éxperimentations sur la base de données Cornell kinFace

Nous présentons ici les résultats de nos expériences sur la base de données Cornell kinFace. Le Tableau III.1 illustre la précision moyenne du descripteur LPQ, pour laquelle nous avons fait varier le rayon ($R = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$), ainsi que celle de l'approche BSIF selon ses différents filtres (3x3, 5x5, 7x7, 9x9, 11x11, 13x13, 15x15), et en fixant le nombre de blocs d'histogrammes à 12.

Ensuite, les expériences ont été menées avec l'extraction de caractéristiques profondes à partir des images originales, en utilisant InceptionV3. Les résultats sont présentés avec les meilleurs scores des descripteurs texturaux (LPQ et BSIF) dans le Tableau III.2 .

Chapitre III. Expérimentations et résultats

Enfin, la fusion au niveau des scores entre les meilleurs résultats précédents est illustrée dans le Tableau III.3.

Descripteurs texturaux	Échelles	Précision moyenne (%)
LPQ	R=1	91.29
	R=2	90.32
	R=3	90.69
	R=4	90.65
	R=5	90.29
	R=6	88.89
	R=7	89.17
	R=8	89.27
	R=9	89.69
	R=10	89.23
BSIF	F=3*3	87.38
	F=5*5	89.52
	F=7*7	90.55
	F=9*9	90.94
	F=11*11	91.67
	F=13*13	91.66
	F=15*15	92.42

TAB. III.1 – Précision moyenne (%) des descripteurs texturaux LPQ et BSIF sur la base de données Cornell kinFace

Descripteurs Texturaux et profonds	Précision moyenne (%)
LPQ	91.29
BSIF	92.42
InceptionV3	71.34

TAB. III.2 – La précision moyenne (%) des meilleures scores de chaque descripteurs sur la base de données Cornell kinFace

Fusion	Précision moyenne (%)
LR fusion	93.75

TAB. III.3 – Résultat final après la fusion des meilleurs résultats de chaque descripteur (Texturaux et profond) avec la régression logistique (LR)

III.4.2 Expérimentations sur la base de données de UB KinFace

La même expérience est menée sur la base de données UB KinFace. Nous présentons les résultats comme suit : le tableau III.4 montre la précision moyenne (%) des descripteurs texturaux LPQ et BSIF sur la base de données UB KinFace et le tableau III.5, la précision moyenne (%) des meilleures scores de chaque descripteurs sur la base de données UB KinFace et enfin le tableau III.6 montre le résultat final après la fusion des meilleurs scores de chaque descripteur (Textural et profond) avec la régression logistique (LR).

Les relations sont les suivantes : **Set1** : paires parent-enfant avec un grand écart d'âge. **Set2** : paires parent-enfant avec un écart d'âge plus réduit.

Descripteurs texturaux	Échelles	Set1	Set2	Précision moyenne(%)
LPQ	R=1	85.94	88.52	87.23
	R=2	85.66	87.75	86.70
	R=3	84.96	86.79	85.87
	R=4	85.40	86.25	85.82
	R=5	84.40	85.77	85.08
	R=6	84.9	86.27	85.61
	R=7	85.20	87.29	86.24
	R=8	82.97	85.78	84.37
	R=9	81.94	83.53	82.73
	R=10	81.70	83.30	82.5
BSIF	F=3*3	80.45	82.27	81.36
	F=5*5	81.22	84.82	83.02
	F=7*7	84.17	86.56	85.36
	F=9*9	83.47	85.55	84.51
	F=11*11	83.18	85.02	84.10
	F=13*13	83.13	84.73	83.93
	F=15*15	82.90	84.02	83.46

TAB. III.4 – Précision moyenne (%) des descripteurs texturaux LPQ et BSIF sur la base de données UB KinFace

Descripteurs Texturaux et Profonds	Set1	Set2	Précision moyenne(%)
LPQ	85.94	88.52	87.23
BSIF	84.17	86.56	85.36
InceptionV3	63.96	65.96	64.96

TAB. III.5 – La précision moyenne (%) des meilleures scores de chaque descripteurs sur la base de données UB KinFace

Fusion	Set1 (%)	Set2 (%)	Précision moyenne (%)
LR Fusion	87.90	91.03	89.46

TAB. III.6 – Résultat final après la fusion des meilleurs scores de chaque descripteur (Textural et profond) avec la régression logistique (LR)

III.5 Synthèse des performances obtenues

Les expériences que nous avons menées en utilisant l'approche proposée, à savoir la fusion par régression logistique basée sur TXQEDA de deux types de caractéristiques (LPQ, BSIF, InceptionV3) sur deux base de données (Cornell KinFace et UB KinFace), nous amènent à conclure ce qui suit :

- **L'efficacité des caractéristiques texturales :** Nous avons appliqué deux descripteurs, LPQ et BSIF. Pour le descripteur BSIF, nous avons obtenu la meilleure précision moyenne de 92.42 % avec le filtre 15x15 et 85.36 % avec le filtre 7x7, et pour LPQ, la meilleure précision moyenne atteinte est de 91.29 % (lorsque R = 1) et 87.23 % (lorsque R = 1) avec les bases de données Cornell KinFace et UB KinFace, respectivement.
- **L'efficacité du descripteur profond :** Nous avons employé le modèles CNN pré-entraînés InceptionV3. Après simulation, nous avons constaté que les résultats obtenus n'étaient pas aussi performants que ceux issus des caractéristiques texturales : par exemple la précision moyenne était de 71.34 % avec Cornell KinFace et de 64.96 % avec UB KinFace, Cependant, nous savons que les caractéristiques extraites ne sont pas identiques, c'est pourquoi nous les conservons afin d'exploiter leur complémentarité.
- **L'efficacité de la fusion par régression logistique opérée au niveau des scores :** Nous utilisons l'approche LR pour combiner, au niveau des scores, les meilleurs résultats issus des différents types de caractéristiques (LPQ, BSIF et InceptionV3). Il est impressionnant de constater que la méthode de fusion par régression logistique affiche d'excellentes performances, atteignant des précisions de

93.75 % sur la base de données Cornell KinFace et de 89.46 % sur UB KinFace (voir Tableau III.3 et III.6). Ces résultats démontrent que la fusion par régression logistique permet de booster significativement la précision.

III.6 Comparaison avec les travaux connexes

Les résultats les plus performants de notre méthode, obtenus par la fusion des scores des descripteurs LPQ, BSIF, InceptionV3, combinés à la technique de réduction de dimension TXQEDA+WCCN, ont été comparés à ceux de méthodes récentes dans le Tableau III.7, en se basant sur les bases de données Cornell KinFace et UB KinFace. Les travaux de comparaison sont sélectionnés en fonction des algorithmes utilisés.

Travail	Année	Algorithmme	Base de données Cornell	Base de données UB
Bessaoudi & al [38]	2019	MSIDA	86.87	83.34
Zhang & al. [39]	2021	AdvKin	81.40	75.00
Mukherjee & al. [40]	2022	BC2DA	83.07	83.30
Goyal & al. [41]	2023	MLDPL	84.16	87.90
Proposé	2025	Fusion	93.75	89.46

TAB. III.7 – Comparaison des performances des méthodes de vérification de parenté sur les bases de données Cornell et UB

Les performances données sur le tableau ci-dessus, montrent que notre approche surpassé les techniques les plus récentes sur les deux bases de données considérées.

III.7 Conclusion

Dans ce dernier chapitre, nous avons exposé les résultats obtenus suite à la mise en oeuvre de nos architectures présentées dans le chapitre précédent. Tout d'abord, nous avons détaillé la mise en oeuvre de notre système, en présentant l'environnement de développement sur lequel le système a été réalisé. Ensuite, nous avons analysé les résultats de chaque expérience en calculant les métriques d'évaluation, facilitant ainsi la comparaison entre les différentes expérimentations.

Les résultats obtenus ont permis de démontrer l'efficacité de la stratégie de fusion adoptée, confirmant sa capacité à améliorer la performance de notre système de vérification de la parenté faciale.

Conclusion Générale

Le travail réalisé dans ce mémoire consiste à mettre en place un système de vérification de la parenté faciale. L'objectif principal était de maîtriser les diverses étapes de traitement : prétraitement, extraction de caractéristiques, calcul des scores et fusion. Nous avons ensuite cherché comment améliorer les performances du système en intégrant de nouvelles techniques.

Pour de tels systèmes, l'environnement ainsi que les différentes sources de variabilité ont une influence majeure sur les performances. Le système de base, que nous avons développé, utilise MTCNN pour la détection des visages. Pour l'extraction des caractéristiques basés sur des descripteurs texturaux (shallow features), nous avons proposé deux descripteurs spécifique LPQ et BSIF. Nous avons ensuite appliqué une méthode optimisée appelée TXQEDA+WCCN pour la réduction de dimension et la compensation des variabilités.

L'évaluation des performances de notre système a été réalisée sur deux bases de données : Cornell KinFace et UB KinFace. Nous avons testé les descripteurs LPQ et BSIF à différentes échelles, et obtenu des taux de vérification corrects de 91.29 % et 87.23 % pour LPQ, et de 92.42 % et 85.36 % pour BSIF, respectivement sur les bases Cornell et UB KinFace.

Par la suite, nous avons remplacé nos descripteurs texturaux par une autre méthode d'extraction de caractéristiques basée sur des caractéristiques profondes (deep features), utilisant l'architecture InceptionV3 avec diverses configurations du système. Le même protocole a été conservé afin d'assurer une comparaison objective avec les descripteurs texturaux. Cette méthode a été appliquée aux mêmes bases de données. Les résultats obtenus étaient respectivement de 71.34 % pour Cornell KinFace, et 64.96 % pour UB KinFace.

Afin de tirer parti des avantages complémentaires des deux approches (profondes et texturaux), nous avons appliqué une fusion des scores au niveau du système. En utilisant une fusion par régression logistique (LR) avec trois extracteurs (InceptionV3, LPQ, BSIF), nous avons atteint des taux de vérification correcte de 93.75 % et 89.46 % respectivement pour les bases Cornell et UB KinFace.

En conclusion, nous estimons avoir développé un système répondant à nos objectifs initiaux. L'utilisation de la fusion de scores entre deux méthodes d'extraction de caractéristiques (profondes et texturales) permet d'obtenir une meilleure robustesse et

Conclusion Générale

d'améliorer significativement les performances du système de vérification de parenté.

Perspectives

Les travaux menés dans le cadre de ce projet représentent un bon début pour plusieurs autres expérimentations futures qui doivent être poursuivies afin de parvenir à un système encore plus robuste et performant.

À partir de ce projet, nous prévoyons de poursuivre nos recherches dans le domaine de la vérification de la parenté faciale, en explorant de nouvelles approches plus avancées telles que l'utilisation de réseaux neuronaux plus profonds, des techniques d'apprentissage auto-supervisé, ainsi que des méthodes de fusion plus intelligentes basées sur des modèles d'ensembles. Par ailleurs, nous envisageons d'étendre l'évaluation du système sur des bases de données plus variées, incluant des images dans des conditions réelles et non contrôlées.

Bibliographie

- [1] Kefeng Li and Quanzhen Huang. Cross-pose face recognition by integrating regression iteration and interactive subspace. *EURASIP Journal on Wireless Communications and Networking*, 2019(1) :105, 2019.
- [2] Huanjie He, Jiuzhen Liang, Zhenjie Hou, Hao Liu, and Xinwen Zhou. Occlusion recovery face recognition based on information reconstruction. *Machine Vision and Applications*, 34, 07 2023.
- [3] Joanna Isabelle Olszewska. Automated face recognition : Challenges and solutions. In S. Ramakrishnan, editor, *Pattern Recognition*, chapter 4. IntechOpen, Rijeka, 2016.
- [4] Wuming Zhang, Xi Zhao, Jean-Marie Morvan, and Liming Chen. Improving shadow suppression for illumination robust face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 41(3) :611–624, 2018.
- [5] Abdelhakim Chergui, Salim Ouchtati, Sébastien Mavromatis, Salah Eddine Be-khouche, Mohamed Lashab, and Jean Sequeira. Kinship verification through facial images using cnn-based features. *Traitemet du Signal*, 37(1) :1–8, 2020.
- [6] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE signal processing letters*, 23(10) :1499–1503, 2016.
- [7] Nafiz Shahriar. What is convolutional neural network—cnn (deep learning). *Electronic resource]. Available : <https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5>*, 2023.
- [8] Ruogu Fang, Kevin D Tang, Noah Snavely, and Tsuhan Chen. Towards computational models of kinship verification. In *2010 IEEE International conference on image processing*, pages 1577–1580. IEEE, 2010.
- [9] Miguel Bordallo Lopez, Abdenour Hadid, Elhocine Boutellaa, Jorge Goncalves, Vasilis Kostakos, and Simo Hosio. Kinship verification from facial images and videos : human versus machine. *Machine Vision and Applications*, 29:873–890, 2018.
- [10] Youssef Belabbaci El Ouanas, Ahmed Benlamoudi, and Mohamed-Chaker Larabi. Fusion of deep and shallow features for face kinship verification. *arXiv preprint arXiv :2312.10462*, 2023.

Bibliographie

- [11] Xiaoting Wu, Xiaoyan Feng, Xinyu Cao, et al. Facial kinship verification : A comprehensive review and outlook. *International Journal of Computer Vision*, 130(6) :1494–1525, 2022.
- [12] J. Paul Robinson, Ming Shao, Tingyu Wang, and Yun Fu. Families in the wild (fiw) : Large-scale kinship image database and benchmarks. In *Proceedings of the 2016 ACM International Conference on Multimodal Interaction (ICMI)*, pages 42–50. ACM, 2016.
- [13] Xiaoqian Qin, Dakun Liu, and Dong Wang. A literature survey on kinship verification through facial images. *Neurocomputing*, 377:213–224, 2020.
- [14] Guodong Guo and Xiaolong Wang. Kinship measurement on salient facial features. *IEEE Transactions on Instrumentation and Measurement*, 61(8) :2322–2325, 2012.
- [15] Simon Nash, Mark Rhodes, and Joanna Isabelle Olszewska. Ifr : Interactively pose corrected face recognition. In *International Conference on Bio-inspired Systems and Signal Processing*, volume 5, pages 106–112. SCITEPRESS, 2016.
- [16] Mohammed Almuashi, Siti Zaiton Mohd Hashim, Dzulkifli Mohamad, Mohammed Hazim Alkawaz, and Aida Ali. Automated kinship verification and identification through human facial images : a survey. *Multimedia Tools and Applications*, 76:265–307, 2017.
- [17] Hu Han, Shiguang Shan, Xilin Chen, and Wen Gao. A comparative study on illumination preprocessing in face recognition. *Pattern Recognition*, 46(6) :1691–1699, 2013.
- [18] Hamou Reda Mohamed and Ahmed Lehireche. La classification non supervisée (clustering) de documents textuels par les automates cellulaires. *University Djillali Liabes of Sidi Bel Abbes*, 2009.
- [19] Christian Schuldt, Ivan Laptev, and Barbara Caputo. Recognizing human actions : a local svm approach. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 3, pages 32–36. IEEE, 2004.
- [20] Xiuzhuang Zhou, Junlin Hu, Jiwen Lu, Yuanyuan Shang, and Yong Guan. Kinship verification from facial images under uncontrolled conditions. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 953–956, 2011.
- [21] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet : A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823, 2015.
- [22] Omkar Parkhi, Andrea Vedaldi, and Andrew Zisserman. Deep face recognition. In *BMVC 2015-Proceedings of the British Machine Vision Conference 2015*. British Machine Vision Association, 2015.
- [23] Joseph P Robinson, Ming Shao, Yue Wu, and Yun Fu. Families in the wild (fiw) large-scale kinship image database and benchmarks. In *Proceedings of the 24th ACM international conference on Multimedia*, pages 242–246, 2016.

Bibliographie

- [24] Jiwen Lu, Xiuzhuang Zhou, Yap-Pen Tan, Yuanyuan Shang, and Jie Zhou. Neighborhood repulsed metric learning for kinship verification. *IEEE transactions on pattern analysis and machine intelligence*, 36(2) :331–345, 2013.
- [25] Xiaoqian Qin, Xiaoyang Tan, and Songcan Chen. Tri-subject kinship verification : Understanding the core of a family. *IEEE Transactions on Multimedia*, 17(10) :1855–1867, 2015.
- [26] Ming Shao, Siyu Xia, and Yun Fu. Genealogical face recognition based on ub kinface database. In *CVPR 2011 workshops*, pages 60–65. IEEE, 2011.
- [27] Yann Le Cun, Lawrence D Jackel, Brian Boser, John S Denker, Hans Peter Graf, Isabelle Guyon, Don Henderson, Richard E Howard, and William Hubbard. Handwritten digit recognition : Applications of neural net chips and automatic learning. In *Neurocomputing : Algorithms, Architectures and Applications*, pages 303–318. Springer, 1990.
- [28] Asifullah Khan, Anabia Sohail, Umme Zahoor, and Aqsa Saeed Qureshi. A survey of the recent architectures of deep convolutional neural networks. *Artificial intelligence review*, 53:5455–5516, 2020.
- [29] Aeri Rachmad, Mohammad Syarief, Juniar Hutagalung, Suci Hernawati, Eka Mala Sari Rochman, and Yuli Panca Asmara. Comparison of cnn architectures for mycobacterium tuberculosis classification in sputum images. *Ingenierie des Systemes d'Information*, 29(1) :49, 2024.
- [30] Manli Sun, Zhanjie Song, Xiaoheng Jiang, Jing Pan, and Yanwei Pang. Learning pooling for convolutional neural network. *Neurocomputing*, 224:96–104, 2017.
- [31] Hidenori Ide and Takio Kurita. Improvement of learning for cnn with relu activation by sparse regularization. In *2017 international joint conference on neural networks (IJCNN)*, pages 2684–2691. IEEE, 2017.
- [32] Ouissal FADEL. Combinaison de classificateurs pour la reconnaissance des anomalies mammaires. 2023.
- [33] Sergey Ioffe and Christian Szegedy. Batch normalization : Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. pmlr, 2015.
- [34] Ville Ojansivu and Janne Heikkilä. Blur insensitive texture classification using local phase quantization. In *Image and Signal Processing : 3rd International Conference, ICISP 2008. Cherbourg-Octeville, France, July 1-3, 2008. Proceedings 3*, pages 236–243. Springer, 2008.
- [35] Juho Kannala and Esa Rahtu. Bsif : Binarized statistical image features. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR)*, pages 1363–1366. IEEE, 2012.
- [36] Oualid Laiadi, Abdelmalik Ouamane, Abdelhamid Benakcha, Abdelmalik Taleb-Ahmed, and Abdenour Hadid. Tensor cross-view quadratic discriminant analysis for kinship verification in the wild. *Neurocomputing*, 377:286–300, 2020.

Bibliographie

- [37] Andrew Hatch, Andreas Stolcke, and Barbara Peskin. Within-class covariance normalization for svm-based speaker recognition. In *Proceedings of Interspeech*, pages 1471–1474, 2006.
- [38] Mohcene Bessaoudi, Abdelmalik Ouamane, Mebarka Belahcene, Ammar Chouchane, Elhocine Boutellaa, and Salah Bourennane. Multilinear side-information based discriminant analysis for face and kinship verification in the wild. *Neurocomputing*, 329:267–278, 2019.
- [39] Lei Zhang, Qingyan Duan, David Zhang, Wei Jia, and Xizhao Wang. Advkin : Adversarial convolutional network for kinship verification. *IEEE transactions on cybernetics*, 51(12) :5883–5896, 2020.
- [40] Moumita Mukherjee and Toshanlal Meenpal. Binary cross coupled discriminant analysis for visual kinship verification. *Signal Processing : Image Communication*, 108:116829, 2022.
- [41] Aarti Goyal and Toshanlal Meenpal. Patch-based dual-tree complex wavelet transform for kinship recognition. *IEEE Transactions on Image Processing*, 30:191–206, 2020.