

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique

Université Batna 2
Faculté des mathématiques et d'informatique
Département d'informatique

Thèse

En vue de l'obtention du diplôme de
Doctorat en Informatique

Approches évolutives multi-biométriques
pour l'identification des personnes

Présentée Par
Aldjia BOUCETTA

Membres du jury :

Président de Jury :	Hamouma MOUMEN	MCA,	Université Batna 2
Rapporteur :	Kamal Eddine MELKEMI	Prof.,	Université Batna 2
Examineurs :	Mohamed-Khireddine KHOLLADI	Prof.,	Université El-Oued
	Mourad BOUZENADA	MCA,	Université Constantine 2

Résumé

Cette thèse porte sur l'étude des approches évolutionnaires pour la biométrie. Précisément, l'objectif est de proposer des approches multi-biométriques efficaces pour l'authentification de personnes basées sur les signatures biométriques.

Dans cette thèse, d'abord, nous proposons une nouvelle méthode de reconnaissance de la main qui combine les moments invariants d'Hu et les moments de Legendre, en tant que nouvel extracteur de caractéristiques.

Puis, nous proposons l'utilisation de PSO comme technique évolutive permettant de combiner la modalité du visage, de l'iris et de palmprint au niveau du score.

Dans ce système, nous utilisons trois algorithmes d'extraction de caractéristiques (transformation de Gabor, transformation DWT et transformation de Contourlet) pour extraire les principales caractéristiques biométriques. Ensuite, PCA, LDA et KFA sont utilisés pour la réduction de dimensionnalité des vecteurs de caractéristiques.

Finalement, nous proposons un système biométrique multimodal efficace basé sur le Deep Learning. Pour cela, nous explorons deux approches basées sur CNN(Convolutional Neural Network) en utilisant un modèle pré-entraînement téléchargé de la littérature (squeezeNet).

Dans le premier approche, nous essayons de classer les objets en fonction de leurs caractéristiques distinctes et prenons donc une décision de classification finale basée sur les autres décisions. Nous pouvons utiliser un élément du modèle, tel que l'extracteur de caractéristiques, ou le fine-tuning en l'adaptant à notre tâche au lieu d'imagenet avec 1000 classes.

Dans la deuxième approche, les vecteurs de caractéristiques sont extraits séparément, en utilisant squeezeNet pour extraire les caractéristiques profondes des modalités, nous prenons les caractéristiques de la couche 66 et nous combinons ces vecteurs pour former un vecteur de caractéristiques composites. Ensuite fait une classification à l'aide de SVM.

Mots clés : Biométrie multimodale ; visage ; Iris ; Palmprint ; Techniques évolutives ; Optimisation ; PSO ; GA ; Deep Learning ; CNN ; SqueezeNet.

Abstract

This thesis deals with the study of evolutionary approaches for biometrics. Specifically, the goal is to provide effective multi-biometrics approaches for authentication of people based on biometric signatures.

In this thesis, we first propose a new method of hand recognition that combines Hu's invariant moments and Legendre moments, as a new feature extractor.

Then, we propose the use of the Particle swarm optimization (PSO) as an evolutionary technique to combine the modality of the face, iris and palmprint at the score level.

In this system, we use three characteristic extraction algorithms (Gabor transformation, DWT transformation and Contourlet transformation) to extract the main biometric characteristics. Then, PCA, LDA and KFA are used for the dimensionality reduction of feature vectors.

Finally, we propose an effective multimodal biometric system based on Deep Learning. For this, we explore two approaches based on CNN (Convolutional Neural Network) using a pre-training model downloaded from the literature (squeezenet).

In the first approach, we try to classify the objects according to their distinct characteristics and thus take a final classification decision based on the other decisions. We can use an element of the model, such as the feature extractor, or fine-tuning by adapting it to our task instead of imagenet with 1000 classes.

In the second approach, the feature vectors are extracted separately, using squeezenet to extract the deep features of the modalities, we take the characteristics of the layer 66 and we combine these vectors to form a vector of composite features. Then done a classification using SVM.

Keywords : Multimodal biometry ; Face ; Iris ; Palmprint ; Evolutionary techniques ; Optimization ; PSO ; GA ; Deep Learning ; CNN ; SqueezeNet.

ملخص

تتناول هذه الرسالة دراسة المقاربات التطورية للقياسات الحيوية. على وجه التحديد ، يتمثل الهدف في توفير مقاييس فعالة متعددة القياسات الحيوية للمصادقة على الأشخاص بناءً على التوقيعات الحيوية.

في هذه الأطروحة، نقتراح أولاً طريقة جديدة للتعرف على اليد تجمع بين لحظات Hu المتغيرة ولحظات Legendre، باعتبارها مستخرجًا جديدًا للميزات.

بعد ذلك، نقتراح استخدام PSO كطريقة قابلة للتطوير لدمج طريقة الوجه والقزحية وكفة اليد على مستوى النتيجة.

في هذا النظام، نستخدم ثلاث خوارزميات لاستخراج المعالم (التحويل غابور ، تحول DWT والتحول كونتوريت) لاستخراج الخصائص الحيوية الرئيسية. بعد ذلك ، يتم استخدام PCA و LDA و KFA لتقليل الأبعاد لمتجهات الميزات.

أخيرًا، نقتراح نظامًا فعالاً متعدد المقاييس الحيوية يعتمد على التعلم العميق. لهذا ، نستكشف نهجين يعتمدان على CNN (الشبكة العصبية التلافيفية) باستخدام نموذج تمهيدي للتدريب (squeezeNet).

في النهج الأول، نحاول تصنيف الكائنات وفقًا لخصائصها المميزة واتخاذ قرار التصنيف النهائي بناءً على القرارات الأخرى. يمكننا استخدام عنصر النموذج، مثل مستخرج الخصائص، أو الضبط عن طريق تكييفه لمهمتنا بدلاً من imagenet مع 1000 فئة.

في الطريقة الثانية، يتم استخراج متجهات المعالم بشكل منفصل، باستخدام squeezeNet لاستخراج السمات العميقة، ونأخذ خصائص الطبقة 66 ونجمع هذه المتجهات لتشكيل متجه من الميزات المركبة. ثم القيام بتصنيف باستخدام SVM.

الكلمات المفتاحية:

القياسات الحيوية متعددة الوسائط; الوجه; القزحية; كفة اليد; التقنيات التطورية ; ; GA ; PSO التعلم العميق ; SqueezeNet ; CNN.

Remerciements

C'est avec un grand plaisir que je réserve ces lignes comme un signe de gratitude et de reconnaissance à tous ceux qui ont contribué, de près ou de loin, au bon déroulement de ma thèse. Qu'ils trouvent ici l'expression de mes sincères remerciements.

Mes remerciements les plus cordiaux sont adressés particulièrement à mon directeur de recherche le professeur Kamal Eddine MELKEMI, qui m'a encouragé dans ce travail, m'a aidé et dirigé dans mes recherches. Pour l'attention qu'il m'a apporté durant toute cette période ainsi que pour sa lecture et ses commentaires qui m'ont permis d'améliorer cette thèse.

Je tiens également à exprimer toute ma gratitude à Monsieur Hamouma MOUMEN doyen de la faculté des mathématiques et d'informatique de m'avoir fait l'honneur d'accepter la présidence de ce jury.

Mes sincères remerciements s'adressent aussi à Monsieur Mohamed-Khireddine KHOLLADI, professeur à l'université d'El-Oued et Monsieur Mourad BOUZENADA, MCA à l'université de Constantine 2 pour l'honneur qu'ils me font en acceptant d'examiner ce travail malgré leurs multiples tâches.

Dédicaces

Je remercie chaudement ma chère mère, mon père, mes chers frères et sœurs.

Surtout, je voudrais exprimer mes sincères remerciements à mon mari qui
m'a aidé et m'a soutenu pendant cette thèse.

A mes chers enfants, Asma, Selma et Saif Eddin pour leur amour.

Enfin, j'adresse mes plus sincères remerciements à tous mes proches et amis. Je remercie en particulier Zahia BENAMOR et Leila BOUSSAAD.

Table des matières

Liste des publications	xv
1 Introduction générale	1
2 Les systèmes biométriques	5
2.1 Introduction	5
2.2 Propriétés souhaitées pour un système biométrique	5
2.3 Modes de fonctionnement d'un système biométrique	6
2.3.1 Le mode d'enrôlement ou d'apprentissage	6
2.3.2 Le mode authentification ou vérification	7
2.3.3 Le mode identification	7
2.4 Principales modalités biométriques	8
2.4.1 La biométrie morphologique (physiologique)	8
2.4.2 La biométrie comportementale	12
2.4.3 La biométrie biologique	14
2.5 Évaluation d'un système biométrique	17
2.6 Limitations des systèmes biométriques unimodaux	19
2.7 Conclusion	20
3 Les systèmes biométriques multimodaux	21
3.1 Introduction	21
3.2 Les avantages des systèmes biométriques multimodaux	21
3.3 Les différentes formes de multi modalité	22
3.3.1 Multi-capteurs :	22
3.3.2 Multi-instances :	23
3.3.3 Multi-algorithmes :	23
3.3.4 Multi-échantillons :	23
3.3.5 Systèmes multi-biométries :	24
3.4 Les niveaux de fusion	24
3.4.1 La fusion pré-classification	24
3.4.2 La fusion post-classification	25
3.5 Normalisation	26
3.6 Les méthodes de fusion de score	29
3.6.1 Méthodes de fusion par combinaison de scores	29
3.6.2 Les méthodes de fusion par classification de scores :	36
3.7 Travaux récents sur biométrie multimodale	41
3.7.1 Fusion au niveau du capteur :	42
3.7.2 Fusion au niveau des caractéristiques :	42
3.7.3 Fusion au niveau des scores :	44
3.7.4 Fusion au niveau des décisions :	44

3.8	Conclusion	46
4	Identification des personnes à base de la main	47
4.1	Introduction	47
4.2	Principaux problèmes de la modalité de la main	47
4.3	Les traits biométriques de la main	49
4.3.1	La forme de la main	49
4.3.2	Géométrie de la main	50
4.3.3	L’empreinte de la main (palmprint)	51
4.3.4	Les doigts	53
4.4	Système proposé de reconnaissance de la forme de la main à l’aide des moments	53
4.4.1	Prétraitement d’image	54
4.4.2	Extraction des caractéristiques	55
	Moments Invariants de Hu	55
	Les moments de Legendre	58
4.4.3	Classificateur K-NN	59
4.5	Les résultats expérimentaux	60
4.5.1	Base de données	60
4.5.2	Résultats expérimentaux	60
4.6	Conclusion	61
5	Approche multibiométrie basée sur le PSO adaptatif	63
5.1	Introduction	63
5.2	Description de l’approche proposée	63
5.2.1	Pré-traitement	64
5.2.2	Techniques d’extraction de caractéristiques	64
5.2.3	Techniques de réduction de dimensionnalité	66
5.2.4	Calculer les scores	69
5.2.5	Combiner les scores	69
5.3	Résultats expérimentaux et discussion	70
5.4	Conclusion	73
6	Nouvelle approche multi-biométrique basée sur le deep learning	79
6.1	Introduction	79
6.2	Etat de l’art	79
6.3	Réseaux de neurones artificiels	81
6.3.1	Réseau neuronal à propagation avant ou feed-forward	84
6.3.2	Réseau de neurones récurrent (feedback network) :	84
6.4	Apprentissage profond (Deep Learning)	85
6.4.1	Définition de l’apprentissage profond	85
6.4.2	Domaines d’application de l’apprentissage profond :	86
6.4.3	Réseau de neurones convolutifs (CNN ou ConvNet) :	88
	Architecture du réseau de neurones convolutifs :	90
6.4.4	Créer et entraîner des modèles de Deep Learning	92
	L’entraînement à partir de zéro	92
	Apprentissage par transfert (Transfer Learning)	92
	Extraction de caractéristiques	93

6.5	Le système biométrique multimodal proposé	93
6.5.1	Première stratégie	93
6.5.2	La deuxième stratégie	95
6.6	Résultats expérimentaux et discussion	96
6.6.1	Base de données utilisées	96
6.6.2	Résultats expérimentaux	97
	Première stratégie :	97
	La deuxième stratégie :	105
6.6.3	Comparaisons de performance	107
6.7	Conclusions	108
	Conclusion générale	109
	Bibliographie	111

Table des figures

2.1	Enrôlement d'une personne dans un système biométrique (MORIZET, 2009)	7
2.2	Authentification d'un individu dans un système biométrique (FLYNN JAIN ANIL, 2008)	7
2.3	Identification d'un individu dans un système biométrique (CHARFI, 2017)	8
2.4	La géométrie de la main (JAIN, ROSS et PRABHAKAR, 2004)	9
2.5	Modalité Visage (JAIN, ROSS et PRABHAKAR, 2004)	10
2.6	(a) Détail de l'iris (b) localisation de l'iris (HONG MIN, LI HAN et PARK, 2015)	11
2.7	La reconnaissance de la rétine (BORAH, SARMA et TALUKDAR, 2015)	11
2.8	Image de la reconnaissance vocale (WANG et al., 2015)	12
2.9	Image signature (ANTAL, SZABO et TORDAI, 2018)	13
2.10	Images sur la démarche (BOLLE et PANKANTI, 1998)	14
2.11	Image de la frappe au clavier (JAIN, ROSS et PRABHAKAR, 2004)	14
2.12	Image du système configuration des veines (LI et JAIN, 2015)	17
2.13	Illustration du FRR et du FAR (MORIZET, 2009).	18
2.14	Courbe ROC (MANSFIELD et al., 2002).	19
2.15	Courbe CMS (CHARFI, 2017).	19
3.1	Les différentes formes de multi modalité (ALLANO et al., 2009).	23
3.2	Les différents niveaux de fusion.	25
3.3	Structure générale d'un GA (MITCHELL, 1998).	34
3.4	Principe de l'SVM : (a) SVM avec un noyau linéaire(b) SVM avec un noyau gaussien (c) SVM avec un noyau polynomial (EJARQUE et al., 2009).	38
3.5	Les projections de deux classes de points sur les axes principaux construits par la méthode LDA (KHIARI, 2016)	40
4.1	Quelques exemples de problèmes environnement (GUO et al., 2012).	48
4.2	Quelques exemples de difficultés d'acquisition.	48
4.3	Quelques exemples de problèmes d'occlusion.	49
4.4	Forme de la main.	49
4.5	Caractéristiques géométriques de la main (GUO et al., 2012).	51
4.6	L'empreinte de la main.	52
4.7	Cinq surfaces de doigts.	53

4.8	Schéma du système proposé.	54
4.9	Exemple d'image du processus de prétraitement.	55
4.10	(A) Image originale, (B) Image à échelle 50%, (C) Image avec rotation 45°.	57
4.11	(A) Image originale, (B) Image miroir.	58
4.12	Quelques exemples d'images de la main dans la base de données	60
5.1	Conception de base du système proposé.	64
5.2	CT d'une image d'iris jusqu'à 4 niveaux.	65
5.3	Trois niveaux de décomposition d'une image.	66
5.4	Courbes ROC des méthodes unimodales et multimodales utilisant gabor et PCA, LDA et KFA.	76
5.5	Courbes ROC des méthodes unimodales et multimodales utilisant DWT et PCA, LDA et KFA.	77
5.6	Courbes ROC des méthodes unimodales et multimodales utilisant CT et PCA, LDA et KFA.	78
6.1	Principe de fonctionnement d'un neurone artificiel.	82
6.2	Exemple d'un réseau de type perceptron multicouche.	85
6.3	Fusion tardive avec des canaux séparés.	94
6.4	Architecture de squeeze-net.	95
6.5	Fusion au niveau des caractéristique.	96
6.6	Exemples de classes vraies et prédites	98
6.7	Matrice de confusion et score de précision de visage.	99
6.8	Matrice de confusion et score de précision de l'iris.	100
6.9	Matrice de confusion et score de précision de palmprint.	100
6.10	Précision de l'entraînement de validation, et perte au fil des itérations de visage.	101
6.11	Précision de l'entraînement de validation, et perte au fil des itérations de l'iris.	102
6.12	Précision de l'entraînement de validation, et perte au fil des itérations de palmprint.	102
6.13	Exemples de classes vraies et prédites.	103
6.14	Matrice de confusion et score de précision de visage.	104
6.15	Matrice de confusion et score de précision de l'iris.	104
6.16	Matrice de confusion et score de précision de palmprint.	105
6.17	Matrice de confusion et score de précision de la deuxième stratégie.	106

Liste des tableaux

4.1	Moments des images (A), (B) et (c)	58
4.2	Performance de reconnaissance	61
4.3	Analyse de performance	61
5.1	Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant la transformation de Gabor et la PCA.	71
5.2	Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant la transformation de Gabor et la LDA.	72
5.3	Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant la transformation de Gabor et KFA.	72
5.4	Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant DWT Transform et PCA.	73
5.5	Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant DWT et LDA.	73
5.6	Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant DWT et KFA.	74
5.7	Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant CT et PCA.	74
5.8	Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant la CT et la LDA.	75
5.9	Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant CT et KFA.	75
6.1	Les résultats obtenus avec stratégie 1 et stratégie 2.	106
6.2	Comparaison du taux de reconnaissance du visage, de l'oreille et de leur fusion dans le scénario 1.	107
6.3	Comparaison du taux de reconnaissance du visage, de l'oreille et de leur fusion dans le scénario 2.	108

Liste des abréviations

PIN	Personal Identification Number
STR	Short Tandem Repeats
FRR	False Reject Rate
FAR	False Accept Rate
EER	Equal Error Rate
ROC	Receiver Operating Characteristics
AUC	Area Under the Curve
DET	Detection Error Tradeoff
CMS	Cumulative Match Score
GA	Genetic Algorithm
PSO	Particle Swarm Optimization
K-NN	K Nearest Neighbours
SVM	Support Vector Machine
ANN	Artificial Neural Network
LDA	Linear Discriminant Analysis
LPQ	Local Phase Quantization
SIFT	Scale Invariant Feature Transform
DCT	Discrete Cosine Transform
LVQ	Linear Vecteur Quantization
PCA	Principal Component Analysis
LBP	Local Binary Patterns
DS	Dempster Shafer
HTER	Half Total Error Rate
ROI	Region Of Interest
CCD	Charge Coupled Device
DFT	Discrete Fourier Transform
SC	Shape Context
IDSC	Inner Distance Shape Contexts
SMCC	Sparse Multiscale Competitive Code
KPBG	Key Point based Block Growing
LLDP	Local Line Directional Patterns
MFRAT	Modified Finite Radon Transform
mRMR	minimum Redundancy Maximum Relevance
KFA	Kernel Fisher discriminant Analysis
CT	Contourlet Transform
LP	Laplacian Pyramid
DFB	Directional Filter Banks
DCA	Discriminant Correlation Analysis
CNN	Convolutional Neural Networks

ADLVQ	Adaptive Deep Learning Vector Quantisation
LGPV	Local Gradient Pattern with Variance
DBN	Deep Belief Network
CCA	Canonical Correlation Analysis
ReLU	Rectified Linear Unit
MLP	Multi Layer Perceptron
GPU	Graphics Processing Unit

Liste des publications

1. Aldjia Boucetta, Kamal Eddine Melkemi : Score Level Fusion of Palm-print, Face and Iris Using Adaptive PSO. International Journal of Applied Metaheuristic Computing (IJAMC). Int. J. of Applied Metaheuristic Computing 10(3) : 175-194 (2019).
2. Aldjia Boucetta, Kamal Eddine Melkemi : Hand shape recognition using Hu and Legendre moments. SIN 2013 : 272-276.
3. Aldjia Boucetta, Kamal Eddine Melkemi : DWT Based-Approach for Color Image Compression Using Genetic Algorithm. ICISP 2012 : 476-484.

Introduction générale

De nos jours, en raison de l'expansion de la société, il est de plus en plus nécessaire de procéder à une vérification / identification de l'identité personnelle sécurisée et fiable en utilisant des moyens automatiques.

La nécessité d'un système fiable, simple, flexible et sécurisé est une préoccupation majeure et un enjeu difficile pour plusieurs applications rendant des services uniquement légitime inscrits utilisateurs (JAIN, ROSS et PRABHAKAR, 2004).

Parmi les applications de ce type, citons le partage des ressources informatiques en réseau, l'octroi d'un accès aux installations nucléaires, la réalisation de transactions financières à distance (téléachat) et le contrôle de l'accès physique.

Les méthodes traditionnelles d'établissements de l'identité d'une personne sont déjà largement utilisées dans le cadre de la vérification de l'identité. Ces méthodes sont basées sur quelque chose que vous connaissez (sécurité basée sur la connaissance), tel que les mots de passe, qui peuvent être partagés ou oubliés, ou quelque chose que vous possédez (sécurité basée sur un jeton), tel que des clés, des cartes magnétiques, des cartes d'identité et des codes PIN (Personal Identification Number), pouvant être partagés, volés, copiés ou perdus.

L'authentification biométrique (également appelée biométrie) est le moyen efficace pour remédier aux divers problèmes posés par les moyens d'authentification traditionnels, améliorer le niveau de sécurité et offrir plus de commodité et plusieurs avantages (FLYNN JAIN ANIL, 2008).

L'authentification biométrique est la reconnaissance automatique de la personne en fonction de ce qui vous êtes, elle fait référence à sa physiologie ou ce que vous produisez, elle fait référence à son comportement caractéristique.

Ses caractéristiques physiologiques distinctives incluent le visage, les empreintes digitales, la géométrie de la main, l'iris, la rétine, l'ADN, etc.

Les caractéristiques comportementales sont des actions réalisées par une personne de manière unique ; ils comprennent la signature, la frappe au clavier, la voix, etc. Ses caractéristiques sont appelées modalités biométriques ou traits.

Un système biométrique est à la base d'un système de reconnaissance de formes qui acquiert des données biométriques auprès de la personne, en extrayant les caractéristiques les plus significatives, en comparant ces caractéristiques à celles stockées dans la base de données, et prendre la décision finale en fonction du résultat de la comparaison (accepter / rejeter).

En règle générale, un système biométrique comporte deux étapes : l'inscription et la reconnaissance. L'inscription désigne l'étape à laquelle le système stocke les informations de référence biométriques sur la personne dans une base de données.

Au stade de la reconnaissance, le système analyse le trait biométrique de l'utilisateur, extrait les caractéristiques et les compare aux informations biométriques de référence stockées dans la base de données. Un score de similarité élevé entre la requête et les données de référence entraîne l'authentification ou l'identification de l'utilisateur.

Il est très important de disposer de critères couramment utilisés pour mesurer les performances des systèmes biométriques, afin que ces systèmes puissent être comparés et que les performances réelles puissent être estimées, et les progrès pourraient être motivés.

En biométrie, les performances sont basées sur la probabilité d'accepter des imposteurs, appelée FAR (False Acceptance Rate); et la probabilité de rejeter les utilisateurs authentiques, appelée FRR (False Rejection Rate). La courbe ROC (Receiver Operating Characteristic) et DET (Detection Error Trade-off) peuvent être utilisées pour une comparaison graphique des performances entre différents systèmes (KHIARI, 2016).

Pour une mesure empirique simple, le EER (Equal Error Rate) est généralement utilisé en biométrie, qui fait référence au point auquel le FRR et le FAR sont identiques à un seuil de décision donné.

Les systèmes biométriques qui utilisent une seule modalité biométrique (système biométrique unimodal ou monomodal) souffrent souvent de plusieurs limitations telles que la non-acceptabilité du public, le taux d'erreurs, la possibilité de l'intrusion de ces systèmes et autres problèmes. Ce qui a donné naissance à la biométrie multimodale qui se base sur la combinaison de diverses informations de différentes sources biométriques. Ces sources peuvent être différentes instances de la même modalité, des modalités biométriques différentes, plusieurs prototypes d'une modalité issues de plusieurs capteurs ou plusieurs informations issues de plusieurs algorithmes d'extraction de caractéristiques d'une seule modalité (NANDAKUMAR, 2008).

Ainsi, des études ont montré que ces systèmes biométriques multimodaux peuvent atteindre de meilleures performances que les systèmes monomodaux. Pour cette raison, les systèmes biométriques multimodaux conviennent à de nombreux domaines d'application.

L'identification biométrique basée sur multiples modalités représente une tendance émergente. La raison la plus convaincante pour combiner différentes modalités est d'améliorer le taux de reconnaissance. Cela peut être fait lorsque les caractéristiques biométriques des différentes modalités biométriques sont statistiquement indépendantes.

Il existe d'autres raisons de combiner deux modalités ou plus : la première est que différentes modalités biométriques pourraient être plus appropriées pour différentes applications. Une autre raison est simplement la préférence de l'utilisateur du système biométrique.

Dans ce contexte, l'objectif général de notre thèse est de proposer un système d'identification des personnes par la fusion de différentes modalités

biométriques.

Dans un système biométrique multimodal, la fusion peut être effectuée à trois niveaux différents : le niveau d'extraction des caractéristiques, la fusion au niveau du score et le niveau de décision (ATREY et al., 2010).

La fusion au niveau de l'extraction des caractéristiques combine différentes caractéristiques biométriques dans le processus de reconnaissance.

La fusion de scores fait correspondre les scores individuels de différents systèmes de reconnaissance pour obtenir un score multimodal.

Les systèmes au niveau de décision effectuent des opérations logiques sur les décisions du système monomodal pour parvenir à une résolution finale.

Il a toutefois été rapporté que l'approche la plus appropriée et la plus efficace pour la biométrie multimodale consiste à fusionner les données au niveau du score (NANDAKUMAR, 2008). Parce que fusionner des scores à cette étape permet un développement parallèle de chaque système unimodal et offre un bon compromis entre richesse d'information et facilité de mise en œuvre.

Étant donné que les scores correspondants issus des différentes modalités sont hétérogènes, la normalisation des scores est nécessaire pour transformer ces scores en un domaine commun, avant de les combiner (ATREY et al., 2010).

Les contributions de cette thèse se résument comme suit :

1. La reconnaissance des mains est l'objectif du premier système biométrique proposé. Nous avons présenté une nouvelle méthode de reconnaissance basée sur la forme de la main qui combine les moments invariants de Hu et les moments de Legendre.
2. La deuxième contribution de cette thèse réside dans le développement d'une nouvelle approche basée sur la fusion au niveau du score du système de reconnaissance multimodal de l'iris, du visage et de palmprint. Les trois scores sont fusionnés à l'aide de la règle de la somme pondérée et de PSO (Particle Swarm Optimization).
3. La troisième contribution de cette thèse est orientée vers l'apprentissage automatique (Machine learning) particulièrement, l'apprentissage par transfert (transfer learning). Nous utilisons un modèle pré-entraînement téléchargé de la littérature (squeezeNet) pour extraire les caractéristiques profondes des images biométriques telle que visages, iris et palmprint.

Ce mémoire de thèse comporte cinq chapitres :

Dans le premier chapitre, nous présentons les différentes notions de bases de la biométrie et un aperçu des caractéristiques d'un système biométrique ainsi que certaines modalités largement utilisées.

De plus, les différentes techniques utilisées pour l'évaluation de ces systèmes, les avantages et les inconvénients de la biométrie monomodale.

Le deuxième chapitre présente un aperçu sur les systèmes biométriques multimodaux en explicitant les différents outils, modes, niveaux de fusion proposés par les travaux de la littérature et en présentant quelques systèmes biométriques multimodaux existants.

Pour le troisième chapitre, nous présentons les différents systèmes d'identification des personnes à base de la main.

Nous proposons ensuite la combinaison des moments de Hu et Legendre, comme nouvel extracteur de caractéristiques pour un système de reconnaissance de la forme de la main. Les propriétés d'invariance en rotation et la position des moments invariants offrent une flexibilité au système en permettant à l'utilisateur de placer naturellement la main sans contrainte supplémentaire.

Le quatrième chapitre présente une fusion au niveau du score du système de reconnaissance multimodal de l'iris, du visage et de palmprint.

Les transformations Gabor, DWT (Discrete Wavelet Transform) et Contourlet sont utilisées pour l'extraction de caractéristiques, puis PCA (Principle Component Analysis), LDA (Linear Discriminant Analysis) et KFA (Kernel Fisher Analysis) pour réduire la taille des vecteurs de caractéristiques. Ensuite, les vecteurs de caractéristiques sont classés en utilisant K-NN (k-Nearest Neighbors). Enfin, les résultats des trois classificateurs sont fusionnés à l'aide de la règle de la somme pondérée et du PSO.

Dans le cinquième chapitre, nous présentons tout d'abord les notions en relation avec l'apprentissage profond (deep learning) tel que réseaux de neurones artificiels, les réseaux de neurones convolutifs.

Ensuite, nous proposons un système biométrique multimodal efficace basé sur le Deep Learning. Pour cela, nous explorons deux stratégies basées sur CNN(Convolutional Neural Network) en utilisant un modèle pré-entraînement téléchargé de la littérature (squeezenet).

Dans la première stratégie, nous avons fait une fusion au niveau de décision et dans la deuxième stratégie, les vecteurs de caractéristiques sont extraits séparément, en utilisant squeezenet pour extraire les caractéristiques profondes des modalités et nous combinons ces vecteurs pour former un vecteur de caractéristiques composites. Ensuite fait une classification à l'aide de SVM.

La thèse se termine en résumant les principaux résultats obtenus et en suggérant des lignes directrices et des recommandations pour les travaux futurs.

Chapitre 1

Les systèmes biométriques

1.1 Introduction

Au fil du temps, l'être humain a toujours essayé d'améliorer sa vie dans plusieurs domaines, notamment en matière de sécurité.

Avec le développement technologique rapide, la sécurité devient l'un des problèmes les plus préoccupants de notre société et pose un problème délicat aux citoyens, aux entreprises et au gouvernement dans la protection des informations sensibles et des données sensibles contre le vol. Pour toutes ses raisons, il est obligatoire de créer une nouvelle technique de contrôle, c'est le système biométrique qui constitue une solution efficace (PERRONNIN et DUGELAY, 2002).

La biométrie est un ensemble des techniques qui consiste à identifier un individu à partir de ses caractéristiques physiques ou comportementales. Ses caractéristiques sont traitées par certains ordres des processus automatisés à l'aide des dispositifs comme des modules de balayage ou des appareils photo. Contrairement à ce que l'on sait ou ce que l'on possède la biométrie est basée sur ce que l'on est permet ainsi d'éviter la duplication, le vol, l'oubli ou la perte.

Dans ce chapitre, nous soulignerons quelques concepts et définitions de base liés à la biométrie et à ses différentes techniques.

1.2 Propriétés souhaitées pour un système biométrique

Chaque caractéristique physiologique et/ou comportementale peut être utilisée comme biométrique pour reconnaître une personne tant qu'elle répond à ces exigences (JAIN, ROSS et PRABHAKAR, 2004) :

1. L'universalité : cela signifie que chaque personne devrait avoir cette caractéristique biométrique.
2. L'unicité : qui indique que la caractéristique biométrique est représentative d'un seul et unique individu.

3. La facilité de mesure : la facilité avec laquelle une personne peut interagir avec un système biométrique tout en étant identifiée ou authentifiée par ce système. Cette mesure est objective et dépend du type d'appareils utilisé pour capturer l'échantillon biométrique (caméra, scanner, etc.).
4. La performance : qui se réfère principalement à la précision de la reconnaissance, à la vitesse d'exécution et à la robustesse aux variations intrinsèques des caractéristiques biométriques, au bruit et aux déformations des dispositifs de capture. Les performances d'un système dépendent également des ressources nécessaires pour atteindre la vitesse et la précision attendues, ainsi que des facteurs opérationnels et environnementaux qui affectent la vitesse et la précision de la reconnaissance.
5. L'acceptation par l'utilisateur : reflète l'étendue des préoccupations et des objections que l'utilisation d'une technique biométrique donnée a tendance à générer. Dans certains pays, la reconnaissance faciale est mal acceptée; dans d'autres pays, la reconnaissance des empreintes digitales a des connotations criminelles. La mesure de l'acceptation reste très subjective et varie d'une personne à l'autre et d'un pays à l'autre, en fonction du régime de protection des données en vigueur, du contexte culturel et des attentes personnelles des utilisateurs.
6. Le contournement : reflète la facilité avec laquelle le système peut être trompé par des méthodes frauduleuses.
7. La permanence ou la stabilité : elle est définie par la constance d'une caractéristique biométrique au cours du développement normal et du vieillissement d'une personne.

En principe, plus une caractéristique est stable, moins il est nécessaire de mettre à jour les caractéristiques personnelles ou de réenregistrer la personne (PERRONNIN et DUGELAY, 2002).

1.3 Modes de fonctionnement d'un système biométrique

Un système biométrique peut fonctionner en modes suivants :

1.3.1 Le mode d'enrôlement ou d'apprentissage

L'enrôlement de personnes est la phase initiale de création du modèle biométrique et de son stockage en relation avec une identité déclarée (voir figure 2.1) (MORIZET, 2009).

Les caractéristiques physiques sont transformées en un modèle représentatif de la personne et spécifique au système de reconnaissance.

Pendant cette phase, des données supplémentaires spécifiques à la personne enregistrée sont enregistrées, par exemple son nom et prénom et un identifiant personnel (PIN). Cette étape est effectuée une seule fois.

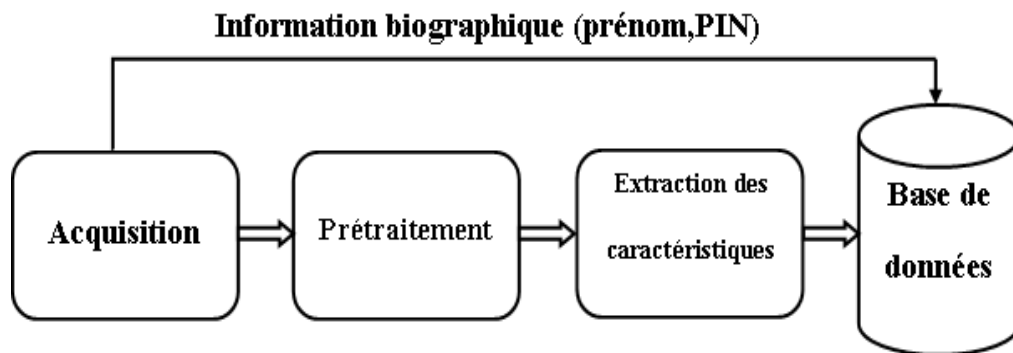


FIGURE 1.1 – Enrôlement d’une personne dans un système biométrique (MORIZET, 2009)

1.3.2 Le mode authentification ou vérification

Le système valide l’identité d’une personne en comparant les données biométriques capturées avec sa propre base de données (FLYNN JAIN ANIL, 2008). Dans tel système, une personne souhaitant être identifiée réclame une identité, généralement via un code PIN (Personal Identification Number), un nom d’utilisateur, une carte à puce, etc. Le système procède à une comparaison individuelle pour déterminer si la réclamation est vraie ou fausse. Ce mode de test s’appelle un test en ‘un contre un’ (voir figure 2.2).

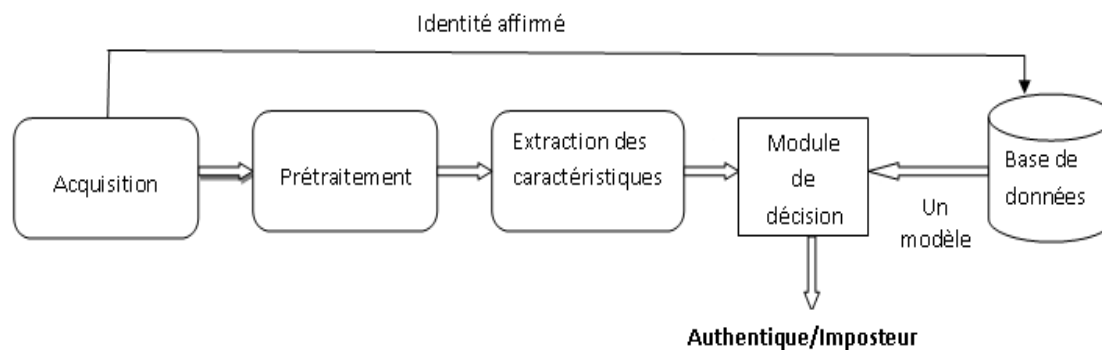


FIGURE 1.2 – Authentification d’un individu dans un système biométrique (FLYNN JAIN ANIL, 2008)

1.3.3 Le mode identification

Le système identifie un individu en recherchant les templates de tous les utilisateurs de la base de données. Par conséquent, le système effectue plusieurs comparaisons pour établir l’identité d’un individu ou échoue si le sujet n’est pas enregistré dans la base de données du système. Ce mode de test s’appelle un test en ‘un contre N’ (voir figure 2.3).

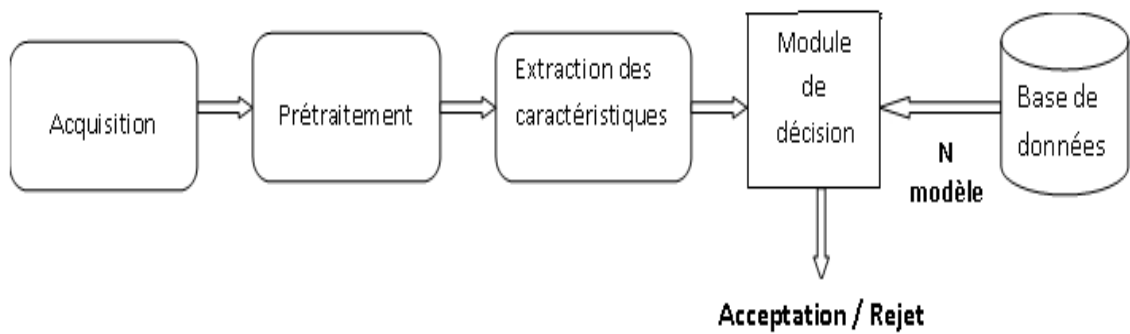


FIGURE 1.3 – Identification d’un individu dans un système biométrique (CHARFI, 2017)

1.4 Principales modalités biométriques

L’utilisation de la biométrie pour l’identification et l’authentification est une technique très efficace, ainsi qu’un concept fondamental lié à la reconnaissance des individus en fonction de leurs caractéristiques biométriques.

Il existe plusieurs techniques biométriques utilisées dans diverses applications, on peut distinguer trois catégories :

1.4.1 La biométrie morphologique (physiologique)

Elle est basée sur les caractéristiques physiques particulières qui, pour tout le monde, sont uniques et permanentes. Des exemples de ces traits sont présentés dans les sections suivantes :

1. L’empreinte digitale :

la reconnaissance d’empreintes digitales est l’une des premières techniques biométriques basées sur le fait que chaque personne a des empreintes digitales uniques (MALTONI et al., 2009).

Cependant, les empreintes digitales sont des mesures biométriques mal acceptées par les utilisateurs en raison de l’association souvent établie avec la criminologie.

Les empreintes digitales sont formées par les crêtes (ridge) et les vallées (furrow) présentes à la surface du bout des doigts.

Les empreintes sont différentes pour chaque doigt d’une même personne (CAO et JAIN, 2015).

L’image d’empreinte digitale d’un individu est capturée à l’aide d’un lecteur d’empreintes digitales, puis les entités sont extraites de l’image, puis un modèle est créé. Si les précautions appropriées sont suivies, le résultat est un moyen d’authentification très précis.

Les techniques d’appariement d’empreintes digitales peuvent être classées en deux catégories : les techniques basées sur la détection de minutie locale et les techniques basées sur la corrélation. L’approche basée sur les minuties consiste d’abord à trouver les points caractéristiques puis à tracer leur emplacement sur l’image du doigt.

Cependant, cette approche présente certaines difficultés lorsque l'image numérique d'impression est de qualité médiocre, car l'extraction précise des points minutie est difficile. Cette méthode ne prend pas en compte la structure globale des crêtes et des sillons.

Les méthodes basées sur la corrélation sont capables de surmonter les problèmes de l'approche basée sur les minuties. Ces méthodes utilisent la structure globale de l'empreinte, mais les résultats sont moins précis que les minuties.

De plus, les techniques de corrélation sont affectées par la translation et la rotation de l'image de l'empreinte digitale. C'est pourquoi les deux approches sont généralement combinées pour augmenter les performances du système.

2. La géométrie de la main :

La géométrie de la main consiste à analyser et à mesurer la forme de la main, c'est-à-dire mesurer la longueur, la largeur et la hauteur de la main d'un utilisateur et créer une image 3D (voir figure 2.4). Des LEDs infrarouges et un appareil-photo numérique sont utilisés pour acquérir les données de la main (CHARFI, 2017).

Cette technique offre un niveau de précision raisonnable et relativement facile à utiliser (PERRONNIN et DUGELAY, 2002).

Cependant, il peut être facilement trompé par des jumeaux ou par des personnes avec des formes de mains proches. Les utilisations les plus courantes de la géométrie de la main comprennent l'enregistrement de présence et le contrôle d'accès.

D'autre part, les systèmes de capture de la géométrie de la main sont relativement volumineux et lourds, ce qui limite leur utilisation dans d'autres applications telles que l'authentification dans les systèmes embarqués : téléphones portables, voitures, ordinateurs portables, etc.

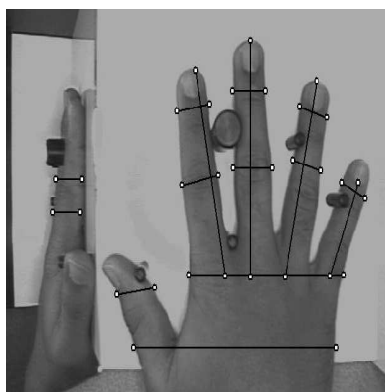


FIGURE 1.4 – La géométrie de la main (JAIN, ROSS et PRABHAKAR, 2004)

3. Le visage :

L'authentification faciale est la méthode la plus naturelle qui associe l'homme à l'interaction visuelle (voir figure 2.5) (GUERFI, 2008).

Récemment, différents types de caméras et d'appareils photo, de qualité et de coût variables, sont apparus sur le marché, permettant d'adapter la qualité des images aux conditions d'utilisation. Pourtant, aujourd'hui, ce n'est que lorsque le sujet est fixe et que les conditions environnementales sont normales (arrière-plan uniforme, éclairage suffisant) que les systèmes informatiques donnent de bons résultats.

Si l'acquisition se déroule dans un environnement naturel, sans contraintes imposées, les performances se dégradent considérablement car les variations personnelles (lunettes, chapeaux, moustaches) ou environnementales (éclairage, éloignement) sont encore mal prises en compte par les systèmes informatiques. Dans ce cas, une série de prétraitements sont souvent nécessaires avant de procéder à la reconnaissance (JAIN, ROSS et PRABHAKAR, 2004).



FIGURE 1.5 – Modalité Visage (JAIN, ROSS et PRABHAKAR, 2004)

4. L'iris :

L'iris est la région annulaire entre la pupille et le blanc de l'œil (CHARFI, 2017) (voir figure 2.6). La reconnaissance par l'iris est une méthode sûre, car celle-ci est toujours différente (même entre jumeaux, entre l'œil gauche et l'œil droit) et reste stable au cœur de la vie.

Cependant, il est obligatoire pour les utilisateurs de placer leurs yeux très près de l'appareil d'acquisition (ELREFAEI et al., 2018).

Son seul inconvénient est son coût relativement élevé, ce qui ne répond pas autant aux applications quotidiennes. Ainsi, son utilisation a été limitée dans les endroits où la sécurité est primordiale et même critique, comme dans les bases nucléaires par exemple.

La reconnaissance par l'iris est utilisée aussi dans le secteur financier pour les employés et les clients, dans les hôpitaux et dans les grands aéroports (HONG MIN, LI HAN et PARK, 2015). Une personne voulant s'identifier place son œil à quelques centimètres du capteur et l'image de l'iris est prise par une caméra. Ensuite, les caractéristiques sont extraites de l'image de l'iris et comparées à celles enregistrées dans la base de données.

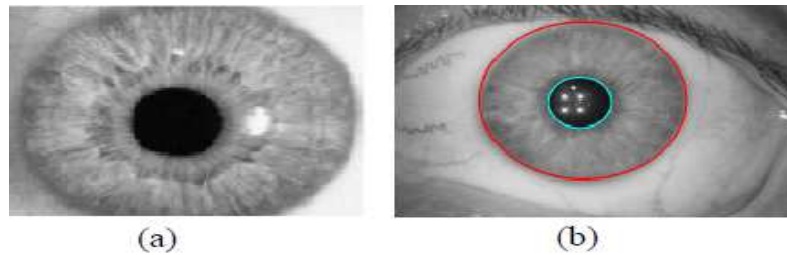


FIGURE 1.6 – (a) Détail de l’iris (b) localisation de l’iris (HONG MIN, LI HAN et PARK, 2015)

5. La rétine :

La rétine est une fine couche de cellules située à l’arrière du globe oculaire des vertébrés. C’est la partie de l’œil qui convertit la lumière en signaux nerveux (voir figure 2.7). Le principe de la biométrie rétinienne consiste à capturer et analyser les motifs des vaisseaux sanguins sur le nerf mince à l’arrière du globe oculaire, qui traite la lumière pénétrant dans la pupille.

La biométrie de la rétine fournit également un haut niveau de reconnaissance. Cette technique est utilisée pour des applications de très haute sécurité : par exemple, dans les applications militaires ou nucléaires (BORAH, SARMA et TALUKDAR, 2015).

Les caractéristiques de la rétine sont liées à la configuration géométrique des vaisseaux sanguins. La technique utilise un équipement spécialisé et un rayon illumine l’arrière de l’œil.

Les systèmes identifient jusqu’à cent quatre vingt douze points de référence. Certains risques pour la santé ont été révélés et limitent l’utilisation de cette technique à des locaux très sensibles (EKKA, PUHAN et PANDA, 2015).

Les systèmes d’acquisition de rétine sont coûteux. L’œil doit être situé très près de la tête de lecture et l’utilisateur doit regarder un point spécifique pendant plusieurs secondes. Cette méthode est donc mal acceptée par le grand public.

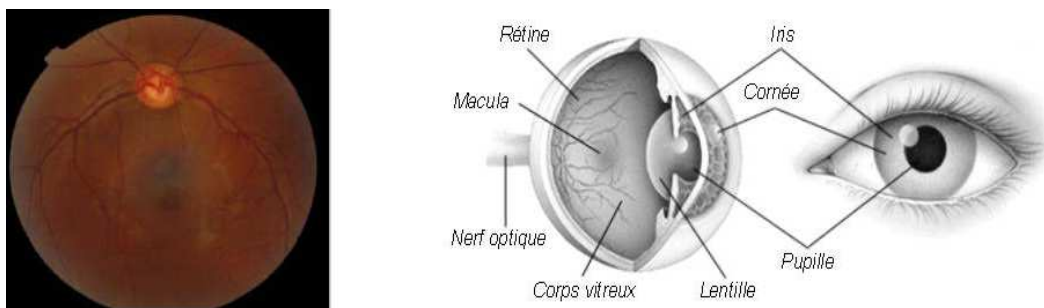


FIGURE 1.7 – La reconnaissance de la rétine (BORAH, SARMA et TALUKDAR, 2015)

1.4.2 La biométrie comportementale

Elle repose sur l'analyse de certains comportements d'une personne comme la trace de sa signature (inclinaison et vitesse de déplacement du stylo, pression exercée), l'empreinte de sa voix, sa démarche et sa façon de taper sur un clavier (vitesse de frappe).

1. La voix :

Grâce à cette modalité, nous pouvons analyser et reconnaître la voix humaine. Il est capturé via un microphone pour permettre la transcription sous la forme d'un texte exploitable par la machine (voir figure 2.8) .

Ensuite, plusieurs caractéristiques résultant de cette modalité (comme la tonalité, la fréquence, l'harmonique, la vitesse et le rythme, etc.) sont extraites afin de les comparer avec celles déjà stockées dans une base de données pour affirmer ou nier l'identité d'un locuteur. (SAINI et KAUR, 2013). Un téléphone ou un microphone peut être utilisé comme



FIGURE 1.8 – Image de la reconnaissance vocale (WANG et al., 2015)

appareil d'acquisition, ce qui rend cette technique relativement peu coûteuse et facilement réalisable.

Cependant, elle est sensible aux changements de voix dus à l'âge, au froid, etc. En outre, le système de reconnaissance vocale peut tomber en panne dans des environnements bruyants (faible rapport signal/bruit). Il peut également être sensible aux attaques (par ex. réinjection du signal vocal vers le système)(JAIN, ROSS et PRABHAKAR, 2004).

Cette modalité est utilisée dans de nombreuses applications, notamment les applications de dictée vocale par ordinateur, la répression par la police, les agences d'espionnage et la téléphonie.

Mais cette modalité trouve des limites en raison de la grande différence entre le langage formel, qui inclut et utilise les machines, et le langage naturel que les humains utilisent. Le défi consiste à trouver un compromis entre ces deux langues (WANG et al., 2015).

2. La signature :

L'identification par la signature est une méthode automatique de mesure des signatures des personnes.

Cette technique est considérée parmi les premières utilisées dans le domaine de la biométrie. Elle se base généralement sur le fait que l'utilisateur signe avec un stylo électronique sur une palette graphique et au même temps elle examine l'ensemble de dynamique comme la vitesse, la direction, et la pression de l'écriture, le temps pendant lequel le stylo est en contact avec le papier, le temps pris pour faire la signature et les positions où le stylo est relevé et abaissé sur le papier (ANTAL, SZABO et TORDAI, 2018).

La signature est une biométrie comportementale qui change sur une période de temps et qui est influencée par les conditions physiques et émotionnelles des signataires. Les signatures de certaines personnes varient considérablement.



FIGURE 1.9 – Image signature (ANTAL, SZABO et TORDAI, 2018)

3. La démarche :

La façon dont une personne marche peut la distinguer des autres. Dans un système de reconnaissance par cette modalité, nous cherchons à identifier un individu par sa façon de marcher et de bouger tout en analysant des images vidéo de la marche du candidat (CONNOR et ROSS, 2018) (voire figure 2.10). Il s'agit donc d'un mode d'identification à distance. Les gens montrent différents traits en marchant, tels que le maintien du corps, la distance entre les deux pieds, la position des articulations comme les genoux et les chevilles et les angles de pivotement (BOLLE et PANKANTI, 1998), ce qui contribue grandement à les identifier.

Ce mode est particulièrement adapté aux applications de vidéosurveillance. Les performances des systèmes basés sur la démarche ne sont pas suffisamment acceptables car elles sont affectées par plusieurs facteurs, notamment le choix des chaussures, la nature des vêtements, l'atteinte des jambes, la surface de marche, etc.

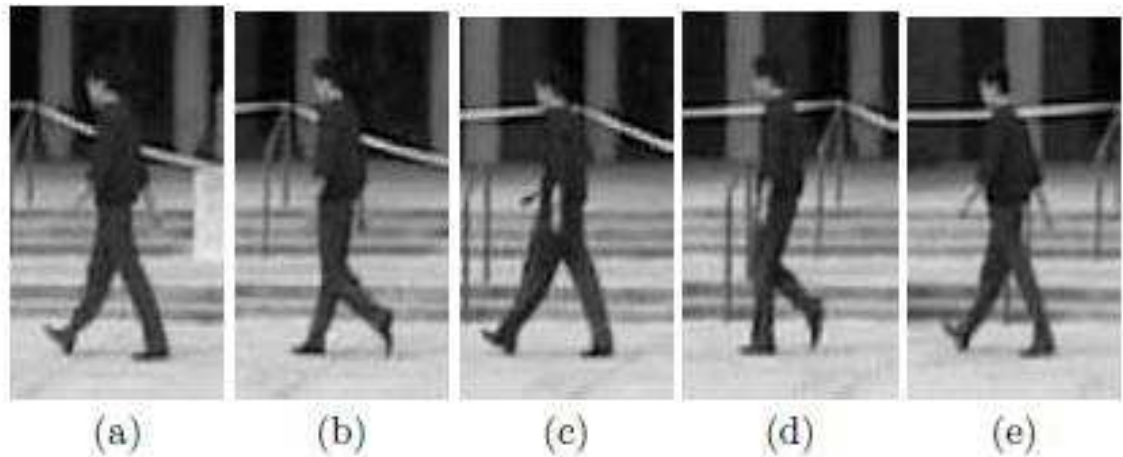


FIGURE 1.10 – Images sur la démarche (BOLLE et PANKANTI, 1998)

4. Dynamique de la frappe au clavier :

La dynamique de la frappe au clavier est une caractéristique de l'individu, c'est la transposition de la graphologie à des moyens électroniques (voir figure 2.11). C'est une méthode basée sur la manière d'utiliser ou de taper sur un clavier. Les paramètres pris en compte sont généralement les durées entre les frappes, la fréquence des erreurs et la durée de la frappe. D'autre part, cela dépend de l'état physique et psychologique de la personne (âge, maladie...). Cela variera la qualité de la frappe (JAIN, ROSS et PRABHAKAR, 2004).



FIGURE 1.11 – Image de la frappe au clavier (JAIN, ROSS et PRABHAKAR, 2004)

1.4.3 La biométrie biologique

Ces méthodes biométriques basées sur l'ADN, l'odeur corporelle, la forme de l'oreille, la forme des lèvres, veines de la main, etc.

1. L'ADN :

Est un fluide biologique qui est analysé par des méthodes simples telles que l'analyse du groupe sanguin, de différentes protéines ou enzymes. La plupart de ces analyses ont été rapidement abandonnées au profit du profilage ADN, également appelé empreint génétique, compte tenu de sa forte discrimination et de sa robustesse. L'analyse d'ADN la plus courante repose sur des séquences répétées en tandem courtes, également appelées microsatellites ou STR (Short Tandem Repeats), qui ne font pas partie des parties codant des protéines et ont des caractères distinctifs. L'ADN constitue une avancée majeure dans le domaine de la criminalistique pour l'identification des personnes inconnues ou dans le but de déterminer la source des échantillons biologiques laissés sur les lieux du crime (LI et JAIN, 2015).

2. L'odeur corporelle :

Chaque corps humain dégage une odeur caractérisée sa composition chimique et qui pourrait être utilisée pour distinguer différents individus (INBAVALLI et NANDHINI, 2014). La première odeur d'une personne contient des constituants stables dans le temps indépendamment de l'alimentation ou des facteurs environnementaux. Seconde odeur contient des constituants qui sont présents en raison de l'alimentation et facteurs environnementaux. L'odeur troisième contient des constituants qui sont présents en raison de l'influence de sources extérieures (c'est-à-dire, lotions, savons, parfums).

Pour une identification individuelle par odeur humaine, l'odeur première doit avoir des constituants stables dans le temps. Les composés présents dans les extraits de sécrétions axillaires mâles et femelles contenant les odeurs caractéristiques présentes dans la région axillaire ont été isolés et identifiés.

3. La forme de l'oreille :

La reconnaissance de l'oreille humaine est une nouvelle technique biométrique. Le criminologue français Bertillon a été le premier à suggérer que les personnes puissent être identifiées par la forme de leur oreille externe. Un peu plus tard, le policier américain Iannarelli a proposé le premier système de reconnaissance auditif basé sur douze caractéristiques. Iannarelli a expérimentalement constaté que dix mille oreilles étaient différentes même chez des jumeaux identiques (IANNARELLI, 1964). La biométrie de l'oreille a été utilisée dans de nombreux systèmes gouvernementaux tels que la médecine légale, la sécurité et les applications de la loi (GHOUALMI, DRAA et CHIKHI, 2016). En fait, l'oreille a suscité l'intérêt de la communauté biométrique en raison des avantages suivants :

Premièrement, sa taille permet d'accélérer la tâche de reconnaissance et d'accroître son efficacité. Deuxièmement, l'oreille a une distribution uniforme des couleurs, ce qui garantit la conservation des informations pertinentes lors de la conversion en une image en niveaux de gris.

Troisièmement, l'oreille n'a pas besoin de beaucoup la collaboration de l'utilisateur (GHOUALMI, DRAA et CHIKHI, 2016). La biométrie auriculaire est donc un bon choix car elle offre un bon compromis entre précision et coût.

La biométrie auriculaire a été utilisée dans de nombreuses applications gouvernementales et commerciales, telles que la médecine légale et la sécurité. Par exemple, aux États-Unis, un système de classification des oreilles fondé sur des mesures manuelles est utilisé depuis plus de 40 ans. En outre, le Service de l'immigration et de la naturalisation des États-Unis donne des spécifications indiquant que l'oreille droite doit être visible (GHOUALMI, DRAA et CHIKHI, 2016).

4. La forme des lèvres :

La lèvre est l'organe sensoriel tactile constituant la partie visible de la bouche. Des études ont montré que les sillons des lèvres humaines sont uniques à chaque personne et peuvent donc être utilisés pour l'identification humaine (RESHMI et KARTHICK, 2013).

Les caractéristiques de la forme des lèvres ont été largement utilisées dans la médecine légale par des experts et par la loi pour l'identification humaine.

Le défi d'utiliser la lèvre comme biométrie réside dans le domaine de l'unicité et du contournement. L'utilisation de la lèvre comme moyen d'identification de l'homme a été la première proposée à travers le concept de tirages sur les lèvres.

5. La biométrie des veines :

On a longtemps pensé que le modèle de la veine dans l'anatomie humaine pouvait être unique aux individus (figure 2.12).

En conséquence, le balayage veineux au fil des ans a connu plusieurs succès, grâce à la numérisation des réseaux de poignets et, plus récemment, au balayage des doigts. Cette technique utilise un "scanner de réseau veineux palmaire", pour être identifié, il est nécessaire de placer la zone affectée au-dessus du lecteur. C'est le modèle utilisé par le réseau de veines pour conserver les points caractéristiques (LI et JAIN, 2015).

La biométrie des veines est considérée comme une modalité fiable et robuste car la structure globale des veines ne changent pas avec le temps (à l'exception de certains changements dus à la dilatation des veines ou à des cas pathologiques). Le "Spoofing" de systèmes biométriques utilisant des veines peut être difficile, mais des expériences récentes ont montré que la modalité peut être imparfaite, en particulier lorsque l'identification et l'authentification sont effectuées en mode non supervisé (absence de contrôle humain).

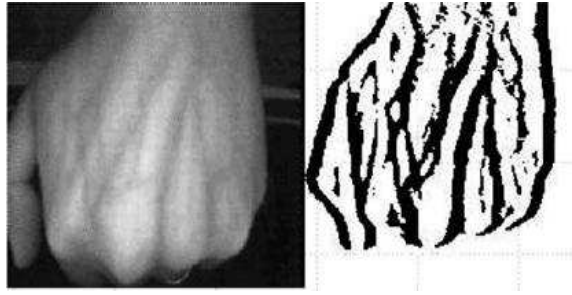


FIGURE 1.12 – Image du système configuration des veines (LI et JAIN, 2015)

1.5 Évaluation d'un système biométrique

L'évaluation des systèmes biométriques est un défi majeur en biométrie pour plusieurs raisons. Tout d'abord, il permet aux chercheurs de mieux tester et évaluer leurs systèmes avec ceux de la littérature. En conséquence, le comportement des utilisateurs peut être pris en compte lors du processus d'évaluation. De plus, elle permet d'identifier, pour chaque système, les applications industrielles basées sur ces performances.

Dans la littérature, il existe plusieurs métriques et plusieurs types de courbes (MORIZET, 2009) pour définir les performances d'un système biométrique, voici quelques-unes les plus utilisées :

1. Le taux de faux rejets ("False Reject Rate" ou FRR) : ce taux détermine la probabilité qu'un système ne reconnaisse pas une personne qui aurait normalement dû être reconnue. C'est un rapport entre le nombre de personnes légitimes dont l'accès a été refusé et le nombre total de personnes légitimes qui se sont manifestées.
2. Le taux de fausse acceptation ("False Accept Rate" ou FAR) : ce taux détermine la probabilité qu'un système reconnaisse une personne qui normalement n'aurait pas dû être reconnue. C'est un rapport entre le nombre de personnes qui ont été acceptées lorsqu'elles n'auraient pas dû l'être et le nombre total de personnes non autorisées qui ont tenté d'être acceptées (EL-ABED, 2011).
3. Le taux d'égal erreur ("Equal Error Rate" ou EER) : ce taux est calculé à partir des deux premiers critères et constitue un point de mesure des performances actuelles. Ce point correspond à l'endroit où $FRR = FAR$, c'est le meilleur compromis entre les faux rejets et les fausses acceptations.

La figure 2.13 illustre le FRR et le FAR à partir des distributions des scores authentiques et des imposteurs tandis que l'EER est représenté sur la figure 2.14.

4. Taux d'identification (identification rate, IR) : est une estimation de la probabilité qu'un sujet soit identifié correctement au moins au rang-k. Par conséquent, le taux d'identification est nécessairement une fonction croissante de k. Un échantillon de test se voit attribuer un rang

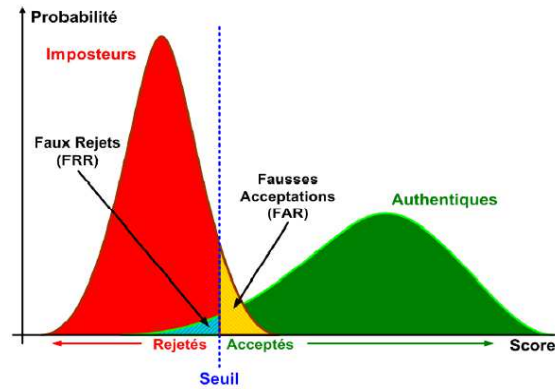


FIGURE 1.13 – Illustration du FRR et du FAR (MORIZET, 2009).

k lorsque le sujet réel est classé en position k par un système d'identification. Ainsi, un résultat de rang 1 est considéré comme une identification correcte. Il peut être formellement défini sur l'ensemble des images de test P comme suit (EL-ABED, 2011) :

$$IR = \frac{R1}{N} \quad (1.1)$$

Où $R1$ est le nombre des images de test correctement reconnues en L à $k = 1$ et N est la taille de l'ensemble de tests.

Nous allons maintenant définir trois types de courbes de performance :

1. La courbe ROC (Receiver Operating Characteristics) : cette courbe représente en ordonnée la proportion de tests positifs parmi les utilisateurs authentique en fonction de la proportion de tests positifs parmi les imposteurs pour toutes les valeurs des seuils de test envisageables. Pour pouvoir déterminer la validité d'un test à travers cette courbe, il est nécessaire de calculer l'aire sous la courbe ROC appelée AUC (Area Under the Curve). Plusieurs méthodes ont été proposées dans (MANSFIELD et al., 2002) pour estimer l'AUC. Ainsi, lorsque le test est parfaitement discriminant, l'aire sous la courbe (AUC) est égale à 1, mais cela n'est jamais possible. En réalité, plus l'AUC est grande, l'algorithme est meilleur. La figure 2.14 illustre un exemple de courbe ROC.

Les courbes ROC sont indépendantes du seuil, permettant la comparaison des performances de différents systèmes dans des conditions similaires, ou d'un système unique dans des conditions différentes.
2. La courbe DET (Detection Error Tradeoff) : une courbe DET trace les taux d'erreur sur les deux axes, donnant un traitement uniforme aux deux types d'erreurs (MORIZET, 2009). Le graphique peut alors être tracé en utilisant des axes logarithmiques.
3. La courbe de scores cumulés CMS (pour Cumulative Match Score) :

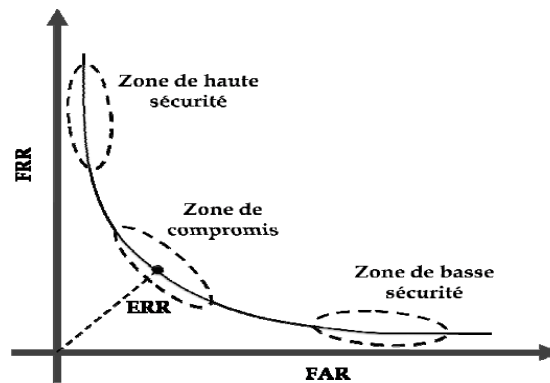


FIGURE 1.14 – Courbe ROC (MANSFIELD et al., 2002).

cette courbe (voir figure 2.15) donne le pourcentage de personnes reconnues selon une variable appelée rang. On dit qu'un système reconnaît au rang 1 lorsqu'il choisit l'image la plus proche suite à la reconnaissance. On dit qu'un système reconnaît au rang 2, lorsqu'il choisit parmi deux images, celle qui correspond le mieux à l'image d'entrée (CHARFI, 2017).

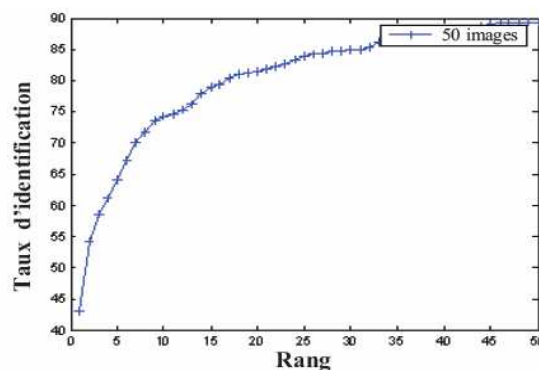


FIGURE 1.15 – Courbe CMS (CHARFI, 2017).

1.6 Limitations des systèmes biométriques unimodaux

Bien que les techniques de reconnaissance biométrique soient très efficaces, nous ne pouvons actuellement garantir un excellent taux de reconnaissance avec les systèmes biométriques unimodaux, basés sur une signature biométrique unique. De plus, ces systèmes sont souvent affectés par les problèmes suivants :

1. Les limites dues à la personne (rhume pour la voix, blessure au doigt etc.).
2. Les capteurs défectueux ou encrassés (faible luminosité, contraste, ...).
3. Positionnement des capteurs/taille des personnes.

4. Variations intra-classes : causée par une personne qui interagit incorrectement avec capteur et cela augmentera FRR.
5. Similitudes intra-classe : désigne le chevauchement des espaces caractéristiques correspondant à plusieurs classes ou individus. Cela peut augmenter le FAR.
6. Cas particulières
 - Faible ouverture de l'œil
 - Forte ouverture de la pupille
7. Non universel : le système biométrique peut ne pas être en mesure d'acquérir des données biométriques significatives d'un sous-ensemble d'utilisateurs.
8. Imitations, faussaires.

1.7 Conclusion

La biométrie devient de plus en plus un outil de reconnaissance des personnes dans diverses applications. Elle gagne sa place en tant que moyen numéro un d'authentification. Cependant, un certain nombre de défis, tels que les attaques contre les systèmes biométriques, restent à surmonter. Malgré cela, l'avenir de la biométrie est prometteur pour l'authentification et l'identification des personnes.

Ce chapitre a été consacré à la présentation de généralités sur la biométrie monomodale . C'est un chapitre d'introduction pour les chapitres suivants.

Nous avons donc introduit les différents concepts de ce domaine biométrique et un aperçu des caractéristiques d'un système biométrique ainsi que certaines modalités largement utilisées ont été données.

De plus, les différentes techniques utilisées pour l'évaluation de ces systèmes où il a été possible d'affirmer que le taux de faux rejets ou d'acceptations erronées sont des indicateurs de performance permettant l'adaptation de l'application au contexte d'utilisation de la modalité biométrique et que le degré de sécurité d'un système biométrique peut être adapté à l'utilisation attendue du système.

Nous avons également pris connaissance des avantages et des limites de la biométrie monomodale.

La fusion de plusieurs modalités dans la même application augmente la fiabilité du système. Le chapitre suivant est consacré à la présentation des différents concepts liée à la fusion de données (les systèmes multimodaux).

Chapitre 2

Les systèmes biométriques multimodaux

2.1 Introduction

Bien que certains systèmes monomodaux aient considérablement amélioré leur fiabilité et leur précision, ils rencontrent souvent des problèmes d'enregistrement dus à la non-universalité de leurs caractéristiques biométriques de base, à la vulnérabilité, à l'usurpation biométrique ou au manque de précision causé par des données bruitées.

Par conséquent, une modalité biométrique unique peut ne pas satisfaire l'exigence de performance souhaitée dans les applications du monde réel. Pour surmonter ces problèmes, les systèmes d'authentification biométriques multimodaux, combinant des informations provenant de plusieurs modalités, semblent constituer une solution fiable pour obtenir de bons résultats d'authentification. Des études ont montré que les systèmes biométriques multimodaux peuvent atteindre de meilleures performances que les systèmes monomodaux. Ces systèmes traitent le problème de la non-universalité, depuis plusieurs modalités pour assurer une couverture suffisante de la population.

Ils ont également découragé la falsification d'identité car il serait difficile pour un imposteur de falsifier plusieurs termes biométriques d'une personne.

2.2 Les avantages des systèmes biométriques multimodaux

Les avantages des systèmes multimodaux par rapport aux systèmes biométriques unimodaux sont énumérés ci-dessous (FLYNN JAIN ANIL, 2008).

1. Tolérance aux pannes : un système biométrique multimodal peut également être considéré comme un système à tolérance de panne qui continue à fonctionner même lorsque certaines sources biométriques deviennent peu fiables en raison d'un dysfonctionnement du capteur

ou du logiciel ou d'une manipulation délibérée de l'utilisateur. La notion de tolérance aux pannes est particulièrement utile dans les systèmes d'authentification à grande échelle impliquant un grand nombre de sujets (comme une application de contrôle des frontières).

2. Indexation de bases de données biométriques à grande échelle : les systèmes biométriques multimodaux peuvent faciliter le filtrage ou l'indexation de bases de données biométriques à grande échelle. Par exemple, dans un système bimodal comprenant un visage et une empreinte digitale, le jeu de caractéristiques de visage peut-être utilisé pour calculer une valeur d'index afin d'extraire une liste des candidats d'identité potentiels d'une grande base de données de sujets. La modalité d'empreinte digitale peut alors déterminer l'identité finale à partir de cette liste de candidats limitée.
3. Non-universalité : les systèmes biométriques multimodaux abordent le problème de la non-universalité rencontrée par les systèmes biométriques unimodaux. Par exemple, si le doigt coupé d'un sujet l'empêche de s'inscrire avec succès dans un système d'empreintes digitales, la disponibilité d'un autre trait biométrique, par exemple l'iris, peut être utilisée pour l'inclusion de l'individu dans le système biométrique.
4. Réduction de l'ambiguïté des déductions : des informations plus complètes ou plus précises permettent un meilleur choix entre les différentes hypothèses.
5. Attaques par usurpation d'identité : il devient de plus en plus difficile pour un imposteur d'usurper de multiples traits biométriques d'un individu légitimement inscrit.

2.3 Les différentes formes de multi modalité

Quelles sont les sources d'information pouvant être considérées dans un environnement multimodal? Nous abordons cette question en introduisant une terminologie décrivant les différents scénarios possibles pour obtenir de multiples sources de données.

Dans les quatre premiers scénarios décrits ci-dessous, la fusion des informations est réalisée à l'aide d'un trait unique, alors que dans le cinquième scénario, plusieurs traits sont utilisés (voir figure 3.1)(MORIZET, 2009).

2.3.1 Multi-capteurs :

Ces systèmes visent à capturer un seul trait biométrique à l'aide de deux capteurs ou plus, afin d'obtenir le plus d'informations possible, par exemple utilisation d'un lecteur d'empreintes digitales capacitives et d'un lecteur d'empreintes digitales résistives (ALLANO et al., 2009).

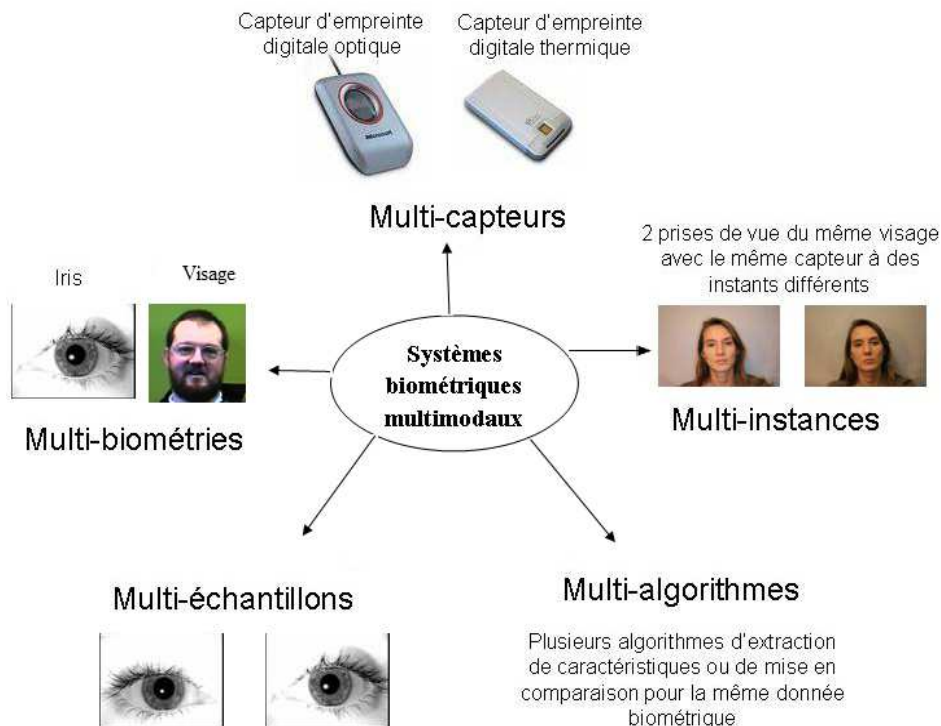


FIGURE 2.1 – Les différentes formes de multi modalité (ALLANO et al., 2009).

2.3.2 Multi-instances :

Une seule modalité biométrique mais plusieurs parties du corps humain sont utilisées et sont également appelées dans la littérature systèmes multi-unités . Un exemple est l'utilisation de plusieurs doigts dans la vérification des empreintes digitales.

2.3.3 Multi-algorithmes :

une seule entrée biométrique est traitée avec différents algorithmes d'extraction de caractéristiques afin de créer des modèles avec un contenu d'information différent. Un exemple est le traitement des images d'empreintes digitales selon des représentations minuties et texturées.

2.3.4 Multi-échantillons :

Un seul capteur peut être utilisé pour acquérir plusieurs échantillons de même trait biométrique afin de prendre en compte les variations éventuelles du trait ou d'obtenir une plus grande précision.

Un exemple est l'utilisation séquentielle de plusieurs impressions du même doigt dans la vérification des empreintes digitales. De même, un système de visage, par exemple, peut capturer le profil frontal du visage d'une personne avec les profils gauche et droit afin de prendre en compte les variations de la pose du visage.

2.3.5 Systèmes multi-biométries :

Les systèmes multi-biométries utilisent les informations de plusieurs modalités biométriques différentes combinées pour obtenir la reconnaissance des individus, par exemple la reconnaissance des visages associée à la reconnaissance de l'iris.

2.4 Les niveaux de fusion

Le système biométrique comporte quatre modules importants. Le module de capteur acquiert les données biométriques d'un utilisateur via des capteurs; le module d'extraction de caractéristiques traite les données biométriques acquises et extrait un ensemble de caractéristiques; le module de score compare l'ensemble de caractéristiques extraites aux modèles stockés à l'aide d'un classificateur ou d'un algorithme de correspondance afin de générer des scores de correspondance; dans le module de décision les scores correspondants sont utilisés soit pour identifier un utilisateur inscrit, soit pour vérifier l'identité de cet utilisateur.

Dans un système multibiométrique, la fusion peut être réalisée en utilisant les informations disponibles dans l'un de ces modules.

Ainsi, quatre niveaux de fusion différents sont possibles : le niveau du capteur, le niveau d'extraction des caractéristiques, le niveau du score et le niveau de décision (Figure 3.2). Allano et al. (ALLANO et al., 2009) ont classé la fusion d'informations dans les systèmes biométriques en deux grandes catégories : la fusion pré-classification et la fusion post-classification. Le niveau de capteur et le niveau caractéristiques sont appelés fusion de pré-classification, tandis que le niveau de score et le niveau de décision sont appelés fusion de post-classification.

2.4.1 La fusion pré-classification

La fusion de pré-classification est la fusion des informations de plusieurs données biométriques au niveau du capteur ou au niveau des caractéristiques extraites par le module d'extraction de caractéristiques.

1. Fusion au niveau capteur : les données brutes, obtenues à partir de la détection de la même caractéristique biométrique avec deux capteurs ou plus, sont combinées (ATREY et al., 2010). La fusion au niveau de capteur est applicable uniquement si les sources multiples représentent des échantillons du même trait biométrique obtenus à l'aide d'un capteur unique ou de capteurs compatibles différents.
2. Fusion au niveau de données ou caractéristiques : à ce niveau de fusion, les traits biométriques sont d'abord prétraités pour extraire séparément les vecteurs de caractéristiques, les combiner et générer un seul vecteur de caractéristiques composites. Ce vecteur sera utilisé par le processus de classification. Par exemple, dans de nombreux cas, les

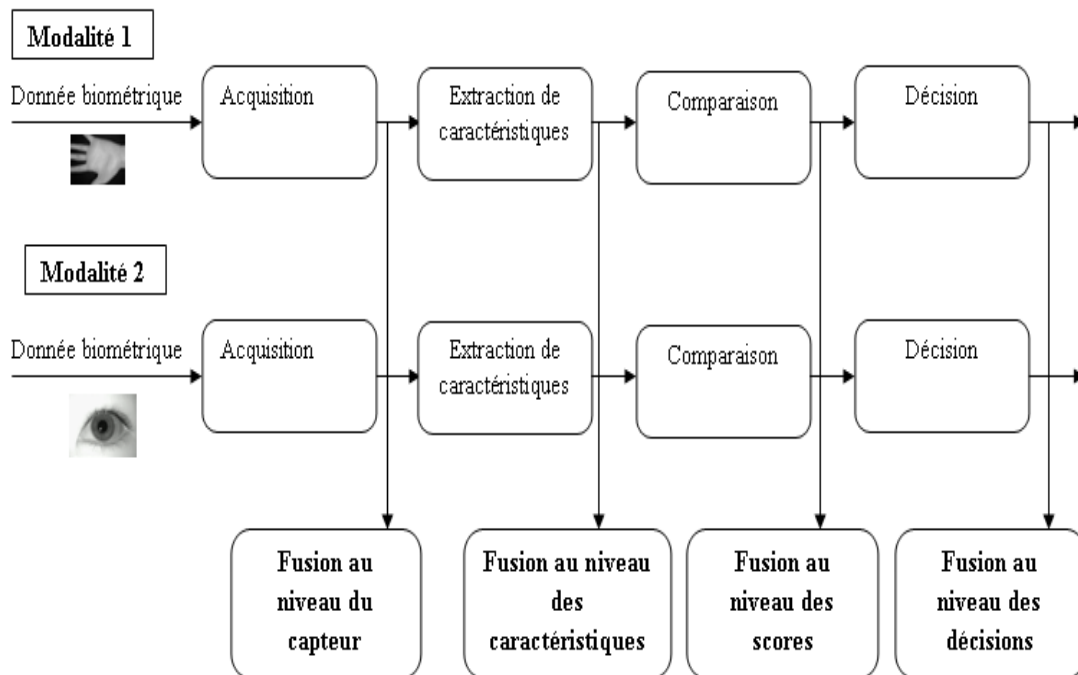


FIGURE 2.2 – Les différents niveaux de fusion.

caractéristiques peuvent ne pas être compatibles en raison de la différence de nature des modalités.

De plus, une telle liaison peut conduire à un vecteur de caractéristiques de très grandes dimensions. Cela augmente la charge de calcul. Il a été signalé qu'une conception de classificateur complexe peut-être nécessaire pour opérer sur l'espace de fonctions à concaténer (NANDAKUMAR, 2008).

2.4.2 La fusion post-classification

La fusion post-classification peut s'effectuer au niveau des scores des modules de comparaison ou au niveau des décisions. Dans les deux cas, la fusion est en fait un problème bien connu dans la littérature sous le nom "systèmes de classificateurs multiples".

1. Fusion au niveau score : lorsque chaque système biométrique génère un score indiquant la proximité des données d'entrée par rapport à un modèle, l'intégration peut être effectuée au niveau du score. C'est ce que l'on appelle la fusion au niveau de mesure ou au niveau de confiance.

Les résultats de score générés par les appareils de correspondance biométrique contiennent les informations les plus riches sur le modèle d'entrée. En outre, il est relativement facile d'accéder et de combiner les scores générés par les différents correspondants. Par conséquent, l'intégration de l'information au niveau du score de correspondance est l'approche la plus courante dans les systèmes biométriques multimodaux (MORIZET, 2009) (NANDAKUMAR, 2008).

2. Fusion au niveau des décisions : avec cette approche, chaque sous-système biométrique effectue de façon autonome les étapes d'extraction des caractéristiques, comparaison et reconnaissance (MORIZET, 2009). Ensuite, chaque système fournit une décision binaire sous la forme OUI ou NON qui peut être représentée par 0 et 1, et le système de fusion de décision consiste à prendre une décision finale basée sur cette série de 0 et 1. Les méthodes les plus utilisées sont des fonctions booléennes. La fusion au niveau de la décision est souvent utilisée en raison de sa simplicité .

Il est difficile de combiner des informations au niveau des caractéristiques car les ensembles de fonctionnalités utilisées par différentes modalités biométriques peuvent être inaccessibles ou incompatibles. La fusion au niveau de la décision est trop rigide car une quantité limitée d'informations est présentée à ce niveau. Par conséquent, l'intégration au niveau des scores est généralement préférée en raison de la facilité d'accès et de combinaison des scores générés par différents classificateurs, mais fusionner les informations à ce niveau est intéressante car il réduit la complexité en permettant à différents classificateurs d'être utilisés indépendamment l'un de l'autre.

2.5 Normalisation

Les méthodes de normalisation des scores visent à transformer individuellement chacun des scores des sous-systèmes pour les rendre homogènes avant de les combiner. En effet, les scores de chaque sous-système peuvent être de nature différente. Certains systèmes produisent des scores de similarité, d'autres produisent des distances (plus la distance est faible, plus la référence et le test sont proches, plus l'utilisateur est un client).

De plus, chaque sous-système peut avoir différents intervalles de variations des scores, par exemple, pour un système, les scores varient entre 0 et 1 et pour un autre, les scores varient entre 0 et 1000.

On comprend bien la nécessité de l'étape de normalisation avant que les scores bruts de différents classificateurs puissent être combinés dans l'étape de fusion. La normalisation aborde le problème des scores incomparables représentant les résultats de différents classificateurs biométriques (NANDAKUMAR, 2008).

Parmi les techniques de normalisation les plus utilisées, nous pouvons citer les suivantes :

1. La méthode Min-Max (MM) : la technique de normalisation Min-Max effectue une transformation linéaire sur les données afin de ne pas modifier le type de distribution initial. Il est utilisé lorsque les valeurs maximales et minimales des données produites par les classificateurs sont connues. Chaque point de données est normalisées utilisant la transformation suivante (MORIZET, 2009).

$$s_{no} = \frac{s - \min(s)}{\max(s) - \min(s)} \quad (2.1)$$

$max(s)$ et $min(s)$ définissent les points d'extrémité du domaine de définition des scores.

Cette méthode conserve la distribution statistique et ne fait que mettre à l'échelle et transformer les données en une plage numérique commune comprise entre $[0, 1]$.

La méthode de normalisation Min-Max n'est pas robuste (c'est-à-dire qu'elle est sensible aux valeurs aberrantes), car les scores minimum et maximum sont estimés à l'aide de l'ensemble de données d'apprentissage, ils sont donc fixes ; par conséquent, il existe un risque de débordements des données dans la phase opérationnelle du système si l'un des scores de l'ensemble de données d'essai dépasse le maximum, par conséquent, nous pouvons obtenir des valeurs négatives.

2. Z-score (ZS) : cette méthode transforme les scores à une distribution avec une moyenne égale 0 et un écart type égale 1 (ATREY et al., 2010).

$$s_{no} = \frac{s - moyenne(s)}{std(s)} \quad (2.2)$$

$moyenne(s)$ et $std(s)$ désignent respectivement la moyenne et l'écart type des scores.

Cependant, l'écart moyen et l'écart type sont sensibles aux valeurs aberrantes et, par conséquent, cette méthode n'est pas robuste. La normalisation du score Z ne garantit pas une plage numérique commune pour les scores normalisés des différents matcheurs.

3. La méthode tangente hyperbolique Tanh (TH) : cette méthode est parmi les techniques statistiques les plus solides. Il place chaque score normalisé dans l'intervalle $[0, 1]$ selon :

$$s_{no} = 0.5 \left[\tanh\left(0.001 \frac{s - u}{\sigma}\right) + 1 \right] \quad (2.3)$$

Où u est la moyenne arithmétique et σ l'écart-type des données, sont déterminés pour chaque sous-système sur une base de développement.

La méthode de normalisation de Tanh n'est pas sensible aux valeurs aberrantes.

4. Normalisation décimale scaling : cette méthode peut être appliquée lorsque les scores des correspondants individuels évoluent dans des intervalles logarithmiques (MORIZET, 2009). Par exemple, les scores de l'un sont compris dans l'intervalle $[0, 1]$ et les scores des autres sont compris dans l'intervalle $[0, 1000]$. La normalisation suivante peut être appliquée :

$$s_{no} = \frac{s}{10^n} \quad (2.4)$$

où $n = \log_{10}max(s)$.

Les faiblesses de cette approche sont le manque de robustesse et l'hypothèse de variation des différents matchers selon un facteur logarithmique.

5. La médiane et l'écart absolu médian (MAD) : la normalisation par la médiane et le MAD est faite selon l'équation 3.5 et s'avère robuste car insensible aux valeurs aberrantes.

Cependant, contrairement à la moyenne et l'écart type, la médiane et le MAD sont des estimateurs faibles de la position et de l'échelle lorsque la distribution du score est non gaussienne. C'est pourquoi cette technique est peu efficace, ne préserve pas la distribution d'entrée en sortie et ne transforme pas les scores normalisés de différents matchers en un intervalle commun.

$$s_{no} = \frac{s - median}{MAD} \quad (2.5)$$

où $MAD = median(|s - median(s)|)$

6. Fonction double sigmoïde : la normalisation double sigmoïde est donnée par l'équation (3.6) où t désigne le point de référence, r_1 et r_2 les bornes entre lesquelles la fonction double sigmoïde est linéaire (KHIARI, 2016).

$$s_{no} = \begin{cases} \frac{1}{1 + \exp(-\frac{1}{2}(s-t/r_1))} & \text{si } s \leq t \\ \frac{1}{1 + \exp(-\frac{1}{2}(s-t/r_2))} & \text{sinon} \end{cases} \quad (2.6)$$

Cette technique garantit la transformation des scores dans un intervalle commun compris entre 0 et 1 mais nécessite un réglage prudent des paramètres t, r_1 et r_2 pour avoir une bonne efficacité. Les scores compris dans la plage $[t - r_1, t + r_2]$ correspondent à la zone de chevauchement des distributions authentique et imposteur et subissent une transformation linéaire, tandis que les scores situés en dehors de cet intervalle sont transformés de manière non linéaire.

7. Normalisation de cohorte sans contrainte : les variations des données biométriques sont considérées comme l'un des principaux problèmes de la fusion multimodale. Ces variations sont reflétées dans les scores biométriques correspondants, et peuvent avoir une influence négative sur l'efficacité et la précision de la reconnaissance biométrique. Ces variations peuvent survenir en raison de conditions de capture non idéales telles que le bruit de fond. Un autre aspect de la difficulté de la biométrie multimodale est le manque d'informations sur la variation dans les différents types de données biométriques.

Pour s'attaquer à ce problème de variation des données, de nombreuses études ont été menées récemment pour améliorer la précision et la robustesse de la biométrie multimodale, grâce à l'introduction de Normalisation de cohorte sans contrainte sur le terrain (ALSAADE et al., 2008). Alors que, ce schéma de normalisation a été largement utilisé en biométrie vocale, la Normalisation de cohorte sans contrainte peut être très utile pour séparer les scores authentiques des scores des imposteurs.

Comme décrit dans (ALSAADE et al., 2008), étant donné un jeton de test de certains types biométriques, le score normalisé fourni par la normalisation de cohorte sans contrainte peut être exprimé par :

$$X_{(v/p)} = \log pT^{(v/p)} - \frac{1}{c} \sum_{c=1}^C \log p_c^{(v/p)} \quad (2.7)$$

Où v et p représentent le type de biométrie (visage et parole), $X_{(v/p)}$ est le score normalisé pour la modalité visage ou parole, $pT^{(v/p)}$ est le score du modèle cible, $p_c^{(v/p)}$ sont les scores obtenus pour un ensemble de modèles concurrents, et C est le nombre de modèles concurrents considérés. Ces modèles concurrents sont sélectionnés de manière dynamique à partir d'un groupe de modèles d'arrière-plan, en fonction de leur proximité avec le jeton de test.

2.6 Les méthodes de fusion de score

La fusion au niveau du score, également appelée mesure de fusion ou niveau de confiance, fait référence à l'utilisation de techniques de fusion pour combiner les scores fournis par les différents classificateurs.

Il existe deux approches pour combiner les scores obtenus par différents systèmes. La première approche consiste à aborder le sujet en tant que problème de combinaison, tandis que l'autre l'aborde en tant que problème de classification. Dans ce qui suit, nous allons donner un aperçu des différentes méthodes de fusion des scores de chaque catégorie.

2.6.1 Méthodes de fusion par combinaison de scores

Les méthodes de cette catégorie fusionnent les résultats en combinant des méthodes mathématiques. Les scores individuels sont combinés pour former un score unique qui est ensuite utilisé pour prendre la décision finale (ATREY et al., 2010). Afin de garantir la cohérence des scores de différents systèmes, ceux-ci doivent passer par une phase de normalisation.

1. Les méthodes de combinaison simples :

Les méthodes de combinaison de scores simples sont des méthodes théoriques très simples qui visent à extraire un score S à partir des N scores disponibles issus de N systèmes. Les méthodes les plus utilisées sont :

- La somme : dans cette technique, le score fusionné est calculé en ajoutant les scores de toutes les modalités impliquées. Le calcul ici est défini comme :

$$S = \sum_{i=1}^n S_i \quad (2.8)$$

- La moyenne : combiner les scores par la moyenne consiste à calculer S tel que :

$$S = \frac{1}{N} \sum_{i=1}^n S_i \quad (2.9)$$

- Le produit : cette technique de fusion calcule le score fusionné en multipliant les scores pour toutes les modalités impliquées. Il est défini mathématiquement comme :

$$S = \prod_{i=1}^N S_i \quad (2.10)$$

- Le minimum : la méthode de la règle minimale sélectionne le score ayant la valeur la plus faible parmi les modalités impliquées. Il est défini mathématiquement comme :

$$S = \text{Min}(s_i) \quad (2.11)$$

- Le maximum : la méthode de la règle maximale sélectionne le score ayant la plus grande valeur parmi les modalités impliquées. Il est défini mathématiquement comme :

$$S = \text{max}(S_i) \quad (2.12)$$

- La médiane : Combiner les scores par la médiane consiste à calculer S tel que :

$$S = \text{Med}(S_i) \quad (2.13)$$

- La somme pondérée : est la technique la plus évoluée de cette catégorie qui nécessite une adaptation par le réglage de paramètres appelée poids.

$$S = \sum_{i=1}^N W_i S_i \quad (2.14)$$

La somme pondérée permet de donner différents poids W_i à chacun des sous-systèmes en fonction de leurs performances individuelles ou de leur intérêt pour le système multimodal.

2. Les méthodes de combinaison par logique floue :

La théorie de la logique floue (sous-ensembles flous) a été proposée par Zadeh (ZADEH, 1965) afin de prendre en charge la logique binaire d'une part et d'améliorer la logique multivolume d'autre part. Contrairement à la logique binaire, très proche du langage machine, la logique floue converge vers un raisonnement humain, basé sur l'intégration et le traitement de caractère approximatif.

- Mesure floue : un jeu de fonction $g : P(Y) \rightarrow [0, 1]$ est appelée une mesure floue si les conditions suivantes sont remplies :

Conditions aux limites : $g(\phi) = 0, g(Y) = 1$

Monotonie : $g(A) \leq g(B)$, si $A \subset B$ et $A, B \in P(Y)$

Continuité : $\lim_{x \rightarrow +\infty} g(A_i) = g(\lim_{x \rightarrow +\infty} A_i)$, si A_i est une suite croissante d'ensembles mesurables.

Sugeno (SUGENO, 1993) a défini à partir de cette définition une valeur de g mesure accompagnant une propriété supplémentaire, comme suit :

$$g(AB) = g(A) + g(B) + \lambda g(A)g(B) \quad (2.15)$$

pour tous les $A, B \subset Y$ et $A \cap B = \phi$, et pour certains $\lambda \geq -1$, le $g\lambda$ mesure floue devient une mesure de probabilité standard. En général, la valeur de λ peut être déterminée en raison de l'état limite de la mesure floue $g\lambda$. Cette condition pour $g(Y) = 1$. Par conséquent, la valeur de λ est déterminée par la résolution de ce qui suit :

$$g\lambda(Y) = \frac{1}{\lambda} \left(\prod_{i=1}^n (1 + \lambda \hat{g}) - 1 \right), \lambda \neq 0 \quad (2.16)$$

Est équivalent à :

$$\lambda + 1 = \left(\prod_{i=1}^n (1 + \lambda \hat{g}) - 1 \right) \quad (2.17)$$

Où $\lambda \in (-1, +\infty)$, $\lambda \neq 0$, et g est la valeur de la fonction de densité floue. La solution peut être facilement obtenue.

- Intégrale floue : l'intégrale floue de la fonction calculée sur Y par rapport à une mesure floue g est définie sous la forme :

$$\int_Y h(y) \circ g(\cdot) = \sup_{\alpha \in [0,1]} [\min[\alpha, g(y|h(y) \leq \alpha)]] \quad (2.18)$$

L'intégrale de Sugeno : lorsque les valeurs des $h(\cdot)$ sont classées dans l'ordre décroissant, $h(y_1) \leq h(y_2) \leq \dots \leq h(y_n)$

L'intégrale floue de Sugeno est calculée comme suit (SHUKLA, TIWARI et KALA, 2010) :

$$\int_Y h(y) \circ g(\cdot) = \max_{i=1:n} [\min(h(y_i), g(A_i))] \quad (2.19)$$

Où $A_i = y_1, y_2, \dots, y_i$ indique un sous-ensemble d'éléments. Les valeurs de $g(A_i)$ supportées par la mesure floue sur les sous-ensembles correspondants peuvent être déterminées de manière récursive sous la forme :

$$g(A_1) = g(y_1) = g^1 \quad (2.20)$$

$$g(A_i) = \hat{g} + g(A_i - 1) + \lambda \hat{g} g(A_i - 1) \quad (2.21)$$

Le calcul de la fonction de densité floue g_i sur la base des données est effectué comme suit :

$$\begin{cases} \hat{g} = \beta p_i, i = 1 \\ g(A_i) = \hat{g} + g(A_i - 1) + \lambda \hat{g} g(A_{i-1}), i = 2, 3, 4 \end{cases} \quad (2.22)$$

Où p_i est le taux de classification dans l'intervalle $[0, 1]$ pour chaque système. $\beta \in [0, 1]$ est un facteur qui met en place un certain équilibre entre les résultats de la classification.

L'intégrale de Choquet : il a été démontré que l'équation de l'intégrale floue (3.18) n'est pas un prolongement correct de l'intégration habituelle de Lebesgue. En d'autres termes, lorsque la mesure est ajoutée, l'expression ci-dessus ne renvoie pas l'intégrale au sens de Lebesgue. Afin de remédier à cet inconvénient, Murofushi et Sugeno (MUROFUSHI et SUGENO, 1989) ont proposé une soi-disant intégrale floue de Choquet calculée comme suit :

$$\int_Y h(y) dg(\cdot) = \sum_{i=1}^n [h(y_i) - h(y_{i+1})] g(A_i) h(y_{n+1}) = 0 \quad (2.23)$$

Shukla et al. (SHUKLA, TIWARI et KALA, 2010) proposent des explications supplémentaires sur la combinaison des scores de logique floue basée sur l'intégrale de Sugeno ou de Choquet.

3. Combinaison en utilisant les techniques d'optimisation :

- Les algorithmes génétiques ou GA (Genetic Algorithm) : les GA sont des algorithmes inspirés de la science génétique développée au XIX^e siècle par Darwin. L'objectif principal de ces algorithmes est de trouver une solution aux problèmes difficiles (problèmes pour lesquels nous ne connaissons pas les méthodes exactes pour les résoudre dans un délai raisonnable). Ces approches permettent l'évolution d'une génération (solution) à une autre par un ensemble d'opérations (mutation et croisement) tout en cherchant à améliorer une fonction objective. Le codage des éléments de la population est l'élément le plus important, puisqu'il permet de modéliser un problème sous la forme de données informatiques manipulables (MITCHELL, 1998).

Des GA sont appliqués dans divers domaines pour résoudre des problèmes d'optimisation ou de recherche. Par exemple : la conception topologique de réseaux téléinformatiques à commutation de paquets, la recherche d'informations, les applications en économie...

Pour utiliser ces approches, les éléments suivants doivent être disponibles (MITCHELL, 1998) :

- (a) Le codage d'un élément de population : une fonction permettant de modéliser les données du problème réel en données utilisables par la GA.

- (b) Une façon de créer la population initiale : la génération de la population initiale est importante car cette génération représente le point de départ de l'algorithme et son choix influe sur la rapidité et l'optimalité de la solution finale.
- (c) Une fonction d'évaluation ou fonction objective : une fonction qui renvoie une valeur d'adaptation pour chaque individu. Cette valeur permet de déterminer la solution pertinente, puisque le problème se limite à rechercher le groupe d'individus ayant les valeurs optimales.
- (d) Des opérateurs qui permettent d'évoluer d'une population à l'autre en améliorant la fonction objective. L'opérateur de croisement recompose les gènes d'individus existant dans la population, tandis que l'opérateur de mutation vise à garantir l'exploration de l'espace d'état.
- (e) Paramètres de dimensionnement : taille de la population, nombre total de générations (critère d'arrêt), probabilités d'application d'opérateurs de croisement et de mutation, etc.

Un GA fonctionne comme suit (voir figure 3.3) :

Étape 1 : Initialisation, nous sélectionnons les individus qui représentent la population initiale.

Étape 2 : Évaluation, on évalue chaque individu par la fonction objective.

Étape 3 : Sélection, nous définissons les individus de la génération P qui seront dupliqués dans la nouvelle population. À chaque génération, il y a deux opérateurs de sélection : la sélection de la reproduction, ou plus simplement la sélection, qui détermine les individus qui vont se reproduire au cours d'une génération et la sélection du remplaçant, ou plus simplement le remplacement, qui détermine les individus qui devront disparaître de la population.

Étape 4 : Reproduction, des opérateurs génétiques de reproduction (croisement et mutation) sont utilisés pour produire la nouvelle génération. Les opérateurs de mutation modifient un individu pour en former un autre, tandis que les opérateurs de croisement génèrent un ou plusieurs enfants à partir de combinaisons de deux parents.

Les GA présente de nombreux avantages :

Ils peuvent rapidement analyser un vaste ensemble de solutions.

Pour un problème d'optimisation, les algorithmes génétiques sont capables de trouver la solution optimale globale dans un espace multidimensionnel sans se soucier des minimaux locaux (MITCHELL, 1998).

L'inconvénient majeur de l'GA est que l'algorithme utilise une très grande quantité de temps de traitement.

- Particle swarm optimization (PSO) : le PSO est un algorithme évolutif basé sur la population, proposé par le docteur Kennedy et Eberhart en 1995, qui repose sur l'intelligence de l'essaim (MARINI et WALCZAK, 2015).

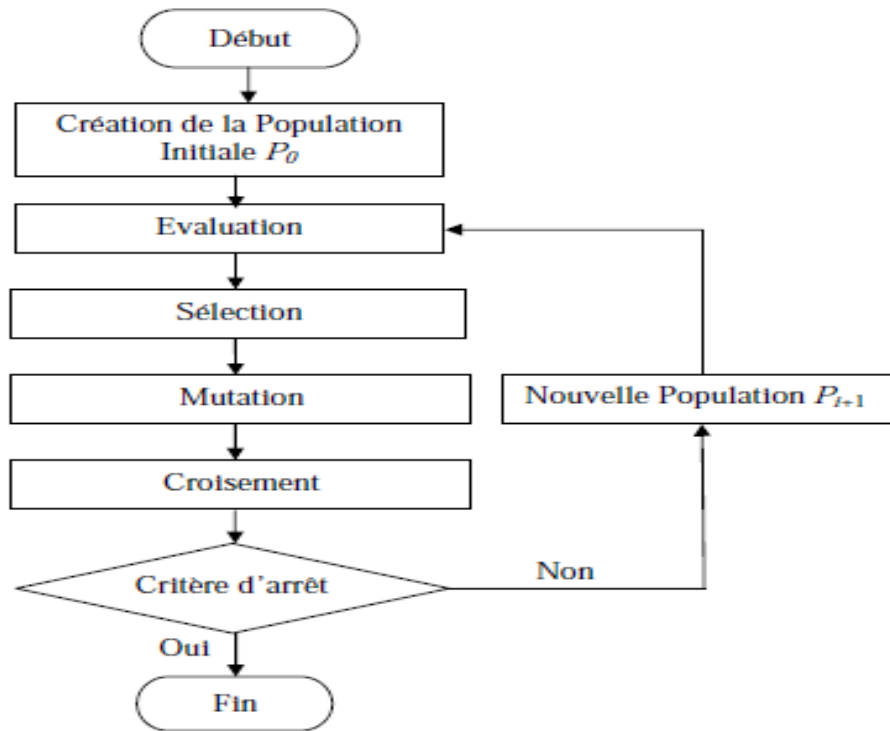


FIGURE 2.3 – Structure générale d'un GA (MITCHELL, 1998).

Il a été inspiré par le comportement social des groupes d'oiseaux lorsqu'ils cherchent de la nourriture. Dans PSO, les solutions potentielles, appelées particules, survolent l'espace à problèmes en recherchant de meilleures régions.

En raison de sa simplicité et de sa simplicité de mise en œuvre, l'algorithme PSO peut être largement utilisé dans des domaines tels que l'optimisation des fonctions, la classification des modèles, l'apprentissage automatique, la formation réseau neutre, le traitement du signal, le contrôle du système vague, le contrôle automatique d'adaptation, etc.

PSO est une technique d'optimisation stochastique basée sur la population qui vise à résoudre un problème d'optimisation dans un espace de recherche. Chaque solution candidate est donc modélisée par particule dans un espace de recherche. Chaque particule ajuste sa trajectoire en utilisant sa mémoire individuelle et les connaissances acquises par ses voisins pour trouver la meilleure solution. Chaque particule teste une solution possible au problème multidimensionnel car se déplace dans l'espace du problème.

Le mouvement des particules est influencé par deux facteurs : la meilleure solution de la particule ($pbest$) et la meilleure solution globale trouvée par toutes les particules ($gbest$).

La particule interagit avec tous les voisins et stocke dans sa mémoire les informations de localisation optimales.

Après chaque itération, $pbest$ et $gbest$ sont mis à jour si une solution

plus optimale est trouvée par la particule ou la population, respectivement. Ce processus se poursuit de manière itérative jusqu'à ce que le résultat optimal ou le critère d'arrêt soit atteint.

Toutes les particules sont évaluées par une fonction de remise en forme à optimiser. Une propriété de vitesse est associée à chaque particule. (MAZOUNI et RAHMOUN, 2011).

— $A^{(s)}$, la position actuelle de la particule s (A -vector).

— $P^{(s)}$, la meilleure position locale de la particule s (P -vector).

— $v_{(s)}$, la vitesse actuelle de la particule s (v -vector).

La meilleure position locale des particules s à une itération donnée est mise à jour comme suit :

$$p^{(s)}(t+1) = \begin{cases} p^{(s)}(t) & \text{si } f(A^{(s)}(t+1)) \geq f(p^{(s)}(t)) \\ A^{(s)}(t+1) & \text{si } f(A^{(s)}(t+1)) < f(p^{(s)}(t)) \end{cases}$$

La vitesse de la particule s est mise à jour en utilisant l'équation (3.24) et la position de la particule s est mise à jour en utilisant l'équation (3.25) pour plus de détails, voir la référence (KALYAN, ANN et K., 2003).

$$v_{(s)}(t+1) = \omega v_{(s)}(t) + c_1 r_1 (pbest(t) - A^{(s)}(t)) + c_2 r_2 (gbest - A^{(s)}(t)) \quad (2.24)$$

$$A^{(s)}(t+1) = A^{(s)}(t) + v_{(s)}(t+1) \quad (2.25)$$

Où $pbest$ est la position de la meilleure valeur de forme de la particule dans l'itération courante, $gbest$ est la position de la particule avec la meilleure valeur de forme physique dans l'essaim, c_1 et c_2 sont les constantes d'accélération et r_1 et r_2 sont deux nombres aléatoires uniformément distribués dans $[0, 1]$.

ω est le poids d'inertie compris entre $[0, 1]$. C'est un paramètre important car il contrôle l'effet du vecteur de vitesse de l'essaim.

L'algorithme PSO s'exécute jusqu'à un nombre d'itérations fixe ou une variation de vitesse proche de zéro.

L'algorithme PSO pose les avantages suivants :

- Le principal avantage de PSO est la convergence rapide.
- Les PSO n'ont pas de calcul de recouvrement et de mutation.
- La recherche peut être effectuée à la vitesse de la particule. Au cours du développement de plusieurs générations, seule la particule la plus optimiste peut transmettre des informations sur les autres particules, et la vitesse de recherche est très rapide.
- PSO est facile à mettre en œuvre et il y a peu de paramètres à ajuster.

Les inconvénient de PSO sont :

- La méthode souffre facilement de l'optimum partiel, ce qui en

fait le moins exact au moment de régler sa vitesse et sa direction.

- Le principal inconvénient de la topologie *gbest* est qu'elle est incapable d'explorer plusieurs régions optimales simultanément.
- Il est parfois facile de piéger PSO dans des optimums locaux et le taux de convergence est considérablement diminué à la fin de la période d'évolution ; lorsqu'une solution presque optimale est choisie, l'algorithme cesse d'optimiser, ce qui limite la précision qu'il peut obtenir.
- Les algorithmes PSO fonctionnent selon l'idée de suivre un chef de file pour balayer l'espace de recherche. Ce mouvement de suivi de certaines particules d'une population peut devenir un désavantage où il y a beaucoup de fronts optimaux locaux.

2.6.2 Les méthodes de fusion par classification de scores :

Les approches de fusion par classification de scores considèrent l'opération de fusion comme étant un problème de classification. Elles cherchent à prendre une décision à partir du vecteur des scores de dimension N . Le but des méthodes de fusion basées sur la classification est d'effectuer une discrimination entre les deux classes (Client et Imposteur) dans l'espace à N dimensions des scores. La discrimination entre ces deux classes peut être plus ou moins complexe selon la nature de classificateur utilisé. Toutefois, les méthodes classiques de classification utilisées en reconnaissance des formes et en apprentissage automatique peuvent être utilisées pour la fusion de scores de ce type (ALLANO et al., 2009). À noter que ces méthodes sont moins performantes que leurs antécédents (les méthodes fusion par combinaisons) à cause de leurs complexités. Dans le domaine des systèmes biométriques multimodaux, plusieurs types de méthodes de classification ont été employés. Nous allons citer quelques-unes les plus utilisées :

1. La méthode KNN (k Nearest Neighbours, ou k plus proches voisins) :

C'est une approche très simple et directe. Elle ne nécessite pas d'apprentissage, mais simplement le stockage des données d'apprentissage. Son principe est le suivant :

Dans le contexte de la classification d'une nouvelle observation x , l'idée fondatrice simple est de faire voter les plus proches voisins de cette observation (MATHIEU-DUPAS, 2010). La classe de x est déterminée par la classe majoritaire parmi les k voisines les plus proches de l'observation de x au sens d'une distance choisie.

2. SVM (Support Vector Machine) :

SVM, également connues sous le nom de séparateurs à Vastes Marges, introduites par Vapnik (VAPNIK, 1995), sont destinées à définir un hyperplan de séparation entre les deux classes, hyperplan qui minimise les erreurs de classification sur un ensemble d'apprentissage.

Cette méthode d'apprentissage supervisé peut apprendre un séparateur plus ou moins complexe en fonction de la nature du noyau sélectionné. Le noyau le plus simple est le noyau linéaire qui correspond à la recherche d'un séparateur linéaire dans l'espace à N dimensions des scores. Le but des fonctions noyau est de transformer l'espace initial en un espace de dimension plus grand dans lequel les données pourraient être séparables linéairement. Le séparateur est donc toujours linéaire dans l'espace transformé par la fonction noyau, mais n'est plus dans l'espace des scores (MORIZET, 2009).

Le but du SVM est donc de trouver un séparateur qui minimise l'erreur de classification sur le jeu d'apprentissage, mais qui donnera également de bons résultats en généralisation des données non utilisées dans l'apprentissage. Pour cela le concept utilisé est celui de marge (d'où les noms de séparateurs à vaste marge). La marge est la distance moyenne au carré entre le séparateur et les éléments d'apprentissage les plus proches, appelés vecteurs de support. Ces éléments sont appelés vecteurs de support car ce n'est que sur ces éléments de l'ensemble d'apprentissage que le séparateur est optimisé.

Pour un ensemble d'échantillons $(x_1, y_1), \dots, (x_n, y_n)$ à deux classes $y_i = \pm 1$, le problème consiste à trouver un hyperplan tel que :

- Les données des étiquettes de classe +1 et -1 se trouvent de chaque côté de l'hyperplan.
- La distance des vecteurs les plus proches de l'hyperplan est maximale.

Ces vecteurs sont appelés des vecteurs de support et la distance qui les sépare de l'hyperplan est la marge optimale.

De manière plus formelle, le but de cette méthode est de trouver un hyperplan $wx + b$, $w \in R$ et $b \in R$ qui sépare les deux classes avec la plus grande marge. La recherche de la marge optimale pour déterminer les paramètres w et b de l'hyperplan entraîne un problème d'optimisation quadratique consistant à minimiser :

$$[\|w\|^2 + C \sum_i \xi_i |y_i (w \cdot \phi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0] \quad (2.26)$$

Où C représente un paramètre de compromis entre la marge et les erreurs, ξ_i représente une variable ressort associé à l'observation x_i , et ϕ est une transformation.

Le problème peut être résolu par la méthode lagrangienne d'optimisation quadratique avec des contraintes (formulation double) pour maximiser la marge (KHIARI, 2016).

$$\sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i, x_j) | 0 \leq \alpha_i \leq C, \sum_i \alpha_i y_i = 0 \quad (2.27)$$

Où α_i est le multiplicateur Lagrangien associé au vecteur x_i . Si la valeur de α_i est non nul alors x_i est un vecteur de support et

$K(x_i, x_j) = \phi(x_i)\phi(x_j)$ est le noyau de transformation.

Le noyau d'un SVM est une fonction symétrique qui permet de projeter les données dans un espace transformé de grandes dimensions dans lequel la séparation de classes est plus facile à effectuer (Figure 3.4) (EJARQUE et al., 2009).

Parmi les noyaux les plus utilisés, nous avons :

- Le noyau linéaire : $K(x_i, x_j) = x_i \cdot x_j$ (Figure 3.4.a)
- Le noyau gaussien : $K(x_i, x_j) = \exp\left(-\frac{|x_i - x_j|^2}{2\sigma^2}\right)$ (Figure 3.4.b)
- Le noyau polynomial : $K(x_i, x_j) = (u x_i x_j + v)^p$ (Figure 3.4.c)

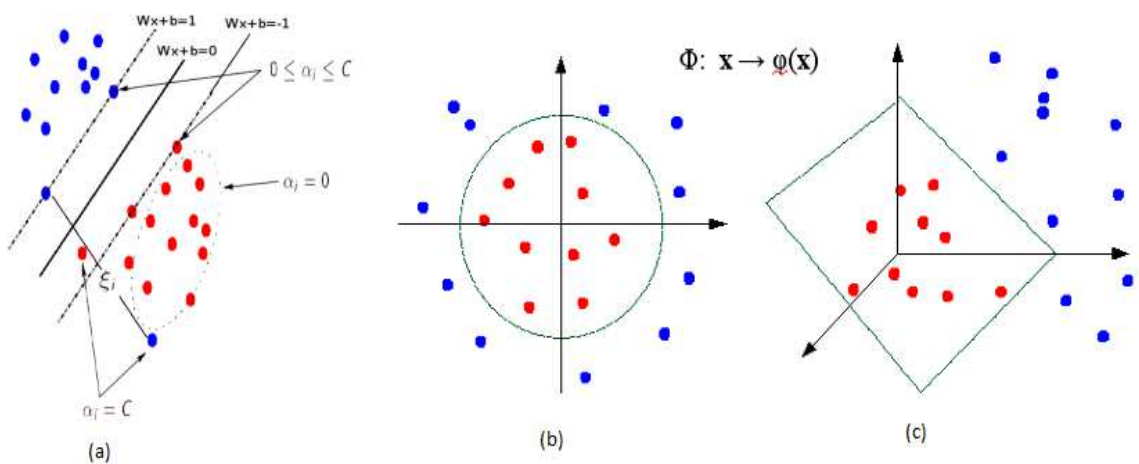


FIGURE 2.4 – Principe de l'SVM : (a) SVM avec un noyau linéaire (b) SVM avec un noyau gaussien (c) SVM avec un noyau polynomial (EJARQUE et al., 2009).

La décision est obtenue selon le signe de la fonction :

$$f(x) = \text{sign} \left[\sum_{i,j} \alpha_i y_i K(x_i, x) + b \right] \quad (2.28)$$

3. Les arbres de décision :

Les arbres de décision (AD) constituent une catégorie d'arbres utilisée dans l'exploration de données en informatique décisionnelle. Ils utilisent une représentation hiérarchique de la structure de données sous forme de séquences de décision (tests) dans le but de prédire un résultat ou une classe. Chaque individu (ou observation), qui doit être affecté à une classe, est décrit par un ensemble de variables testées dans les nœuds de l'arborescence. Les tests sont effectués dans les nœuds internes et les décisions sont prises dans les nœuds feuilles. Leur arborescence les rend également lisibles par un être humain, contrairement à d'autres approches où le prédicteur construit est une boîte noire (HAWARAH, 2008).

Un arbre de décision est une structure souvent utilisée pour représenter des connaissances. Cela permet de remplacer un expert humain

dans la détermination des propriétés d'un objet, c'est l'opération de classification.

En général, à chaque nœud feuille, il existe une valeur de classe. Les arbres de décision sont construits à partir d'un ensemble d'apprentissage. Un chemin de la racine à un nœud est une série d'attributs (questions) avec leurs valeurs (réponses).

Dans toutes les méthodes de construction d'un arbre de décision, nous trouvons les trois opérateurs :

- Décidez si un nœud est terminal : tous les exemples appartiennent à la même classe.
- Sélectionnez un test à associer à un nœud.
- Assigner une classe à une feuille. La classe majoritaire est affectée à une feuille.

4. Les réseaux de neurones :

Le réseau de neurones artificiels (Artificial Neural Network - ANN) a été introduit en tant que modèle rudimentaire de traitement de l'information dans le cerveau humain.

Ainsi, la structure élémentaire d'un ANN est un réseau de petits nœuds informatique connecté les uns aux autres par des liaisons dirigées et pondérées. Les nœuds représentent les neurones et les liens pondérés représentent la force des connexions synaptiques reliant les neurones les uns aux autres.

Un réseau neuronal artificiel est généralement composé de plusieurs couches (KHIARI, 2016) :

- **Une couche d'entrée** : retransmet les entrées sans distorsion.
- **Des couches intermédiaires appelées couches cachées** : transformation non linéaire $x_i \rightarrow \phi(x_i)$ des entrées. Ces couches permettent de discriminer des classes d'objets séparables de manière non linéaire.
- **Une couche de sortie** : transformation linéaire $\phi(x_i) \rightarrow y_i$ des fonctions cachées. Pour un problème de classification, nous aurons autant de sorties que de classes et chaque sortie est interprétée comme une probabilité postérieure.

Il existe deux types de réseaux de neurones : les monocouches et les multicouches. Les réseaux monocouches ont été rapidement ignorés car ils ne peuvent pas séparer ni les classes non linéairement séparables ni les multi-classes. Les réseaux multicouches sont une solution efficace à ces problèmes grâce à leurs propriétés d'approximation universelles (PRAVEEN KUMAR NAYAK, 2013), qui leur permettent d'approcher n'importe quelle fonction avec une précision arbitraire.

Les réseaux de neurones peuvent jouer le rôle d'un pur classificateur dans lequel leurs utilitaires résident dans leur capacité d'apprentissage automatique, leur permettant de résoudre des problèmes sans recourir à des règles complexes, tout en étant tolérants aux erreurs. Mais nous pouvons aussi les considérer pour réaliser la fusion de données afin de séparer deux populations données, à savoir les clients et les imposteurs dans notre cas (KHIARI, 2016).

5. LDA (Linear Discriminant Analysis) :

LDA est utilisée pour trouver la combinaison linéaire de caractéristiques qui sépare mieux les classes d'objets ou d'événement. Les combinaisons résultantes peuvent être utilisées comme classificateur linéaire ou, généralement, dans la réduction des caractéristiques (KHIARI, 2016).

L'analyse discriminante linéaire recherche un ensemble optimal de vecteurs de projection discriminant W , pour mapper l'espace de données d'origine sur un espace de fonctions de dimension inférieure, en maximisant le critère de Fisher $J(W)$, ce qui signifie que le chevauchement entre les classes dans l'espace des caractéristiques de la dimension inférieure est minimal (SHASHOA et al., 2016).

La figure 3.5 est un exemple de classification de deux nuages de points. L'axe principal de LDA est l'axe de projection qui maximise la séparation entre les deux classes. Il est clair que cette projection est optimale pour la séparation des deux classes. Plus précisément, pour tous les échantillons de toutes les classes, deux mesures sont définies.

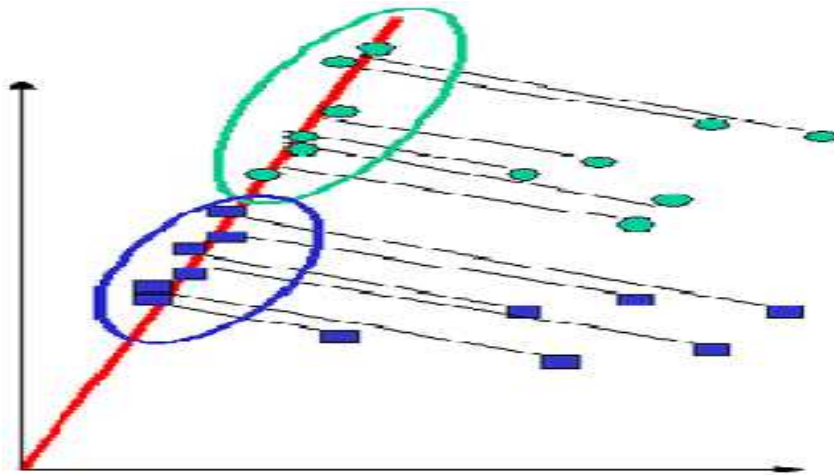


FIGURE 2.5 – Les projections de deux classes de points sur les axes principaux construits par la méthode LDA (KHIARI, 2016)

La première mesure est la matrice de dispersion intra-classe S_w (within class scatter matrix) et la seconde mesure est la matrice de dispersion inter classe S_b (between class scatter matrix) qui sont définies par (KHIARI, 2016) :

$$S_w = \sum_{i=1}^L \frac{1}{n_i} \sum_{j=1}^{n_i} (X_j^i - m^i)(x_j^i - m^i)^T \quad (2.29)$$

Avec x_j^i est le j échantillon de la classe i , m^i la moyenne de la classe et n_i le nombre d'échantillons de la classe i .

$$S_b = \sum_{i=1}^L (m^i - \bar{m})(m^i - \bar{m})^T \quad (2.30)$$

Avec \bar{m} est la moyenne de tous les échantillons.

LDA une fonction de W qui peut être donnée comme (SHASHOA et al., 2016) :

$$J(W) = \frac{|W^T S_b W|}{|W^T S_w W|} \quad (2.31)$$

W pris de telle sorte que $J(W)$ soit maximum. La solution $\{w_i | i = 1, 2, \dots, N\}$ est un ensemble de vecteurs propres généralisés associés aux valeurs propres généralisées $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$ du problème de valeur propre généralisée suivante :

$$S_b w_i = \lambda_i S_w w_i \quad (2.32)$$

Dans cette relation, w_i représente les vecteurs propres à la colonne de W correspondant aux plus grandes valeurs propres λ_i . Pour les vecteurs propres w_i , ces vecteurs propres sont les colonnes de la matrice de transformation W et la réduction de la dimension des points de données est réduite à l'aide de la transformation suivante :

$$h(X) = W^T X \quad (2.33)$$

2.7 Travaux récents sur biométrie multimodale

De manière générale, le choix et le nombre de modalités biométriques sont largement déterminés par la nature de l'application, les coûts consommés par multiples modalités et la corrélation entre les modalités considérées. Les informations non corrélées sont préférées pour l'amélioration des performances. Par exemple, dans un téléphone mobile équipé d'un appareil photo, il pourrait être plus facile de combiner les modes de visage et de voix d'un utilisateur. Dans une application ATM (Asynchronous Transfer Mode), il serait plus facile de combiner les termes de l'empreinte digitale et le visage de l'utilisateur. Ainsi, une diversité de systèmes biométriques multimodaux a été proposée dans la littérature en utilisant différentes modalités biométriques basées sur différentes approches appropriées à ces données biométriques.

Ces dernières années, plusieurs approches ont été proposées dans la littérature pour un système d'authentification biométrique multimodal avec différents traits biométriques et avec différents mécanismes de fusion. La biométrie multimodale a fait l'objet d'une attention considérable de la part des communautés de recherche et du marché. Comme le coeur du système biométrique multimodal repose sur la fusion des informations provenant de différents traits biométriques, tous les travaux rapportés sur le système biométrique multimodal ont été limités à quatre niveaux de fusion différents.

Cette section passe en revue les résultats de certaines enquêtes récentes menées à ce jour dans le domaine de la fusion biométrique multimodale avec une brève description du travail effectué.

2.7.1 Fusion au niveau du capteur :

Patil et al. (PATIL et DHOLE, 2016) ont présenté une fusion du visage et palm Print au niveau du capteur, au niveau d'extraction des caractéristiques, au niveau du score et au niveau de décision. Pour l'extraction de caractéristiques, l'analyse LDA est utilisée pour la biométrie du visage et de palm Print.

Gowda et al. (GOWDA, KUMAR et IMRAN, 2018) ont proposés un système de vérification biométrique multimodal, le système est appliqué à tous les niveaux de fusion (niveau de capteur, niveau de caractéristiques, niveau de score et niveau de décision). Ils ont utilisé des méthodes physiologiques telles que le visage, l'empreinte de la paume, l'empreinte digitale, le poignet.

Pour l'extraction des caractéristiques de texture, ils ont utilisé un opérateur de texture bien connue, LPQ (Local Phase Quantization).

Ils ont effectué une fusion au niveau du capteur en adoptant des ondelettes HAAR, une fusion au niveau des caractéristiques avec la normalisation Z-Score, une fusion au niveau score utilisant la règle de somme simple et une fusion au niveau décision avec la règle AND. Pour le système de reconnaissance biométrique mis en œuvre, la stratégie de fusion au niveau des scores est plus performante que les autres techniques de fusion en termes de EER.

2.7.2 Fusion au niveau des caractéristiques :

Kisku et al. (KISKU, GUPTA et SING, 2011) ont proposé des approches biométriques multimodales qui intègrent le visage et le palmprint au niveau de l'extraction des caractéristiques. Le système a atteint un taux de reconnaissance de 98,75 % avec un taux de fausse acceptation de 0%.

Hassan et al. (HASSAN, SABER et AHMED, 2012) ont présenté une étude sur les systèmes biométriques multimodaux basés sur les veines de palmiers et la signature. L'auteur extrait ensuite les caractéristiques des deux modalités à l'aide d'opérations morphologiques et de l'algorithme SIFT (Scale Invariant Feature Transform). Ensuite, la fusion au niveau des caractéristiques pour les deux vecteurs utilise une règle de somme simple.

Les vecteurs de caractéristiques fusionnées sont soumis à une transformée en DCT (Discrete Cosine Transform) afin de réduire leurs dimensions. Le classificateur LVQ (Linear Vecteur Quantization) est utilisé avec des paramètres modifiés pour classer les différentes personnes dans la base de données.

Eskandari et al (ESKANDARI, TOYGAR et DEMIREL, 2014) ont proposé un système biométrique multimodal visage-iris utilisant la fusion au niveau du score et au niveau de caractéristiques. PCA(Principal Component Analysis), LDA, PCA sous-modèle, PCA modulaire et LBP(Local Binary Patterns) sont des méthodes d'extraction de caractéristiques globales et locales appliquées aux images du visage et de l'iris. La sélection de caractéristiques est effectuée à l'aide de PSO dans l'étape de fusion au niveau des caractéristiques afin de réduire la dimension des vecteurs de caractéristiques. Le schéma de fusion correspondant au niveau du score utilisant la règle de la somme pondérée et la normalisation Tanh.

Ahmad et al. (AHMAD et al., 2014) ont proposé la fusion d'informations de la biométrie du visage et de l'empreinte palmaire au niveau de caractéristique.

L'utilisation des méthodes PCA et LDA réduit considérablement la dimension du vecteur de caractéristiques en supprimant les données de redondance et de bruit tout en augmentant la puissance discriminante dans l'espace des caractéristiques fusionnées.

Dans cet article (BHARADI, PANDYA et NEMADE, 2014), un système biométrique multimodal basé sur l'iris et les empreintes digitales est proposé. L'extraction d'éléments de texture à l'aide d'ondelettes hybrides est effectuée. Les caractéristiques des empreintes digitales et de l'iris sont extraites à l'aide de la décomposition à plusieurs niveaux de l'image échantillon capturée à l'aide d'une nouvelle famille d'ondelettes appelée hybride. Dans cet article, le classificateur KNN est utilisé pour la reconnaissance unimodale des empreintes digitales et la reconnaissance multi-instance de l'iris. Les vecteurs caractéristiques de l'iris et de l'empreinte digitale sont combinés à l'aide de la technique de fusion décisionnelle.

L'analyse FAR-FRR est effectuée et les résultats suggèrent que les ondelettes hybrides ont une bonne capacité d'extraction des caractéristiques de texture et que le système proposé a obtenu une ration de classification correcte pouvant atteindre 77%.

Dans (BOUZOUINA et HAMAMI, 2017) la fusion dans le niveau de caractéristiques de deux extracteurs vecteurs, appelés filtres de Gabor et moment Zernike, via la sélection de caractéristique utilisant un algorithme génétique afin de combiner les propriétés des méthodes globales et locales. Plus tard, les auteurs ont étudié une fusion au niveau des scores des modalités visage et iris après normalisation à l'aide de la méthode de classification SVM.

Les performances de ce système sont testées à l'aide de la base de données CASIA-IrisV3-Interval. Le taux de reconnaissance obtenu avec cette base de données est de 98.8 %.

Modak et al. (MODAK et JHA, 2017) ont proposé un système biométrique multimodal combinant le visage et la main au niveau des caractéristiques.

Le système proposé traite la fusion dans deux scénarios, l'un étant la fusion des coefficients PCA et LDA du visage et l'autre est la fusion de la biométrie du visage et de la main. Le résultat de l'expérience montre que les performances de la fusion basée sur LDA sont beaucoup plus élevées que celles de l'adaptateur basé sur PCA.

Bayram et al. (BAYRAM et BOLAT, 2018) combinent les modalités biométriques de l'oreille, de visage et de visage thermique en utilisant deux méthodes de fusion qui correspondent au niveau du score et à la fusion au niveau des caractéristiques.

Pour chaque modalité, trois méthodes d'extraction de caractéristiques sont utilisées et quatre classificateurs différents (perceptron multicouche, arbre de décision, SVM et réseau neuronal probabiliste) sont formés.

2.7.3 Fusion au niveau des scores :

Dans (WANG et HAN, 2008), l'auteur a proposé un système robuste d'authentification biométrique multimodale intégrant les données de l'iris, du visage et l'empreinte palmaire au niveau du score à l'aide de plusieurs SVM.

En 2011, Mezai et al. (MEZAI, HACHOUF et BENGHERABI, 2011) ont combiné les scores d'appariement du visage et de la voix en utilisant la théorie de Dempster-Shafer (DS) pour la vérification de personne. Cette méthode qui était largement utilisée dans la fusion des classificateurs mais était peu utilisée dans la fusion biométrique multimodale. DS transforme les scores obtenus à l'aide d'algorithmes de vérification faciale et vocale en preuves, puis les combine. Leurs expériences ont été menées sur la base de données XM2VTS Benchmark, et les résultats obtenus ont montré que le HTER (Half Total Error Rate) de la fusion proposée variait de 0,433% à 2,875%. Cependant, les performances des classificateurs de visage et de voix varient de 1,88% à 6,22% et de 1,148% à 6,208% respectivement.

Maryam et al. (MARYAM et ONSSEN, 2012) ont proposé fusion du visage et de l'iris pour obtenir un système de reconnaissance robuste.

Les scores de visage et d'iris sont normalisés en utilisant la normalisation de la tangente hyperbolique (\tanh), puis la fusion de ces deux modalités est effectuée en utilisant la règle de la somme pondérée.

Ola M. Aly et al. (OLA M. ALY et ONSI, 2013) ont proposé une nouvelle approche pour la combinaison adaptative de la biométrie iris, palmprint, fingers_knuckle afin de garantir plusieurs niveaux de sécurité. La règle de fusion au niveau du score est adaptée à l'aide de PSO.

Cherifi et al. (CHERIFI, HAFNAOUI et NAIT-ALI, 2015) ont proposé d'utiliser un algorithme hybride GA-PSO pour trouver les poids optimaux associés aux modalités fusionnées au niveau du score.

Parkavi et al. (PARKAVI, BABU et KUMAR, 2017) ont combiné l'empreinte digitale et l'iris au niveau du score pour l'identification automatique d'un individu. Les techniques de minutiae et de détection de contour sont utilisées dans ce but.

2.7.4 Fusion au niveau des décisions :

Zhu et al. (ZHU et ZHANG, 2010) ont proposé un système d'identification biométrique multimodal basé sur la géométrie du doigt, l'impression d'articulation et l'empreinte de la paume de la main. Après avoir acquis l'image de la main, les auteurs procèdent à la localisation de la paume de la main sur une image et des doigts à une autre image. Ils binarisent l'image de la main pour détecter les contours et les points clés permettant de localiser les deux régions d'intérêt ROI (Region Of Interest). Ensuite, ils extraient les caractéristiques de chaque ROI : les caractéristiques géométriques des 4 doigts, l'impression d'articulation et les points clés de la paume de la main.

L'impression d'articulation est déterminée en fonction de l'opérateur Sobel pour détecter les limites et la transformation en ondelettes pour la suppression du bruit. Ensuite, l'opérateur KLT est utilisé pour détecter les points clés de la paume qui seront également caractérisés.

Lors de la phase de comparaison, les auteurs ont utilisé le calcul de la distance d'Euler entre les vecteurs de caractéristiques géométriques du doigt. En outre, les auteurs ont utilisé la fonction cosinus pour générer un taux de similarité entre les deux empreintes d'articulation en question. Pour générer la décision, les auteurs suivent une architecture hiérarchique en commençant par vérifier la géométrie du doigt en question avec celle de référence, si la décision est positive, les auteurs vérifient ensuite la similarité entre les deux impressions d'articulation en question pour terminer la validation de la similitude entre la paume en question et la paume de référence en se basant sur les points-clés extraits.

Abdolahi et al. (ABDOLAH, MOHAMADI et JAFARI, 2013) ont proposé un système biométrique multimodal qui fusionne l'iris et les empreintes digitales au niveau de la décision, basé sur la logique floue. Après amélioration de l'image, binarisation et squelettisation, les auteurs procèdent à l'extraction des caractéristiques des deux modalités. La méthode d'extraction des caractéristiques de l'empreinte digitale utilisée est basée sur les minuties.

L'extracteur d'iris est basé sur le filtre de Gabor pour générer un code binaire. Dans le module de comparaison de processus d'identification d'iris, une distance de Hamming est utilisée pour compter le nombre de bits différents entre les images d'iris. La même méthode de comparaison est utilisée dans le module du système d'identification d'empreinte digitale pour générer le nombre de minuties qui ne sont pas confondues entre les deux images d'empreinte digitale.

Un système d'interférence floue est utilisé pour générer la décision finale de chaque système biométrique monomode. Cette décision est générée à la base d'une fonction d'appartenance triangulaire, la distance de Hamming calculée et une règle définie par les auteurs sont utilisées pour générer une décision pour chaque processus. Par la suite, les auteurs ont utilisé des règles floues avec les deux décisions générées pour fournir la décision finale de leur système biométrique multimodal. Le taux d'erreur de ce système est de 1.7%.

Barbu et al. (BARBU, CIOBANU et LUCA, 2015) ont proposé un système biométrique multimodal basé sur l'iris, la voix et le visage. La fusion biométrique a été réalisée au niveau de la décision, le modèle biométrique multimodal obtenu a fourni un taux de reconnaissance amélioré.

Sireesha et al. (SIREESHA et SANDHYARANI, 2016) a combiné deux modalités biométriques, à savoir iris et les empreintes digitales. Les caractéristiques sont extraites de ces données biométriques et combinées pour former le vecteur de caractéristiques final. Elles sont ensuite classées à l'aide de différents classificateurs. La sortie de ces classificateurs est améliorée sur la base du classificateur de décision finale proposé dans ce cas, qui effectue une fusion de décision sur la base des décisions individuelles fournies par chaque classificateur.

2.8 Conclusion

Ce chapitre a été consacré à exposer les différentes formes de multimodalité, les différents scénarios et niveaux de fusion, les modes de fonctionnement d'un système biométrique multimodal. Dans la dernière section, nous avons un aperçu des travaux récents sur les systèmes biométriques multimodaux.

Pour comprendre le principe de fonctionnement du système biométrique, nous avons commencé à proposer un nouveau système unimodal basé sur la modalité de la main. Ce système est décrit dans le chapitre suivant.

Chapitre 3

Identification des personnes à base de la main

3.1 Introduction

La reconnaissance et la vérification personnelles sont des fonctions très importantes dans de nombreux systèmes de contrôle d'accès. La technologie biométrique est une nouvelle méthode de reconnaissance de l'identité d'une personne basée sur une base de données déjà établie de caractéristiques physiologiques ou comportementales (JAIN, ROSS et PRABHAKAR, 2004; DUTA, 2009). Habituellement, ces caractéristiques physiologiques doivent être uniques, invariantes, continues et permanentes. Récemment, diverses techniques de biométrie ont été proposées dans la littérature, telles que les empreintes digitales, la géométrie de la main, les empreintes de paume, la reconnaissance du visage, l'iris et la parole (JAIN, ROSS et PRABHAKAR, 2004).

Parmi tous les types de techniques biométriques, la biométrie de la main a attiré l'attention des chercheurs et a été étudiée depuis longtemps (DUTA, 2009). Comme souvent noté dans les littératures, la biométrie de la main est attrayante en raison de ses avantages suivants (YÖRÜK, DUTAĞACI et SANKUR, 2006) :

1. La main peut être facilement capturée d'une manière relativement conviviale en utilisant une caméra CCD (Charge-Coupled Device) conventionnelle.
2. Cette technologie est plus acceptable par le public dans la vie quotidienne principalement parce qu'il manque de lien étroit avec les applications médico-légales (DUTA, 2009) (JAIN, ROSS et PRABHAKAR, 2004). Enfin, il y a eu un intérêt récent dans la fusion de différentes données biométriques pour augmenter les performances du système (ROSS et JAIN, 2003).
3. Les informations sur les mains peuvent être extraites en utilisant des images basses résolutions.

3.2 Principaux problèmes de la modalité de la main

La biométrie de la main est une modalité conçue pour les applications à sécurité moyenne. Par conséquent, elle présente divers problèmes décrits

ci-dessous :

1. Problèmes liés aux conditions environnementales :

Il s'agit de la variation de l'arrière-plan et des conditions d'éclairage (éclairage bas / fort) (GUO et al., 2012). En fait, ce sont des défis très importants pour les systèmes de reconnaissance de la forme de la main et de palmprint.

Quelques exemples de problèmes liés aux conditions environnementales sont illustrés à la figure 4.1.



FIGURE 3.1 – Quelques exemples de problèmes environnement (GUO et al., 2012).

2. Problèmes liés à l'acquisition :

L'acquisition peut causer des bruits dus à la pression des utilisateurs sur la plaque de verre ou aux déformations attribuées aux dispositifs d'acquisition sans contact. En fait, les performances d'un système biométrique sont sensibles à la qualité et aux variations des échantillons (CHARFI, 2017).

Par conséquent, les échantillons bruyants peuvent affecter la précision du système. La figure 4.2 illustre quelques exemples de difficultés d'acquisition.



FIGURE 3.2 – Quelques exemples de difficultés d'acquisition.

3. Problèmes liés au temps d'exécution :

Les systèmes de reconnaissance biométriques basés sur la main doivent permettre de prendre une décision aussi rapidement que possible afin de faciliter leur utilisation dans les applications en temps réel (KHIARI, 2016).

4. Problèmes liés aux occlusions :

La main est un objet articulé ayant plusieurs degrés de liberté. En effet, les occlusions peuvent comprendre des doigts repliés, un poing fermé, une paume fermée non parallèle au plan de l'image (CHARFI, 2017). De plus, la taille de la main peut changer ; la position et l'orientation

affectent la clarté des régions d'intérêt de l'empreinte de la paume ou des régions des doigts.

La présence d'objets tels que bagues, bracelets, etc. est également l'une des difficultés inévitables dans le cas d'une utilisation libre. La figure 4.3 présente quelques exemples de problèmes d'occlusion.



FIGURE 3.3 – Quelques exemples de problèmes d'occlusion.

3.3 Les traits biométriques de la main

Plusieurs systèmes de reconnaissance des mains ont été proposés. Ils décrivent les différentes parties de la main qui constituent l'identité de la personne. Cette section présente certaines caractéristiques de la main, à savoir la forme de la main, sa géométrie, l'empreinte de la main (palmprint) et les doigts.

3.3.1 La forme de la main

La forme de la main (ou la silhouette de la main, qu'elle est présentée à la figure 4.4) a suscité peu d'intérêt dans la littérature, afin de reconnaître les individus.



FIGURE 3.4 – Forme de la main.

Le principal avantage de cette modalité est qu'elle est simple et facile à capturer. De plus, elle n'est pas très sensible à l'état de la main, c'est-à-dire que la saleté et les coupures n'empêcheront pas le fonctionnement du système.

Cependant, il présente également des défauts qui peuvent rendre le système inexact.

En fait, différentes personnes peuvent avoir des formes de main presque similaires. En outre, les bijoux et les accessoires peuvent représenter un défi pour extraire les informations de forme de la main.

Historiquement, Jain et Duta (JAIN et DUTA, 1999) ont été les premiers à analyser une forme déformable et à développer une méthode dans laquelle les formes de main sont stockées et comparées en fonction de l'erreur d'alignement moyenne.

Amayeh et al. (AMAYEH, 2006) ont présenté la conception et la mise en œuvre d'un système de vérification basé sur la forme de la main en utilisant les moments de Zernike d'ordre élevé. Le système fonctionne sur des images de silhouette de main 2D acquises par placer la main sur une table d'éclairage plane sans aucun point d'appui, augmentant ainsi la facilité d'utilisation.

Juan et al. (JUAN MANUEL RAMIREZ-CORTES et BAEZ-LOPEZ, 2008) ont proposé l'utilisation de l'opérateur morphologique *pecstrum* ou du spectre spectral, comme extracteur de caractéristiques pour un système de reconnaissance de la main basé sur la forme. Les propriétés d'invariance à la rotation et à la position du *pecstrum* apportent de la souplesse au système en permettant à l'utilisateur de poser naturellement la main sans contraintes supplémentaires.

Dutagaci et al. (DUTAGACI, SANKUR et YÖRÜK, 2008) ont proposé un algorithme de normalisation pour ajuster les doigts aux positions standard et extraites les caractéristiques de forme en appliquant LDA sur les coefficients de la DFT (Discrete Fourier transform) du contour. Hu Rong-Xiang et al. (HU et al., 2012) ont proposé une méthode de reconnaissance de la forme de la main basée sur deux représentations de formes classiques, à savoir les SC (Shape Context) et les IDSC (Inner Distance Shape Contexts)). Il a une bonne capacité à capturer les caractéristiques discriminantes de la forme de la main et peut traiter le problème de correspondance inexacte des points de repère de la main.

3.3.2 Géométrie de la main

La géométrie des mains a également fait l'objet de beaucoup d'attention pour la biométrie des mains. Les caractéristiques géométriques présentent les principales caractéristiques de la géométrie de la main qui sont adoptées dans la plupart des systèmes biométriques de la main.

Ces caractéristiques ont le mérite d'être relativement invariantes quant aux orientations des doigts et au positionnement global de la main. Parmi ces différentes mesures géométriques, on peut citer la taille de la paume, la longueur de la main ou des doigts, la largeur et le périmètre de la main et des doigts, etc. (GUO et al., 2012), comme indiqué dans la figure 4.5.

En fait, c'est le plus haut nombre adopté dans la littérature pour améliorer les performances du système de reconnaissance de personnes.

Bien que les caractéristiques géométriques soient faciles à extraire, elles ne sont pas assez discriminantes pour être utilisées pour des tâches d'identification/vérification de haute sécurité.

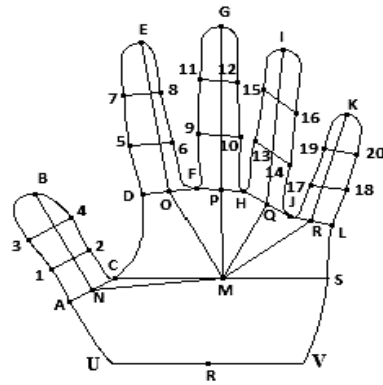


FIGURE 3.5 – Caractéristiques géométriques de la main (GUO et al., 2012).

De nombreuses recherches sur la reconnaissance personnelle basées sur la géométrie de la main ont été examinées et étudiées.

Par exemple, Bulatov et al. (BULATOV et al., 2004) décrit un système sans chevilles dans lequel 30 mesures géométriques sont extraites des images de la main. En plus des largeurs, des périmètres et des zones des doigts; ils incorporeraient aussi les rayons de cercles des doigts et le rayon du plus grand cercle inscrit de la paume.

Saxena Nidhi et al. (SAXENA et al., 2012) ont proposé un système de la géométrie de la main composé de la longueur et de la largeur des doigts et de la largeur d'une paume. Les utilisateurs peuvent placer leurs mains librement sans avoir besoin de chevilles pour fixer le placement des mains.

Luque-Baena et al. (LUQUE-BAENA et al., 2013) ont extrait 403 caractéristiques géométriques, y compris des surfaces, des périmètres, une mesure de la circularité, la compacité, la proportion d'aspect, etc.

Agbinya Johnson (AGBINYA, 2018) a proposé une nouvelle équation de la géométrie de la main qui modélise la relation entre les dimensions de la paume définie par le bout des doigts et la base du doigt. Celles-ci sont liées à la longueur de la main. La longueur est approximativement la longueur du majeur mesurée à partir de l'articulation du poignet et la largeur de la paume. On montre qu'il existe une constante qui relie la géométrie de la main de chaque être humain.

3.3.3 L'empreinte de la main (palmprint)

L'empreinte de la main est l'une des techniques de reconnaissance biométrique la plus couramment utilisée, en particulier pour la criminalité. L'empreinte de la main présente la partie interne de la main (la partie invisible lorsque la main est fermée) (figure 4.6).

L'utilisation des caractéristiques palmprint dans le processus d'identification a été initialement proposée par Shu et Zhang (WEI SHU et ZHANG, 1998).

La modalité Palmprint peut présenter différents types de fonctionnalités pouvant être exploitées pour la reconnaissance des personnes (WEI SHU

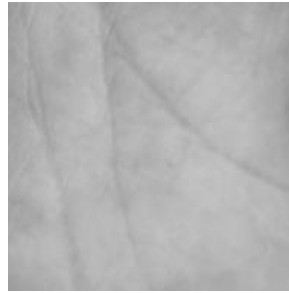


FIGURE 3.6 – L’empreinte de la main.

et ZHANG, 1998) : lignes principales, rides ou lignes secondaires et crêtes. Toutes ces caractéristiques peuvent caractériser une personne puisqu’ils sont discriminantes et stables dans le temps. Concernant les rides ou les lignes secondaires, elles sont plus fines et plus irrégulières que les lignes principales.

Plusieurs approches de reconnaissance des empreinte de la main ont été proposées dans la littérature. Par exemple, Zuo et al. (ZUO et al., 2010) ont proposé une représentation compacte des caractéristiques d’orientation des lignes de la paume sur plusieurs échelles. En outre, ils ont étudié une nouvelle méthode appelée SMCC (Sparse Multiscale Competitive Code).

Wang et al. (WANG, LEI et WANG, 2012) ont proposé une méthode de texture globale pour la reconnaissance des empreintes de la main basée sur la décomposition à l’aide des fonctionnalités d’ondelettes 2D-Gabor. Au début, les images d’empreinte de la main sont prétraitées et normalisées dans la position, l’orientation et l’éclairage.

Nordin et al. (NORDIN et HUSSEIN, 2014) ont étudié une méthode de reconnaissance des empreintes de la main basée sur les moments invariants et les ondelettes. Ce modèle a montré des résultats prometteurs sans les effets de rotation, de translation et de mise à l’échelle des objets, car il est associé à l’utilisation d’une bonne description des caractéristiques de forme.

Wu et al. (WU et ZHAO, 2015) ont proposé une méthode de vérification de l’empreinte de la main basée sur les descripteurs SIFT (Scale Invariant Feature Transform). En fait, un modèle est construit en approximant des images palmaires déformées de manière non linéaire avec des régions stables déformées linéairement par morceaux. Par conséquent, la méthode KPBG (KeyPoint based Block Growing) est proposée. Elle est basée sur une approche par consensus itératif d’échantillon M-estimateur utilisant les caractéristiques SIFT.

Luo et al. (LUO et al., 2016) ont présenté un nouveau descripteur de structure LBP appelé LLDP (Local Line Directional Patterns) pour la reconnaissance d’empreintes de la main. Le but de ce descripteur est de coder la structure d’un voisinage local à partir des informations analysées de la ligne directionnelle, calculées dans 12 directions à l’aide des filtres MFRAT (Modified Finite Radon Transform) et Gabor.

Attallah et al.(ATTALLAH, SERIR et CHAHIR, 2018) ont proposé une nouvelle approche de la reconnaissance de l’empreinte de la main, basée sur la fusion des caractéristiques spirale et les filtres LBP, et sur l’utilisation du

mRMR (minimum Redundancy Maximum Relevance) pour identifier les meilleures caractéristiques.

3.3.4 Les doigts

La forme de la main varie d'une personne à l'autre en raison de l'articulation des doigts. En conséquence, certains chercheurs ont proposé de segmenter les doigts de la main pour modéliser séparément la forme des doigts (voire figure 4.7).

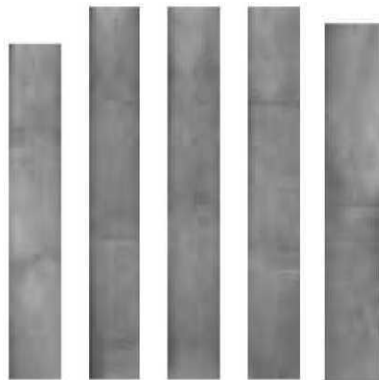


FIGURE 3.7 – Cinq surfaces de doigts.

Oden et al. (ODEN, ERCIL et BUKE, 2003) ont proposé de modéliser la forme du doigt de chaque personne en utilisant des fonctions polynomiales implicites du quatrième degré. De plus, les invariants de Keren ont été extraits des polynômes ajustés pour être utilisés comme invariants de caractéristiques dans les transformations affines (KEREN, 1994).

Cependant, Ribaric et Fratric (RIBARIC et FRATRIC, 2005) ont segmenté la main en six sous-images décrivant la région de la paume et 5 surfaces des doigts. À partir de ces sous images, les caractéristiques les plus importantes sont extraites sur la base de la transformation de Karhunen-Loeve pour acquérir des doigts propres fusionnés à des paumes propres.

Dans un autre travail, Xiong et al. (XIONG et al., 2005) ont identifié plusieurs doigts alignés selon un modèle elliptique, par transformations euclidiennes.

3.4 Système proposé de reconnaissance de la forme de la main à l'aide des moments

Dans cette section, nous proposons la combinaison des moments de Hu et de Legendre, en tant que nouvel extracteur de caractéristiques pour un système de reconnaissance de la main basé sur la forme.

Les propriétés d'invariance à la rotation et la position des moments invariants offrent une flexibilité au système en permettant à l'utilisateur de poser

naturellement la main sans contraintes supplémentaires. Ceci est un avantage par rapport à certains systèmes, dans lesquels une position de la main fixe à l'aide de piquets est requise.

Le système de reconnaissance de la main proposé est indiqué dans la figure 4.8 et se compose de trois modules principaux, qui sont :

Le prétraitement de l'image, l'extraction des caractéristiques et le classificateur du plus proche voisin (classificateur K-NN).

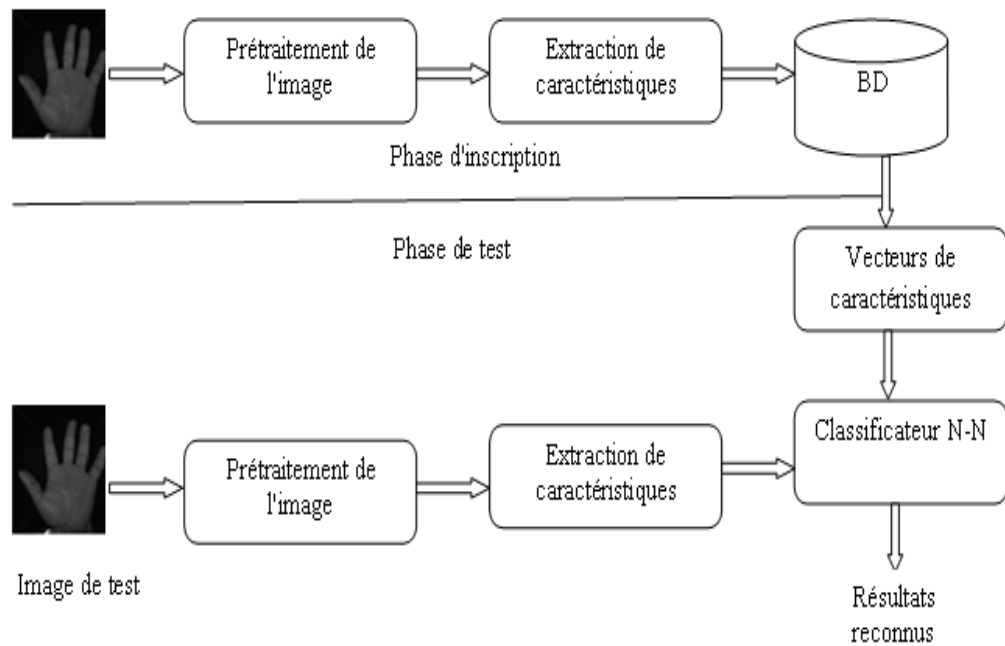


FIGURE 3.8 – Schéma du système proposé.

3.4.1 Prétraitement d'image

Le module d'inscription comprend toutes les opérations de prétraitement requises pour faciliter l'extraction des caractéristiques au cours du module d'extraction des caractéristiques, et disponible pour le module d'identification.

L'étape de prétraitement est requise pour l'image lorsque cette image est à des fins de reconnaissance ou d'inscription.

Comme l'image acquise est une image en niveaux de gris, elle est convertie en image binaire par seuillage. La valeur de seuil est calculée automatiquement à l'aide de la méthode Otsu (OTSU, 1978).

Ensuite, la bordure de la silhouette de la main est lissée en utilisant une ouverture et une fermeture morphologiques. Le résultat est montré dans la figure 4.9.

L'ouverture lisse généralement l'objet de contour, casse les isthmes étroits et éliminent les saillies minces.



FIGURE 3.9 – Exemple d’image du processus de prétraitement.

La fermeture a également tendance à lisser des sections de contours, mais, opposée à l’ouverture, elle fusionne les pauses étroites et les gouffres longs et minces, élimine les petits trous et comble les trous dans le contour.

3.4.2 Extraction des caractéristiques

Après toutes les étapes de prétraitement nécessaires à l’extraction des caractéristiques, le système proposé est exécuté soit en phase d’inscription, soit en phase d’identification. Au cours de la phase d’inscription, la main est stockée dans la base de données des caractéristiques de la main de la personne, tandis que dans la phase d’identification, une correspondance est requise.

Le moment invariant sera la stratégie de base de l’extraction des caractéristiques de la main, mais nous ne devons pas oublier qu’il existe de nombreuses méthodes pour extraire des caractéristiques de la main, mais le moment est le plus fiable et facile.

Dans l’approche proposée, deux types de vecteurs de caractéristiques sont calculés, les vecteurs de caractéristiques invariants de Hu et les vecteurs de caractéristiques de Legendre.

Les moments Hu fournissent un vecteur de caractéristiques de taille sept et les moments Legendre contribuent au vecteur de caractéristiques de taille 12.

Celles-ci sont décrites dans la section suivante.

Moments Invariants de Hu

Hu a introduit la famille d’invariant portant son nom en utilisant les moments géométriques.

Soit une image bidimensionnelle binaire $M \times M f(x, y); x, y = 0, \dots, M - 1$, le moment géométrique d’ordre $(p + q)$ est défini par :

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} x^p y^q f(x, y) \quad (3.1)$$

Avec $p, q = 0, 1, 2, \dots$

Pour conserver la dynamique constante m_{pq} des images de tailles différentes, le plan-images $M \times M$ est d’abord projeté sur un carré défini par $x \in [-1, +1]$ et $y \in [-1, +1]$.

Cela implique que les emplacements de la grille ne seront plus des entiers mais des valeurs réelles comprises dans la plage $[-1, +1]$. Donc, la définition de m_{pq} devient (HU, 1962) :

$$m_{pq} = \sum_{x=-1}^{+1} \sum_{y=-1}^{+1} x^p y^q f(x, y) \quad (3.2)$$

Ainsi définis, les moments géométriques ne sont pas invariants en rotation et en translation. Par contre les moments géométriques centrés sont invariants en translation.

Soit x_0 et y_0 le barycentre de l'image, on définira les moments géométriques centrés par la fonction :

$$u_{pq} = \sum_{x=-1}^{+1} \sum_{y=-1}^{+1} (x - x_0)^p (y - y_0)^q f(x, y) \quad (3.3)$$

L'invariance à l'échelle s'obtient en normalisant ces moments. On obtient alors les moments centraux normalisés η_{pq} définis par :

$$\eta_{pq} = \frac{u_{pq}}{u_{00}^\gamma} \quad (3.4)$$

Avec $\gamma = \frac{p+q}{2} + 1$

Cependant les moments géométriques centrés ne sont pas invariants en rotation. Pour résoudre cette problématique Hu forma une famille d'invariants en rotation au moyen de combinaisons de moments géométriques centrés normalisés.

Hu (HU, 1962) a dérivé un ensemble de sept moments qui sont invariants à la translation, l'orientation et d'échelle. Les invariants de Hu sont étendus par Maitra (MAITRA, 1979) pour être invariants sous le contraste de l'image. Plus tard, Flusser et Suk (FLUSSER et SUK, 1993) ont dérivé le moment invariant, qui est invariant sous la transformation affine générale.

Les équations des moments de Hu sont définies comme suit :

$$\phi_1 = \eta_{20} + \eta_{02} \quad (3.5)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (3.6)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (3.7)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (3.8)$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (3.9)$$

$$\begin{aligned} \phi_6 = & (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ & + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned} \quad (3.10)$$

$$\begin{aligned} \phi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2(\eta_{21} + \eta_{03})^2] \end{aligned} \quad (3.11)$$

Ces sept moments sont dérivés de moments du deuxième et du troisième ordre.

Les moments de Hu sont des descripteurs très efficaces grâce à leur invariance et à leur discriminant. Ils ont plusieurs caractéristiques qui peuvent être résumées comme suit :

1. Les sept moments sont invariants aux translations, aux changements d'échelle et aux rotations.

Dans l'exemple présenté sur la figure 4.10, les moments sont appliqués à l'image d'origine et à leurs images obtenues par mise à l'échelle et rotation. Les moments calculés pour l'image à l'échelle 50% et l'image avec une rotation de 45° sont très proches des moments calculés pour l'image d'origine.

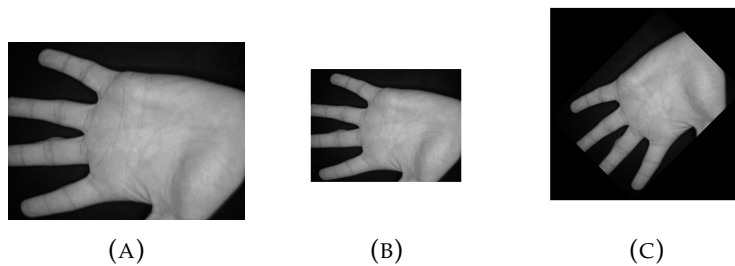


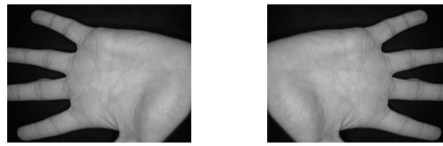
FIGURE 3.10 – (A) Image originale, (B) Image à échelle 50%, (C) Image avec rotation 45°.

Ces résultats confirment que les moments sont invariants : aux changements d'échelle et aux rotations en plus de leur invariance à la translation (DAVIS et BRADSKI, 1999).

2. Les moments de Hu sont également invariants aux réflexions, l'invariance des réflexions peut être problématique lorsqu'il s'agit de reconnaître des images en miroir (voir figure 4.11).

TABLE 3.1 – Moments des images (A), (B) et (c)

Moments	Image originale	Image à échelle 50%	Image avec rotation 45°
Moments 1	0.001253500572694	0.001253372743153	0.001253440173500
Moments 2	0.000000197102029	0.000000197052697	0.000000196929484
Moments 3	0.000000000004506	0.000000000004510	0.000000000004520
Moments 4	0.000000000055269	0.000000000055237	0.000000000055280
Moments 5	0.000000000000000	0.000000000000000	0.000000000000000
Moments 6	0.000000000000025	0.000000000000025	0.000000000000025
Moments 7	0.000000000000000	0.000000000000000	0.000000000000000



(A)

(B)

FIGURE 3.11 – (A) Image originale, (B) Image miroir.

Le septième moment est calculé pour l'image originale (A) et l'image miroir (B) dans la Figure 4.11, les résultats sont comme suit :

Moments 7 de l'image (A) = 0.000000000000000

Moments 7 de l'image (B) = 0.000000000000000

On voit que les septièmes moments de l'image d'origine (A) et de l'image miroir (B) ont la même valeur.

3. Les moments de Hu peuvent être utilisés comme descripteurs de la rectangularité, de l'ellipticité et de la circularité.

Les moments de Legendre

Les moments de Legendre ont été introduits par Teague (ALNASIRI et al., 2009) (SRIDHA et KRISHNA, 2012). Les moments de Legendre appartiennent à la classe des moments orthogonaux et ont été utilisés dans plusieurs applications de reconnaissance de motifs.

Ils peuvent être utilisés pour atteindre une mesure de redondance quasi nulle dans un ensemble de fonctions de moment, de sorte que les moments correspondent à des fonctions indépendantes caractéristiques de l'image.

La définition des moments de Legendre est une forme de projection de la fonction d'intensité de l'image dans les polynômes de Legendre. Les moments bidimensionnels de Legendre d'ordre $(p + q)$, avec la fonction d'intensité d'image $f(x, y)$, sont définis comme suit :

$$L_{pq} = \frac{(2P + 1)(2q + 1)}{4} \int \int_{-1}^1 P_p(x)P_q(y)f(x, y)dx dy \quad (3.12)$$

Où polynôme Legendre, $P_p(x)$, d'ordre p est donné par (ALNASIRI et al., 2009) :

$$P_p(x) = \sum_{k=0}^p \left\{ (-1)^{\frac{p-k}{2}} \frac{1}{2^p} \frac{(p+k)!x^k}{\left(\frac{p-k}{2}\right)! \left(\frac{p+k}{2}\right)!k!} \right\} \quad (3.13)$$

La relation de récurrence des polynômes de Legendre, $P_p(x)$, est donnée comme suit (SRIDHA et KRISHNA, 2012) :

$$P_p(x) = \frac{(2p-1) \times p_{p-1}(x) - (p-1)p_{p-2}(x)}{p} \quad (3.14)$$

Où $P_0(x) = 1, P_1(x) = x$ et $p > 1$. Depuis la région de la définition des polynômes de Legendre est l'intérieur de $[-1, 1]$, une image carrée de $N \times N$ pixels avec la fonction d'intensité $f(i, j)$, $0 \leq i, j \leq (N-1)$, est mise à l'échelle dans la région de $-1 < x, y < 1$. Dans le résultat de ceci, l'équation 4.12 peut maintenant être exprimée sous forme discrète comme (YORUK et al., 2006) :

$$I_{pq}(x) = \lambda_{pq} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p_p(x_i) p_q(y_j) f(i, j) \quad (3.15)$$

Où la constante de normalisation, $\lambda_{pq} = \frac{(2p+1)(2q+1)}{N^2}$ x_i et y_j dénote les coordonnées de pixels normalisées dans la plage de $[-1, 1]$, qui sont donnés par (ALNASIRI et al., 2009) :

$x_i = \frac{2i}{N-1} - 1$ et $y_j = \frac{2j}{N-1} - 1$ Dans cette méthode les Legendre moment d'ordre quatre sont pris en compte pour le calcul des moments.

3.4.3 Classificateur K-NN

Le classificateur K-NN est utilisé pour comparer le vecteur de caractéristiques de l'image prototype et les vecteurs de caractéristiques stockés dans la base de données. Il est obtenu en trouvant la distance entre l'image prototype et la base de données. Soit $C_1, C_2, C_3 \dots C_k$ les k clusters de la base de données. La classe est trouvée en mesurant la distance $T(x(q), C_k)$ entre $x(q)$ et le $k^{\text{ième}}$ cluster C_k . Le vecteur de caractéristiques avec une différence minimale est le vecteur correspondant le plus proche. Il est donné par (ALNASIRI et al., 2009) :

$$T(x(q), C_k) = \min \{ \|x(q) - x\| : x \in C_k \} \quad (3.16)$$

3.5 Les résultats expérimentaux

3.5.1 Base de données

La base de données utilisées ici est CASIA-MS-PalmprintV1 (*CASIA-MS Database*). Cette base de données a été acquise à l'aide d'un dispositif d'imagerie sans contact et contient des images provenant de 100 utilisateurs. Six images ont été acquises chez chaque utilisateur et ces images ont été acquises lors de deux sessions d'acquisition de données différentes (trois images par session) avec un intervalle minimum d'un mois.

La base de données contient 7200 images à la main capturées à partir de 100 personnes différentes en utilisant un auto-conçu. Nous avons sélectionné 100 images de la main gauche; chaque personne a 6 et totalise jusqu'à 600. Nous utilisons ensuite une image de la main de chaque personne comme modèle. Les 500 autres sont utilisés comme échantillons d'apprentissage.

Les images sont redimensionnées dans 87×116 pixels. Quelques exemples d'images originales sont donnés dans la figure 4.12.

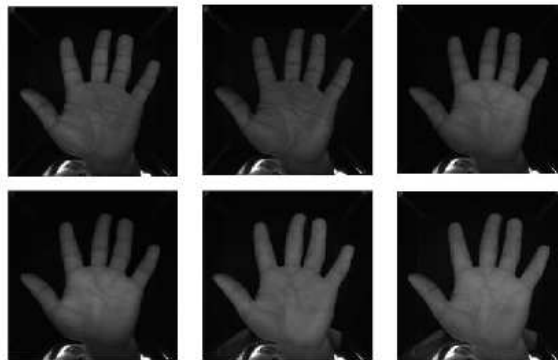


FIGURE 3.12 – Quelques exemples d'images de la main dans la base de données

3.5.2 Résultats expérimentaux

Afin de montrer les performances de l'algorithme proposé, plusieurs expériences ont été testées.

L'implémentation des différents algorithmes de notre système ainsi que le déroulement de tous les tests sont faits sous Matlab 2013.

Les expériences de reconnaissance des mains, basées sur des images de pré-traitement des mains, ont été effectuées sur huit tailles de population sélectionnées, à savoir des sous-groupes de population composés de 20, 35, 50, 100, 200, 300, 400 et 500 individus. Des tailles de population différentes nous a aidé à percevoir la performance de reconnaissance avec un nombre croissant d'individus. Un algorithme de boosting a été appliqué de sorte que plusieurs formations différentes de sous-ensembles (de tailles 20, 35, 50, 100, 200, 300,

400 et 500) ont été créées par choix aléatoire et leurs résultats moyennés ont été calculés. Les résultats sont donnés dans le tableau 4.2.

TABLE 3.2 – Performance de reconnaissance

Taille du groupe	Pourcentage de reconnaissance
20	100
35	99.03
50	98.08
100	97.20
200	96.89
300	96.00
400	95.40
500	94.03

Les résultats expérimentaux donnés dans le tableau 4.3 montrent que le taux de reconnaissance des moments combinés de Legendre et des moments invariants de Hu est supérieur au taux de reconnaissance des moments de Hu et de Legendre séparés.

TABLE 3.3 – Analyse de performance

Taille du groupe	Pourcentage de reconnaissance		
	Hu	Legendre	Hu+Legendre
20	98.34	99.56	100
35	96.45	98.23	99.03
50	95.85	97.65	98.08
100	94.30	96.65	97.20
200	93.00	96.00	96.89
300	92.20	95.36	96.00
400	91.03	94.32	95.40
500	90.24	93.68	94.03
Moyenne	93.93	96.43	97.08

3.6 Conclusion

Nous avons présenté une nouvelle méthode de reconnaissance basée sur la forme de la main qui combine les moments invariants de Hu et les moments de Legendre. En fait, l'utilisation d'un seul vecteur de caractéristiques n'est pas discriminante pour l'identification.

L'algorithme proposé présente certains avantages. Tout d'abord, l'orthogonalité et l'invariance géométrique permettent de minimiser la redondance de l'information et d'augmenter le pouvoir de discrimination. Une précision plus élevée peut être obtenue en utilisant des classificateurs plus complexes tels que les réseaux de neurones.

Puis, notre méthode ne nécessite que les contraintes les plus élémentaires de la configuration de l'acquisition et de la coopération avec les utilisateurs. Elle convient donc parfaitement aux applications réelles.

Cependant, le temps de calcul de la méthode proposée est encore un peu long pour une application réelle.

De toute évidence, les résultats expérimentaux démontrent l'efficacité de la méthode proposée.

La méthode est testée sur une base de données de 100 sujets en utilisant un classificateur K-NN, obtenant un taux moyen de reconnaissance de 97.08%.

Bien que l'approche proposée basée sur la forme de la main a donné des résultats acceptables, néanmoins les systèmes biométriques unimodaux souffrent de divers problèmes tels que le bruit dans les données détectées.

Les systèmes biométriques multimodaux constituent une approche relativement nouvelle pour surmonter ces limitations en assurant des améliorations importantes par rapport au système biométrique unimodal.

Le chapitre suivant présentera une méthode d'identification multimodale basée sur le PSO.

Chapitre 4

Approche multibiométrie basée sur le PSO adaptatif

4.1 Introduction

Les systèmes biométriques unimodaux souffrent de divers problèmes tels que le bruit dans les données détectées, la non-universalité, les similitudes entre classes et les attaques parodiques. Les systèmes biométriques multimodaux constituent une approche relativement nouvelle pour surmonter ces limitations en assurant des améliorations importantes par rapport au système biométrique unimodal, en matière de plus grandes précisions et de résistance à l'usurpation d'identité.

Comme illustré au chapitre II, il existe plusieurs façons de fusionner des systèmes biométriques. La fusion peut se faire au niveau des capteurs, des caractéristiques, des scores ou au niveau de la décision. Dans nos travaux, la fusion au niveau des scores est préférée car elle offre un meilleur compromis en termes d'information et de simplicité de la fusion.

Ce chapitre est consacré à la présentation du système d'authentification multimodale que nous avons proposé.

4.2 Description de l'approche proposée

Les différentes étapes de notre système biométrique multimodal sont présentées dans la figure 5.1.

Dans ce système, nous utilisons trois algorithmes d'extraction de caractéristiques (transformée de Gabor, DWT et transformée de Contourlet) pour extraire les principales caractéristiques biométriques.

Ensuite, pour réaliser la réduction de dimensionnalité des vecteurs de caractéristiques, PCA (Principal Component Analysis), LDA et KFA (Kernel Fisher Discriminant Analysis) sont utilisées.

Ensuite, l'algorithme de classification K-NN est appliqué afin de classer ses caractéristiques.

Enfin, la fusion des scores de ces trois modalités est effectuée en utilisant la règle de WS (Weighted Sum) et PSO.

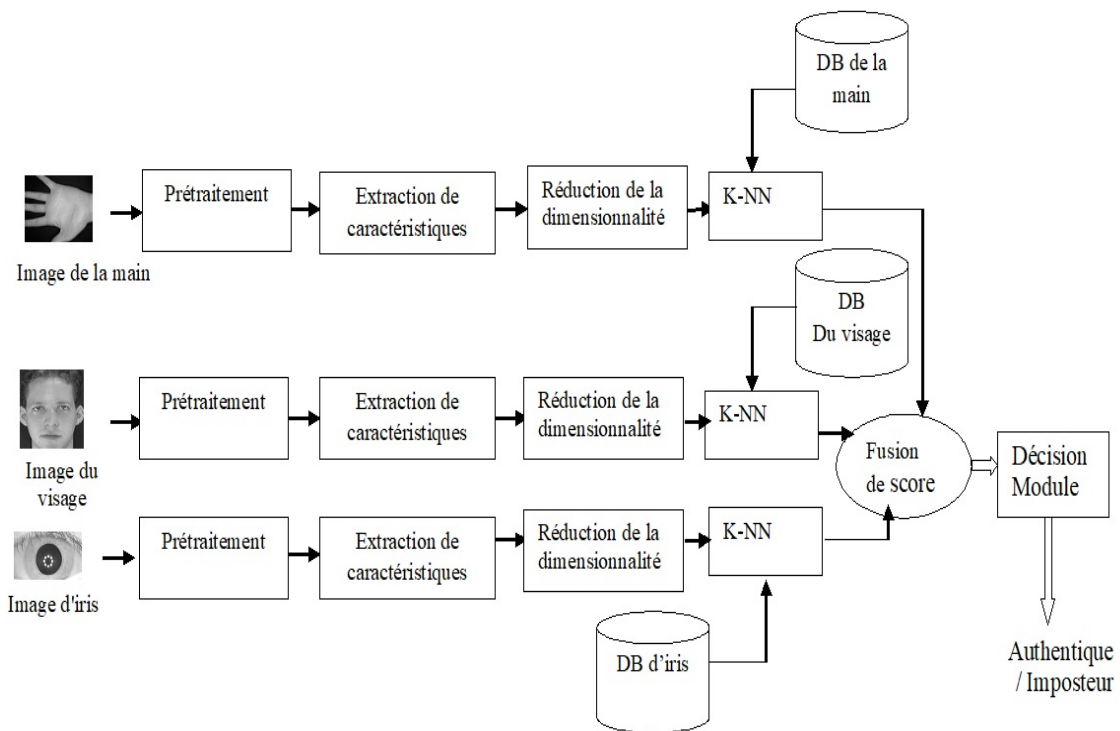


FIGURE 4.1 – Conception de base du système proposé.

4.2.1 Pré-traitement

Dans cette étape, les images des trois traits sont prétraitées. Le prétraitement des images de l'iris se fait avec l'algorithme de détection des contours de Canny (JOHAR et KAUSHIK, 2015). Le prétraitement des images de visage se fait par égalisation d'histogramme. Le prétraitement des images de la main consiste en la détection de régions d'intérêt (ROI) (RAJAGOPAL et MANOHARAN, 2015).

4.2.2 Techniques d'extraction de caractéristiques

Dans cette étape, les algorithmes d'extraction de caractéristiques suivants sont utilisés pour extraire les caractéristiques biométriques saillantes et former le modèle d'entité :

1. Filtre de Gabor :

Le filtre de Gabor est une méthode courante pour capturer des propriétés visuelles importantes, les caractéristiques multi-échelles, caractéristiques de fréquence spatiale multidirectionnelle (HAGHIGHAT MOHAMMAD et MOHAMED, 2013).

Dans le domaine spatial, le filtre de Gabor à deux dimensions est une

fonction du noyau gaussienne modulée par une onde plane sinusoidale complexe, représentée par l'équation suivante (HAGHIGHAT MOHAMMAD et MOHAMED, 2013) (MESHGINI, 2013) :

$$G(x, y) = \frac{f^2}{\pi\gamma\eta} e^{(-\frac{x^2+y^2\gamma^2}{2\sigma^2})} e^{(j2\pi f\acute{x}+\phi)} \quad (4.1)$$

$$\acute{x} = x \cos \theta + y \sin \theta \quad (4.2)$$

$$\acute{y} = -x \sin \theta + y \cos \theta \quad (4.3)$$

Où f est la fréquence du facteur sinusoidal, θ représente l'orientation de la normale aux bandes parallèles d'une fonction de Gabor, ϕ est le décalage de phase, σ est l'écart type et γ est le rapport d'aspect spatial. Dans cette méthode proposée, le filtre à ondelettes de Gabor appliqué sur les images de visages, iris et palmprint pour extraire les caractéristiques.

2. CT (Contourlet transform) :

La CT est une extension directionnelle simple des ondelettes (BISWAS et SIL, 2015). Elle se compose de deux étapes principales telles que la LP (Laplacian Pyramid) et les DFB (Directional Filter Banks).

Dans la première étape, LP est utilisé pour diviser l'image d'entrée en une image passe-bas et une image passe-bande, puis décomposer chaque image passe-bande par DFB.

Dans cette méthode proposée, chaque image prétraitée a été décomposée en utilisant une CT jusqu'à 4 niveaux, tandis que le filtre PKVA est utilisé comme LP et DFB.

La figure 5.2 montre un exemple de la CT de l'image de l'iris.

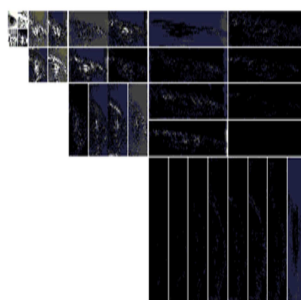


FIGURE 4.2 – CT d'une image d'iris jusqu'à 4 niveaux.

3. Transformée en ondelettes discrète (DWT) :

Les images pré-traitées du palmprint, du visage et de l'iris sont décomposées à l'aide de DWT jusqu'à trois niveaux afin d'obtenir le vecteur des coefficients d'ondelettes.

Le DWT est un outil qui peut être appliqué aux données discrètes pour obtenir une représentation à l'échelle différente des données originales (NIRGUDE et GENGAJE, 2015).

Un DWT 1-D consiste en un filtre passe-bas (L) et un filtre passe-haut (H) divisant une ligne de pixels en deux lignes de demi-taille. L'application des filtres aux images 2D dans les directions horizontale et verticale produit quatre sous-bandes (LL, LH, HL et HH).

La sous-bande LL est la partie la plus importante qui se rapproche l'image originale et les détails manquants sont filtrés dans les sous-bandes restantes. Les sous-bandes contiennent les bords horizontaux (LH), vertical (HL) et diagonal (HH) de la taille de l'échelle définie par l'ondelette. Les niveaux de décomposition supplémentaires sont appliqués uniquement à la sous-bande LL de l'image transformée au niveau précédent.

La figure 5.3 illustre la décomposition à trois niveaux d'une image échantillon.

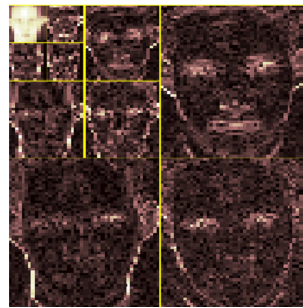


FIGURE 4.3 – Trois niveaux de décomposition d'une image.

4.2.3 Techniques de réduction de dimensionnalité

1. PCA : est un processus mathématique qui utilise une transformation orthogonale pour transférer un ensemble d'observations des variables corrélées dans un ensemble de variables non corrélées appelées composantes principales (BEHERA et GOVINDAN, 2014).

Le principe de PCA pourrait être décrit dans ces étapes (TURK et PENTLAND, 1991) :

- Chaque image de la base de données est convertie en un vecteur à une dimension
- Calculer la moyenne de chaque vecteur en utilisant l'équation suivante :

$$X_m = \frac{1}{M} \sum_{i=1}^M X_i \quad (4.4)$$

- Soustraire la moyenne de tous les vecteurs pour produire un ensemble de vecteurs à moyenne nulle, il est donné par cette équation :

$$\phi_i = X_i - X_m \quad (4.5)$$

- Calculer la matrice de covariance $C(N^2 \times N^2)$ en utilisant l'équation :

$$C = AA^T \quad (4.6)$$

Où, $A = [\phi_1, \phi_2, \dots, \phi_M](N^2 \times M$ matrice)

- Calculer les vecteurs propres u_i de la matrice de covariance C .
La matrice AA^T est très grande, considérez donc la matrice $A^T A$ ($M \times M$ matrix) et calculez les M meilleurs vecteurs propres u_i de AA^T tels que :

$$u_i = Ar_i \quad (4.7)$$

Où r_i est les vecteurs propres de $A^T A$

- Sélectionner uniquement les K vecteurs propres correspondant aux valeurs propres les plus élevées et formant un vecteur de caractéristique.
2. LDA : est une technique basée sur l'apparence utilisée pour la réduction de dimensionnalité et a enregistré une grande performance en reconnaissance biométrique (OH, YOO et PEDRYCZ, 2013). Cette méthode fonctionne sur le même principe que la méthode PCA.

L'objectif de LDA est d'effectuer une réduction de dimensionnalité tout en préservant toutes les informations discriminatoires sur la classe et de trouver la direction dans laquelle les classes sont mieux séparées (SHUANG XU, 2012).

En d'autres termes, l'objectif de LDA est de maximiser la dispersion entre les classes S_B et de minimiser la matrice de dispersion intra-classe S_W dans le sous-espace projectif.

La matrice de dispersion intra-classe S_W et la matrice de dispersion inter-classes S_B sont définies comme suit (OH, YOO et PEDRYCZ, 2013) :

$$S_W = \sum_{j=1}^C \sum_{i=1}^{N_j} (\Gamma_i^j - \mu_j)(\Gamma_i^j - \mu_j)^T \quad (4.8)$$

Où Γ_i^j est le i échantillon de la classe j , μ_j est la moyenne de la classe j , C est le nombre de classes, N_j est le nombre d'échantillons en classe j .

$$S_B = \sum_{j=1}^C (\mu_j - \mu)(\mu_j - \mu)^T \quad (4.9)$$

μ représente la moyenne de toutes les classes. Le sous-espace de LDA est couvert par un ensemble de vecteurs $W = [W_1, W_2, \dots, W_d]$, satisfaisant :

$$W = \arg \max = \left| \frac{W^T S_B W}{W^T S_W W} \right| \quad (4.10)$$

W peut être construit en calculant les vecteurs propres de la matrice $S_w^{-1} \cdot S_B$

$$W = eig(S_w^{-1} \cdot S_B) \quad (4.11)$$

3. KFA : est une version centralisée de LDA (VINAY et al., 2015).

Elle fonctionne en projetant les données d'entrée dans un espace de caractéristiques implicites en utilisant du mappage non linéaire, $\phi : R^N \rightarrow R^f, f > n$ et ensuite, l'analyse linéaire de Fisher est adoptée pour cet espace de caractéristiques (VINAY et al., 2015).

Dans cette méthode nous avons utilisé le même algorithme pour exécuter KFA que dans (LIU et al., 2002).

Pour appliquer KFA dans un espace noyau, il faut trouver les valeurs propres λ et les vecteurs propres W^ϕ de :

$$\lambda S_W^\phi W^\phi = S_B^\phi W^\phi \quad (4.12)$$

Qui peut être obtenu par :

$$\begin{aligned} W_{OPT}^\phi &= \arg \max_{W^\phi} \left| \frac{(W^\phi)^T S_B^\phi W^\phi}{(W^\phi)^T S_W^\phi W^\phi} \right| \\ &= [W_1^\phi \dots W_m^\phi] \end{aligned} \quad (4.13)$$

où $\{W_i^\phi, i = 1, 2, \dots, m\}$ sont les vecteurs propres correspondant aux m valeurs propres les plus grandes $\lambda_i, i = 1, 2, \dots, m$.

S_B^ϕ est la matrice de dispersion entre classes et S_W^ϕ est la matrice de dispersion intra-classe dans R^f sont définis comme suit :

$$S_B^\phi = \sum_{i=1}^c l_i u_i^\phi (u_i^\phi)^T \quad (4.14)$$

$$S_W^\phi = \sum_{i=1}^c \sum_{j=1}^{l_i} \phi(X_{ij}) \phi(X_{ij})^T \quad (4.15)$$

où l_i est le nombre d'échantillons dans la classe i , c est le nombre de classes et u_i^ϕ est la moyenne d'une classe i dans R^f .

La fonction du noyau peut être définie comme suit (LIU et al., 2002) :

$$K(x_{tr}, y_{us}) = \phi(x_{tr}) \cdot \phi(y_{us}) \quad (4.16)$$

où x_{tr} est l'échantillon r de la classe t , y_{us} est l'échantillon s de la classe u .

Soit K une matrice de $m \times m$ et définie par les éléments $K_{tu}^{t=1, \dots, c}$, où K_{tu} est une matrice composée de produits scalaires dans l'espace des objets R^f :

$$K = K_{tu}^{t=1, \dots, c} \text{ où } K_{tu} = (k_{rs})_{s=1, \dots, l_u}^{r=1, \dots, l_t} \quad (4.17)$$

Remarque : K_{tu} est une matrice $l_t \times l_u$ et K une matrice symétrique $m \times m$.

Nous définissons également une matrice Z qui est une matrice diagonale $m \times m$:

$$Z = (Z_t)_{t=1,\dots,c} \quad (4.18)$$

D'après la théorie de la reproduction des noyaux, toute solution $W^\phi \in R^f$ doit se situer dans la plage de tous les échantillons d'entraînement de R^f , par exemple :

$$W^\phi = \sum_{p=1}^c \sum_{q=1}^{l_p} \alpha_{pq} \phi(X_{pq}) \quad (4.19)$$

Il en résulte que la solution pour Eq.5.19 peut être obtenue en résolvant :

$$\lambda K K \alpha = K Z K \alpha \quad (4.20)$$

Eq. 5.13 peut être écrit comme :

$$\begin{aligned} W_{OPT}^\phi &= \arg \max_{W^\phi} \left| \frac{(W^\phi)^T S_B^\phi \cdot W^\phi}{(W^\phi)^T S_W^\phi \cdot W^\phi} \right| \\ &= \arg \max \left| \frac{\alpha K Z K \alpha}{\alpha K K \alpha} \right| = [W_1^\phi \dots W_m^\phi] \end{aligned} \quad (4.21)$$

Nous utilisons la même technique que LDA, pour éviter la singularité dans le calcul W_{OPT}^ϕ , pour plus de détails voir (LIU et al., 2002) and (VINAY et al., 2015).

4.2.4 Calculer les scores

Chaque modèle extrait est associé aux modèles correspondants dans la base de données.

L'algorithme de classification K-NN est appliqué afin de classer ses caractéristiques et d'évaluer l'approche proposée.

4.2.5 Combiner les scores

Les scores peuvent être combinés en utilisant deux méthodes de combinaison de score, y compris la WS et PSO.

1. Règle de WS :

Il est basé sur cette équation :

$$Weighted_sum = \sum_{i=1}^N W_i S_i \quad (4.22)$$

Où : N est le nombre de scores à fusionner. S_i est le score correspondant au trait i et W_i est le poids pour chaque score ; il est calculé comme suit :

$$W_i = \frac{EER_i}{\sum_{i=1}^N EER_i} \quad (4.23)$$

Où EER_i est le taux d'erreur égales du système biométrique unimodal (OLA M. ALY et ONSI, 2013).

2. PSO : est une heuristique basée sur la population, où les solutions candidates sont appelées particules. Toutes les particules sont évaluées par une fonction de fitness à optimiser. Une propriété de vitesse est associée à chaque particule (MAZOUNI et RAHMOUN, 2011)(KALYAN, ANN et K., 2003).

Dans ce travail, PSO est utilisé pour combiner les scores fournis par les trois classificateurs.

L'algorithme commence par générer un groupe de particules aléatoires (poids), $W_i = (W_1, W_2, W_3)$, qui sont utilisés pour calculer le score fusionné (f) tel que :

$$f = W_1 \times S_1 + W_2 \times S_2 + W_3 \times S_3 \quad (4.24)$$

Où S_1, S_2, S_3 sont les scores normalisés des trois modalités impliqués (MAZOUNI et RAHMOUN, 2011), et que W est un facteur de pondération dans l'intervalle $[0, 1]$.

Pour chaque particule W_i , nous calculons la fonction de fitness $EER(f)$, puis nous sélectionnons les poids appropriés ou optimaux W_i pour lesquels $EER(f)$ est minimal.

4.3 Résultats expérimentaux et discussion

Tous les algorithmes ont été implémentés dans MATLAB (R2013a).

Dans cette méthode, nous construisons une base de données multibiométrique pour nos expériences basées sur la base de données CASIA-Iris-Interval_V3 Iris (*CASIA Database*), la base de données ORL (*ORL database*) et la base de données CASIA Palmprint (*CASIA Database*). La base de données multi-biométrique a été créée artificiellement en combinant ces trois bases de données.

La base de données virtuelles multi-biométrique se compose de huit images d'iris, de huit images de visage et de huit images palmprint par personne avec un total de 40 personnes.

La base de données CASIA Palmprint contient 7 200 images de paume capturées à partir de 100 personnes différentes à l'aide d'un appareil d'imagerie spectrale multiple auto-conçu. Toutes les images palmprint sont des fichiers JPEG de niveau gris 8 bits (*CASIA Database*).

CASIA-IrisV3-Interval contient 2655 images d'iris correspondant à 249 individus, avec une résolution de 320 x 280 pixels (*CASIA Database*).

Base de données visage ORL composée de 400 images en échelle de gris de taille 112×92 . Il y a 40 personnes, chaque individu a 10 images. Les images ont été prises à des moments différents, en variant l'éclairage, les expressions du visage (yeux ouverts / fermés, souriant / non souriant) et les détails du visage (lunettes / pas de lunettes) (*ORL database*).

Dans toutes les expériences de cette étude, quatre échantillons de chaque modalité sont sélectionnés au hasard pour l'entraînement et les autres échantillons sont utilisés pour les tests.

La performance a été mesurée en matière de EER et de taux de reconnaissance au rang 1 (Rank-1).

Dans ces expériences, les paramètres de PSO ont été déterminés comme suit :

- La taille de l'essaim est 20. Le nombre d'itérations est 50.
- On trouve expérimentalement que ω a été fixé à 0.4 pour obtenir de meilleures performances.
- Les constantes d'accélération c_1, c_2 sont définies sur 2.
- Limitation de la vitesse V_{max} est défini sur 2.

Les résultats expérimentaux obtenus pour les systèmes unimodaux et multimodaux utilisant différentes méthodes, en matière de taux de reconnaissance au rang 1 (Rank-1) et de EER, sont résumés dans les tableaux (5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9). Dans les systèmes multimodaux, deux combinaisons au niveau de score sont considérées, comprenant la WS et PSO.

TABLE 4.1 – Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant la transformation de Gabor et la PCA.

Système biométrique	Taux de reconnaissance au rang 1 %		EER %	
Iris unimodal	80.63		3.69	
Visage unimodal	88.75		0.62	
Palmprint unimodal	88.75		5.01	
Iris + visage	WS	PSO	WS	PSO
	90.00	97.50	0.67	0.00
Iris+Palmprint	97.50	98.13	0.51	0.13
Palmprint + Visage	91.88	97.50	3.76	0.008
Iris+visage+Palmprint	98.75	100.00	0.05	0.00

Comme le montrent les tableaux (5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9), une fusion des scores de l'iris, du visage et de palmprint conduit à une plus grande précision de reconnaissance comparé au système unimodal et système bimodaux. Comme indiqué dans ces tableaux, le meilleur résultat de la fusion des scores de l'iris, du visage et de palmprint est obtenu en appliquant une fonction PSO, une performance de 100% de reconnaissance au rang 1 et une valeur de 0 % de EER en utilisant le filtre de Gabor combiné avec PCA, LDA et KFA.

TABLE 4.2 – Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant la transformation de Gabor et la LDA.

Système biométrique	Taux de reconnaissance au rang 1 %		EER %	
Iris Unimodal	90.63		2.60	
Visage Unimodal	99.38		0.48	
Palmprint Unimodal	93.13		4.99	
Iris + Visage	WS	PSO	WS	PSO
	95.00	100.00	1.25	0.000
Iris+palmpoint	96.25	97.50	0.62	0.09
Palmprint + Visage	93.12	100.00	4.36	0.008
Iris+Visage+Palmprint	98.12	100	0.14	0.00

TABLE 4.3 – Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant la transformation de Gabor et KFA.

Système biométrique	Taux de reconnaissance au rang 1 %		EER %	
Iris Unimodal	92.50		2.65	
Visage Unimodal	98.75		0.62	
Palmprint Unimodal	93.13		5.00	
Iris + Visage	WS	PSO	WS	PSO
	95.63	100.00	0.62	0.00
Iris+palmpoint	95.63	99.38	0.62	0.04
Palmprint + Visage	93.75	100.00	3.75	0.00
Iris+Visage+Palmprint	97.50	100.00	0.11	0.00

Nous avons également comparé le système multimodal proposé avec des systèmes unimodaux utilisant ROC. Figures 5.4, 5.5 et 5.6 montre les courbes ROC des systèmes unimodaux et multimodaux pour toutes les expériences.

En analysant les résultats expérimentaux dans les figures 5.4, 5.5 et 5.6, on peut conclure que le système multimodal proposé pour la reconnaissance de l'iris, du visage et de palmprint en utilisant PSO obtient les meilleurs résultats par rapport à la règle de fusion à somme pondérée et aux systèmes unimodaux.

TABLE 4.4 – Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant DWT Transform et PCA.

Système biométrique	Taux de reconnaissance au rang 1%		EER %	
Iris Unimodal	62.50		11.25	
Visage Unimodal	82.50		3.75	
Palmprint Unimodal	90.62		5.63	
Iris + Visage	WS	PSO	WS	PSO
	81.25	92.50	4.36	1.88
Iris+palmprint	92.50	94.37	2.50	1.86
Palmprint + Visage	95.00	96.87	2.49	1.23
Iris+Visage+Palmprint	96.25	98.75	1.40	0.61

TABLE 4.5 – Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant DWT et LDA.

Système biométrique	Taux de reconnaissance au rang 1%		EER %	
Iris Unimodal	73.75		8.13	
Visage Unimodal	90.63		3.75	
Palmprint Unimodal	91.88		5.63	
Iris + Visage	WS	PSO	WS	PSO
	90.00	95.62	1.95	0.66
Iris+palmprint	94.37	94.37	1.27	0.93
Palmprint + Visage	94.37	98.12	1.25	0.71
Iris+Visage+Palmprint	97.50	98.75	0.62	0.15

4.4 Conclusion

Ce chapitre présente une fusion au niveau du score du système de reconnaissance multimodal de l'iris, du visage et de palmprint.

La transformation de Gabor, la transformation DWT et la transformation CT sont utilisées pour l'extraction de caractéristiques, puis les PCA, LDA et KFA sont utilisés pour réduire la dimension des vecteurs de caractéristiques. Ensuite, les vecteurs de caractéristiques sont classés à l'aide de K-NN. Enfin, les résultats des trois classificateurs sont fusionnés à l'aide de la règle de WS et de PSO.

À partir des résultats expérimentaux, on peut trouver que la fusion au niveau du score avec PSO est une méthode fiable et réalisable pour la fusion de l'iris, du visage et de palmprint.

L'objectif futur de ce travail est de mettre en œuvre le système biométrique multimodal en combinant d'autres traits biométriques tels que l'empreinte digitale, l'oreille, ...

Bien que l'approche proposée basée sur le PSO donnée des résultats acceptables, néanmoins les recherches dans ce domaine sont orienté vers l'apprentissage automatique. Particulièrement, l'apprentissage profond (deep learning).

TABLE 4.6 – Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant DWT et KFA.

Système biométrique	Taux de reconnaissance au rang 1 %		EER %	
Iris Unimodal	72.50		8.75	
Visage Unimodal	86.88		5.62	
Palmprint Unimodal	91.88		5.63	
Iris + Visage	WS	PSO	WS	PSO
	91.25	95.00	1.83	1.23
Iris+palmprint	92.50	95.00	2.00	1.35
Palmprint + Visage	96.25	97.50	1.78	1.24
Iris+Visage+Palmprint	98.12	98.75	0.62	0.27

TABLE 4.7 – Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant CT et PCA.

Système biométrique	Taux de reconnaissance au rang 1 %		EER %	
Iris Unimodal	56.88		12.51	
Visage Unimodal	78.75		3.75	
Palmprint Unimodal	90.63		5.00	
Iris + Visage	WS	PSO	WS	PSO
	76.25	90.62	5.72	2.49
Iris+palmprint	87.50	91.87	1.89	1.86
Palmprint + Visage	93.75	95.00	1.87	1.24
Iris+Visage+Palmprint	95.62	98.75	1.23	0.61

Le chapitre suivant présente la méthode proposée d'identification multimodale basée sur le deep learning.

TABLE 4.8 – Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant la CT et la LDA.

Système biométrique	Taux de reconnaissance au rang 1 %		EER %	
Iris Unimodal	68.12		9.44	
Visage Unimodal	88.75		3.76	
Palmprint Unimodal	91.25		5.51	
Iris + Visage	WS	PSO	WS	PSO
	89.37	96.25	3.74	0.80
Iris+palmpoint	90.62	94.37	1.99	1.49
Palmprint + Visage	94.37	97.50	1.25	0.61
Iris+Visage+Palmprint	98.75	99.37	0.62	0.072

TABLE 4.9 – Taux de reconnaissance au rang 1 et EER pour les systèmes unimodaux et multimodaux utilisant CT et KFA.

Système biométrique	Taux de reconnaissance au rang 1 %		EER %	
Iris Unimodal	66.88		9.99	
Visage Unimodal	86.25		5.07	
Palmprint Unimodal	90.63		5.00	
Iris + Visage	WS	PSO	WS	PSO
	88.75	95.00	2.45	0.92
Iris+palmpoint	89.37	93.75	2.50	1.86
Palmprint + Visage	95.62	96.87	1.25	1.23
Iris+Visage+Palmprint	98.75	98.75	0.62	0.14

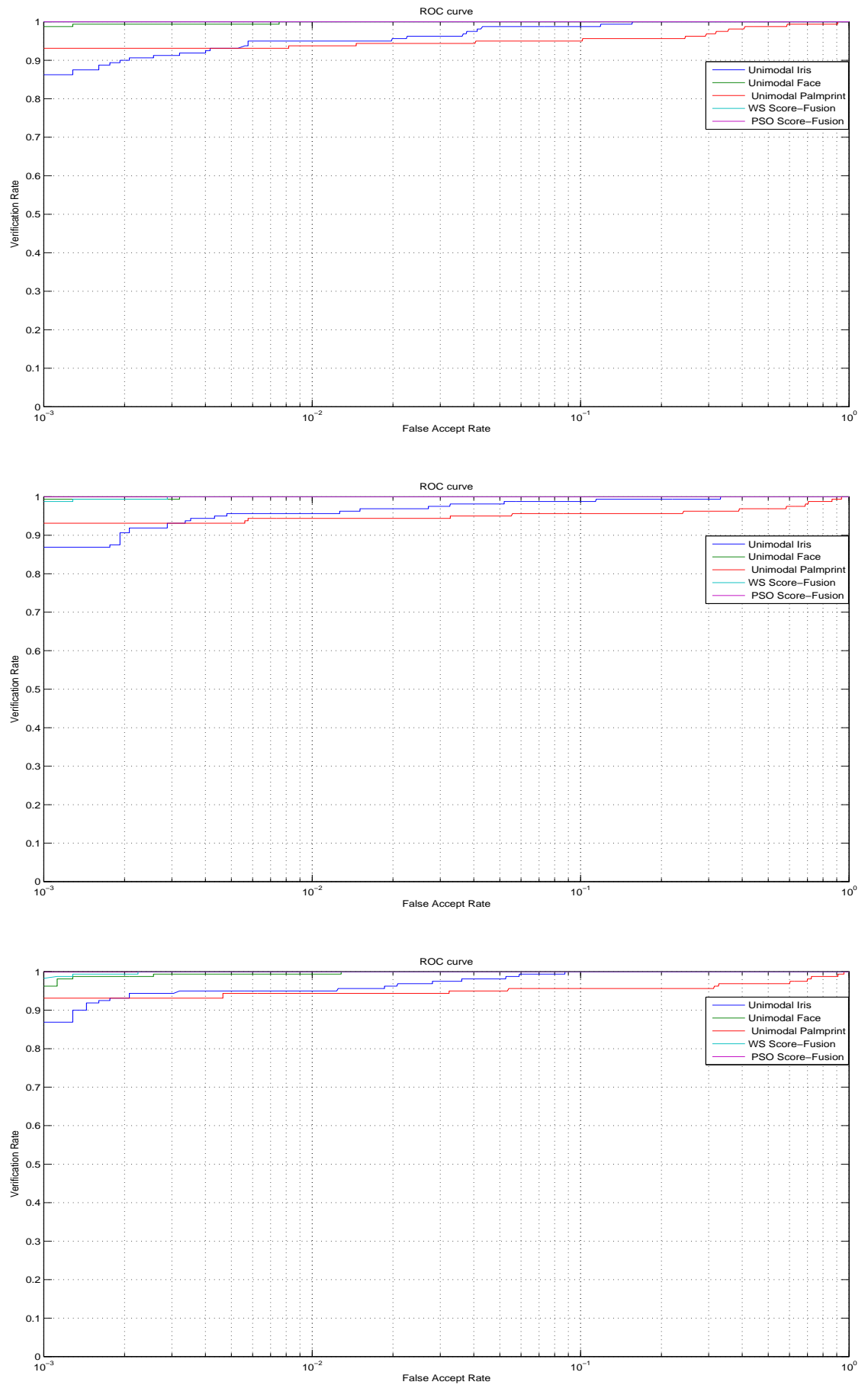


FIGURE 4.4 – Courbes ROC des méthodes unimodales et multimodales utilisant gabor et PCA, LDA et KFA.

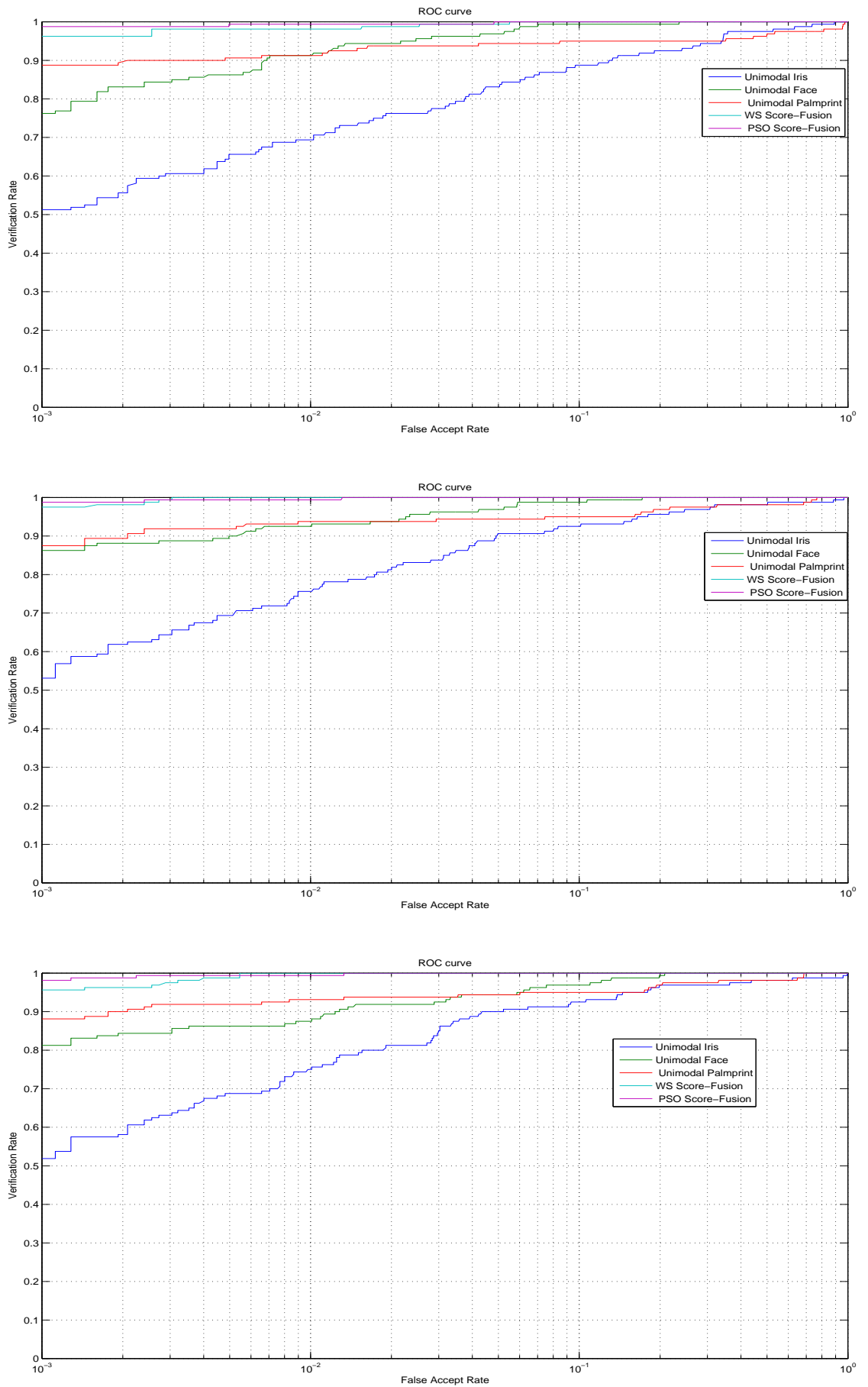


FIGURE 4.5 – Courbes ROC des méthodes unimodales et multimodales utilisant DWT et PCA, LDA et KFA.

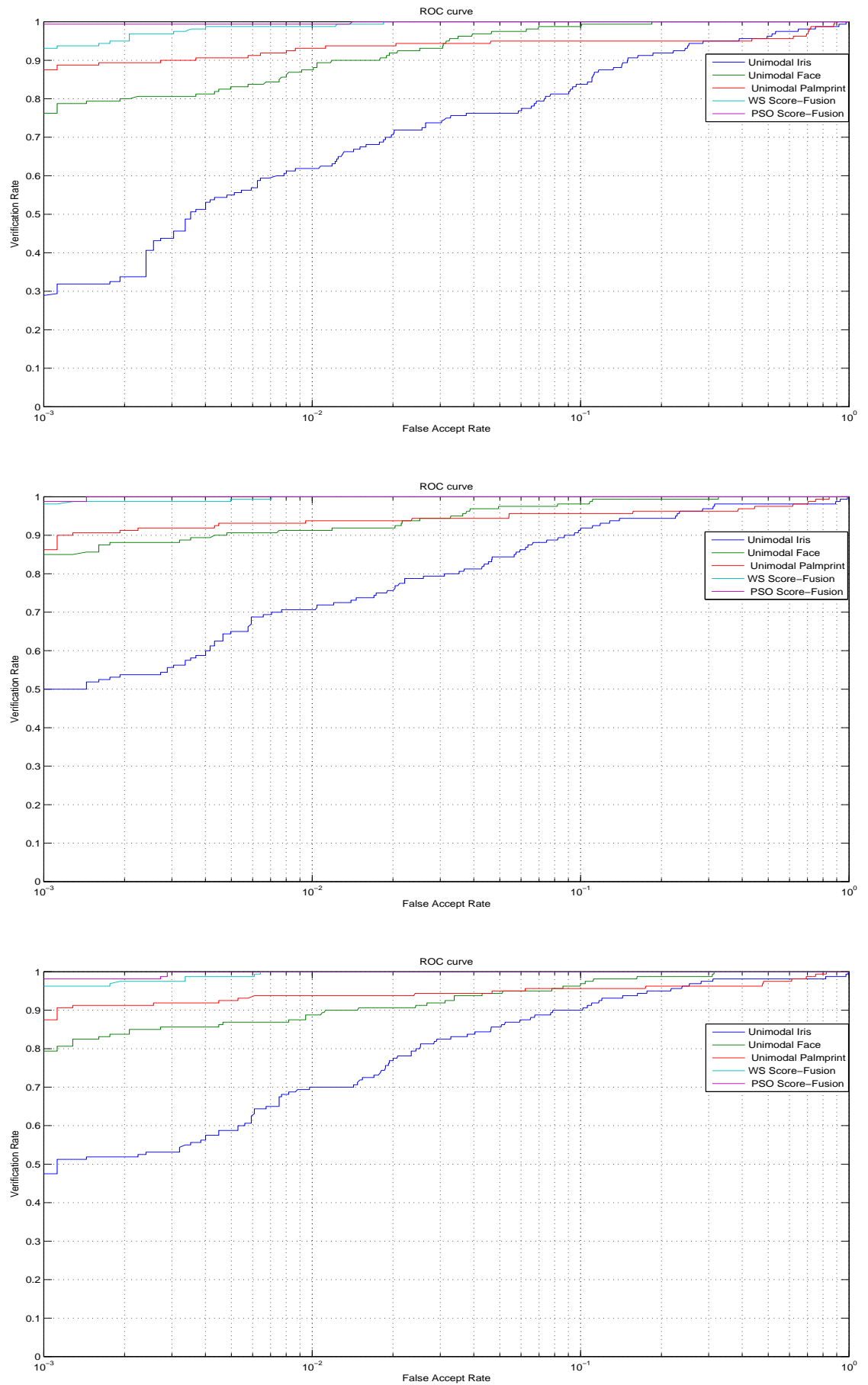


FIGURE 4.6 – Courbes ROC des méthodes unimodales et multimodales utilisant CT et PCA, LDA et KFA.

Chapitre 5

Nouvelle approche multi-biométrique basée sur le deep learning

5.1 Introduction

Dans ce chapitre nous allons présenter tout d'abord les notions en relation avec l'apprentissage profond (deep learning) tel que réseaux de neurones artificiels, les réseaux de neurones convolutifs.

Ensuite, nous proposons un système biométrique multimodal efficace basé sur le Deep Learning. Pour cela, nous explorons deux approches basées sur CNN(Convolutional Neural Network) en utilisant un modèle pré-entraînement téléchargé de la littérature (squeezenet).

Dans le premier approche, nous essayons de classer les objets en fonction de leurs caractéristiques distinctes et prenons donc une décision de classification finale basée sur les autres décisions. Nous pouvons utiliser un élément du modèle, tel que l'extracteur de caractéristiques, ou le fine-tuning en l'adaptant à notre tâche au lieu d'imagenet avec 1000 classes.

Dans la deuxième approche, les vecteurs de caractéristiques sont extraits séparément, en utilisant squeezenet pour extraire les caractéristiques profondes des modalités, nous prenons les caractéristiques de la couche 66 et nous combinons ces vecteurs pour former un vecteur de caractéristiques composites. Ensuite fait une classification à l'aide de SVM.

5.2 Etat de l'art

Les systèmes biométriques multimodaux ont fait l'objet de nombreuses recherches et plusieurs approches ont été proposées pour construire ces systèmes en combinant efficacement les données biométriques de plusieurs capteurs.

Un réseau profond avec une architecture de réseau neurone à plusieurs couches cachées est une étude prometteuse d'exploitation d'informations efficaces pour la collecte de données, ce qui inspire l'apprentissage profond multimodal.

Récemment, plusieurs travaux ont été proposés pour l'apprentissage profond multimodal.

Les années dernières, plusieurs articles ont traité de nouvelles approches biométriques utilisant les techniques d'apprentissage profond.

Omara et al. (OMARA IBRAHIM, 2017) ont proposé un système biométrique multimodal incluant le visage et les oreilles. Les caractéristiques profondes extraites pour les images de visage et d'oreille sont basées sur

VGG-M Net et DCA (Discriminant Correlation Analysis), ils sont utilisés pour la fusion et la réduction de la dimension de vecteur de fusion, puis SVM est utilisé pour la classification. Le système a atteint un taux de reconnaissance de 100%.

Kurban et al. (KURBAN, YILDIRIM et BILGIC, 2017) ont proposé un système biométrique multimodal qui intègre l'image de visage et de geste énergétique. Le modèle d'apprentissage profond de VGG a été utilisé comme extracteur de caractéristiques pour la base de données de visages, et une méthode d'imagerie énergétique a été utilisée pour extraire les caractéristiques gestuelles. Ensuite, PCA est utilisée pour effectuer une réduction dimensionnelle des vecteurs de caractéristiques.

Enfin, les scores similaires ont été produits avec la méthode de la distance euclidienne à déviation standard. Ces scores ont été fusionnés à l'aide de la règle de somme.

Shams et al. (SHAMS, SARHAN et TOLBA, 2017) ont présenté un système de récupération biométrique multimodal des empreintes digitales et des visages basés sur ADLVQ (Adaptive Deep Learning Vector Quantisation).

Le système proposé dans cet article extrait les caractéristiques des modalités d'entrée par LGPV (Local Gradient Pattern with Variance). Ensuite, l'algorithme K-means est appliqué à la quantification vectorielle.

Enfin, ses caractéristiques sont classées à l'aide d'un réseau de neurones profonds basés sur les connaissances acquises lors de la quantification vectorielle.

Al-Waisy et al. (AL-WAISY et al., 2017) ont proposé un système d'identification biométrique multimodal utilisant une architecture parallèle, afin de fusionner les résultats obtenus avec le visage et des iris gauches et droits. Premièrement, l'architecture DBN (Deep Belief Network) à 3 couches est utilisée pour extraire les traits du visage. Les deux premiers DBN servent de détecteurs de caractéristiques et le dernier est utilisé comme modèle discriminant associé aux unités softmax à des fins de la classification multi-classes.

Deuxièmement, pour la reconnaissance de l'iris, un système d'apprentissage profond efficace est utilisé. Il repose sur la combinaison d'un réseau neuronal convolutif et d'un classificateur softmax pour extraire les caractéristiques discriminantes de l'image de l'iris.

Dans (TALREJA, VALENTI et NASRABADI, 2017), les auteurs ont présenté un système multibiométrique sécurisé qui combine le visage et l'iris. Il utilise des CNNs pour l'extraction de caractéristiques.

Deux architectures de fusion, une architecture entièrement connectée et une architecture bilinéaire, sont mises en œuvre pour développer une représentation partagée multibiométrique robuste.

Ensuite, pour générer un modèle annulable multibiométrique, la dimension du vecteur de caractéristique fusionnée finale est réduite en appliquant un processus de sélection de caractéristique.

Dans (AL-WAISY et al., 2018), les auteurs ont présenté un système biométrique multimodal rapide basé sur les iris droits et gauches de la même personne utilisant l'une des méthodes de fusion au niveau des scores. Une approche d'apprentissage profond est une combinaison d'un CNN et d'un classificateur softmax pour l'extraction des caractéristiques et la classification.

Yang et al. (YANG et al., 2018) ont proposé un nouveau type de modèle de reconnaissance biométrique multimodale basé sur les méthodes empilées ELMs (stacked Extreme Learning Machines) et les méthodes CCA (Canonical Correlation Analysis).

Le modèle, qui a une structure symétrique, présente un potentiel élevé pour la biométrie multimodale. Le modèle fonctionne comme suit :

Tout d'abord, il apprend la représentation de la couche cachée d'images biologiques à l'aide de machines d'apprentissage extrêmes, couche par couche.

Deuxièmement, la méthode CCA est appliquée pour mapper la représentation sur un espace de caractéristiques, qui est utilisé pour reconstruire la représentation de caractéristiques d'image multimodale.

Troisièmement, les caractéristiques reconstruites sont utilisées comme entrées d'un classificateur pour l'entraînement supervisées et la sortie.

Pour vérifier la validité et l'efficacité de la méthode, les auteurs adoptent nouveaux jeux de données hybrides obtenus à partir de jeux de données d'image de face et d'images de veine de doigt.

Kim et al. (KIM, MIN SONG et RYOUNG PARK, 2018) ont présenté un système biométrie multimodale qui combine veine de doigt et forme de doigt utilisant un capteur de caméra infrarouge proche sur la base d'un réseau de neurones à convolution profond. Les distances correspondantes calculées en fonction des caractéristiques de la veine et de la forme des doigts obtenues à l'aide de modèles ResNet ont été fusionnées à l'aide de diverses méthodes de fusion, telles que la somme du poids, le produit pondéré et le perceptron.

Soleymani et al. (SOLEYMANI et al., 2018) ont proposé une architecture CNN commune avec fusion au niveau de caractéristiques pour la reconnaissance multimodale utilisant de multiples modalités de visage, d'iris et d'empreintes digitales.

Plutôt que de fusionner les réseaux au niveau de la couche softmax, les caractéristiques de compression optimale de toutes les modalités sont fusionnées aux couches entièrement connectées sans perte de précision des performances, mais avec une réduction significative du nombre de paramètres réseau.

5.3 Réseaux de neurones artificiels

Dans le corps humain, les réseaux de neurones sont les éléments constitutifs du système nerveux qui contrôle et coordonne les différentes activités

humaines. Chaque neurone ou cellule nerveuse est constitué d'un corps de cellule appelée Cyton et d'une fibre appelée Axon.

Les neurones sont interconnectés par les structures fibreuses appelées dendrites à l'aide de connexions spéciales appelées synapses. Les impulsions électriques (appelées potentiels d'action) sont utilisées pour transmettre des informations d'un neurone à l'autre à travers le réseau.

Les réseaux de neurones artificiels ont été développés sur le principe de fonctionnement similaire des réseaux de neurones humains.

Le bloc constitutif de base de chaque réseau neuronal artificiel est un neurone artificiel, c'est-à-dire un modèle mathématique simple (fonction). Un tel modèle comporte trois ensembles simples de règles : multiplication, sommation et activation. À l'entrée du neurone artificiel, les entrées sont pondérées, ce qui signifie que chaque valeur d'entrée est multipliée par le poids individuel (KRENKER, BESTER et KOS, 2011).

Dans la partie centrale du neurone artificiel, la fonction somme additionne toutes les entrées pondérées et les biais. À la sortie du neurone artificiel, la somme des entrées pondérée précédemment et du biais passe par la fonction d'activation, appelée également fonction de transfert (figure 6.1).

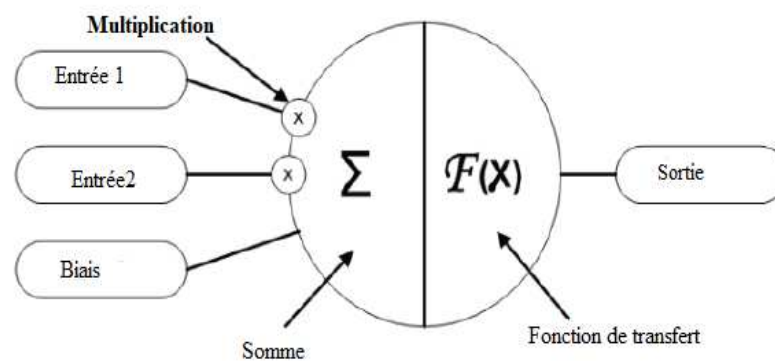


FIGURE 5.1 – Principe de fonctionnement d'un neurone artificiel.

Bien que les principes de fonctionnement et le jeu de règles simples des neurones artificiels ne ressemblent à rien de spécial, le potentiel et la puissance de calcul de ces modèles prennent vie lorsque nous commençons à les interconnecter dans des réseaux de neurones artificiels.

Un neurone artificiel est une fonction f_j de l'entrée $x = (x_1, \dots, x_d)$ pondérée par un vecteur de poids de connexion $w_j = (w_{j,1}, \dots, w_{j,d})$, complété par un biais neuronal b_j et associé à une fonction d'activation ϕ , à savoir (KRENKER, BESTER et KOS, 2011)

$$y_i = f_j(x) = \phi((w_j, x) + b_j)$$

(5.1)

Plusieurs fonctions d'activation peuvent être envisagées. Voici quelques fonctions d'activation que vous trouverez souvent dans la pratique (ALTENBERGER et LENZ, 2018) :

1. La fonction linéaire ou identité : cette fonction est notée par :

$$\phi(x) = x \quad (5.2)$$

Les réseaux de neurones à couche unique utilisent une fonction pas à pas lors de la conversion d'une fonction d'entrée à variation continue en une sortie binaire (0 ou 1) ou en une sortie bipolaire (1 ou -1).

2. Fonction binaire : cette fonction utilise un seuil. Une fonction binaire avec un seuil T est donnée par :

$$f(x) = \begin{cases} 1 & \text{si } x \geq T \\ 0 & \text{sinon} \end{cases}$$

3. La fonction sigmoïde (ou logistique) : c'est une fonction d'activation de la forme :

$$\phi(x) = \frac{1}{1 + \exp(-x)} \quad (5.3)$$

Sa plage est comprise entre 0 et 1. C'est une courbe en forme de S. Elle est facile à utiliser et possède toutes les propriétés intéressantes des fonctions d'activation : non linéaire, continuellement différentiable, monotone et plage de sortie fixe.

4. La fonction tangente hyperbolique (tanh) : sa formule mathématique est :

$$\phi(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}$$

(5.4)

Tanh est aussi comme sigmoïde logistique mais mieux. La plage de la fonction tanh est comprise entre (-1 et 1).

La fonction tanh sert principalement à la classification entre deux classes.

5. La fonction d'activation ReLU (Rectified Linear Unit) : la ReLU est la fonction d'activation la plus utilisée. Elle est définie comme :

$$\phi(x) = \max(0, x) \quad (5.5)$$

La fonction ReLU n'est pas linéaire, ce qui signifie que nous pouvons facilement propager les erreurs en arrière et que plusieurs couches de neurones sont activées par la fonction ReLU.

Il existe de nombreuses classes de réseaux de neurones et ces classes ont également des sous-classes. Ici, je vais énumérer les plus utilisés :

5.3.1 Réseau neuronal à propagation avant ou feed-forward

Le réseau neuronal à propagation avant était le premier et le plus simple type de réseau neuronal artificiel conçu. Dans ce réseau, les informations ne sont acheminées que dans un sens, des nœuds d'entrée aux nœuds de sortie, en passant par les couches masquées, et jusqu'aux nœuds de sortie. Il n'y a pas de cycles ou de boucles dans le réseau. Les exemples les plus connus sont le perceptron simple et sa version multi-couche (FINE, 1999).

(a) Le perceptron simple :

Le perceptron simple est un réseau monocouche, acyclique (il n'a pas de boucle) et dont la dynamique (activité) est déclenchée par la réception des informations saisies (ALTENBERGER et LENZ, 2018).

Ce réseau est dit simple car il ne contient aucune couche cachée, ce qui signifie qu'il ne se compose qu'une couche d'entrée et une couche de sortie.

Ces structures lui permettent d'être considéré comme un classifieur linéaire : autrement dit, il peut classer les données en fonction de deux caractéristiques.

(b) Le perceptron multicouche ou Multi-layer Perceptron(MLP) :

Le perceptron multicouche se compose d'une couche d'entrée, d'une couche de sortie et d'une ou plusieurs couches cachées (voir figure 6.2).

Chaque neurone d'une couche possède des connexions dirigées vers les neurones de la couche suivante.

Dans de nombreuses applications, les unités de ces réseaux appliquent une fonction sigmoïde en tant que fonction d'activation.

Les perceptrons multicouches sont beaucoup plus utiles et une bonne raison est qu'ils sont capables d'apprendre des représentations non linéaires (la plupart des cas, les données qui nous sont présentées ne sont pas linéairement séparables) (TOUZET, 1992).

MLP est largement utilisé pour résoudre des problèmes nécessitant un apprentissage supervisé ainsi que des recherches sur les neurosciences computationnelles et le traitement distribué parallèle. Les applications incluent la reconnaissance vocale, la reconnaissance d'images et la traduction automatique (KRENKER, BESTER et KOS, 2011).

5.3.2 Réseau de neurones récurrent (feedback network) :

Dans le réseau de neurones récurrent, les connexions entre les unités forment un cycle dirigé (elles propagent les données en avant, mais aussi en arrière, des étapes de traitement ultérieures aux étapes précédentes) (TOUZET, 1992).

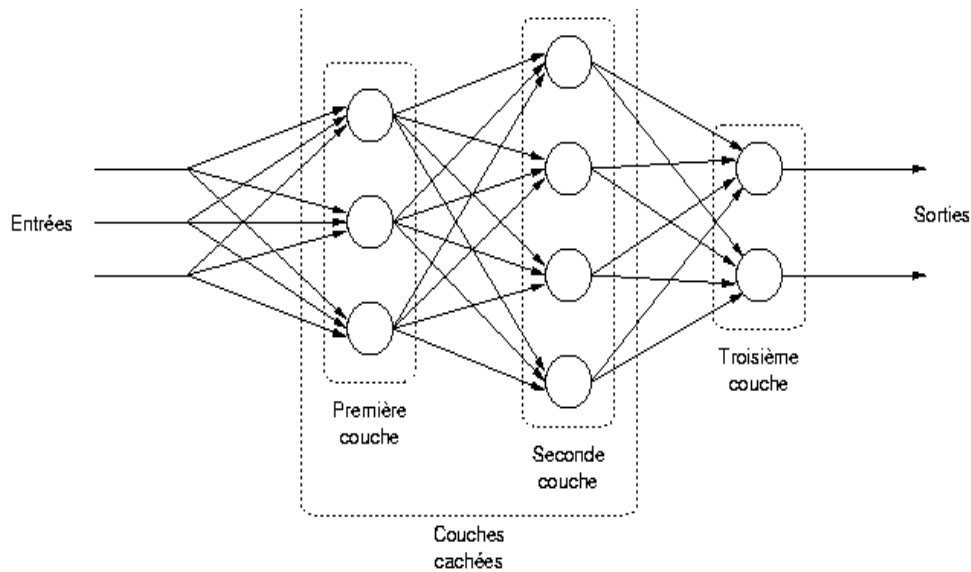


FIGURE 5.2 – Exemple d’un réseau de type perceptron multi-couche.

Cela lui permet de présenter un comportement temporel dynamique. Contrairement aux réseaux neuronaux à propagation avant, les réseaux de neurones récurrents peuvent utiliser leur mémoire interne pour traiter des séquences d’entrées arbitraires.

D’autres réseaux de neurones artificiels récurrents tels que Hopfield, Elman, Jordan, les réseaux bidirectionnels et autres ne sont que des cas particuliers de réseaux de neurones artificiels récurrents (KRENKER, BESTER et KOS, 2011).

5.4 Apprentissage profond (Deep Learning)

5.4.1 Définition de l’apprentissage profond

L’intelligence artificielle est une discipline qui recherche des méthodes de résolution de problèmes très complexes en logique ou en algorithmes.

L’apprentissage automatique est un domaine d’étude de l’intelligence artificielle qui s’appuie sur des approches statistiques pour donner aux ordinateurs la possibilité d’apprendre à partir de données.

Par conséquent, l’apprentissage profond, (en anglais deep learning, deep structured learning, hierarchical learning) est un ensemble de méthodes d’apprentissage automatiques qui tentent de modéliser avec un niveau élevé d’abstraction des données à travers des architectures articulées de différentes transformations non linéaire (SCHMIDHUBER, 2015).

En apprentissage automatique, un programme analyse un ensemble de données afin d’en déduire des règles permettant de tirer des conclusions sur les nouvelles données.

Grâce à un apprentissage profond, les images brutes sont directement introduites dans un réseau de neurones profonds qui apprend automatiquement les fonctionnalités (ANTIPOV, 2017).

L'apprentissage profond est basé sur les réseaux de neurones artificiels, composés de milliers des neurones qui effectuent chacun de petites opérations simples. Les résultats d'une première couche de neurones servent d'entrée aux calculs d'une deuxième couche et ainsi de suite.

Le terme profond se rapporte généralement au nombre de couches cachées du réseau de neurones. Les réseaux de neurones classiques ne comportent que 2 à 3 couches cachées, tandis que les réseaux profonds peuvent en compter jusqu'à 150.

Les progrès de l'apprentissage profond ont été possibles grâce à l'augmentation de la puissance des ordinateurs et au développement de grandes bases de données.

Le réseau de neurones à convolution est l'un des types de réseaux de neurones profonds les plus courants (LECUN, BENGIO et HINTON, 2015). Un CNN convole les caractéristiques apprises avec les données d'entrée et utilise des couches convolutives 2D. Cette architecture est donc parfaitement adaptée au traitement de données 2D telles que des images.

5.4.2 Domaines d'application de l'apprentissage profond :

Dans cette section, nous passons en revue les principales applications de l'apprentissage profond. L'apprentissage profond a connu une croissance constante au cours des dernières années. En particulier, les principales avancées ont été l'application de l'apprentissage profond en analyse multimédia, y compris traitement de l'image, de l'audio et du langage naturel, qui a permis des avancées significatives dans l'état de la technique des systèmes autonomes.

Parmi les applications de l'apprentissage profond en peut citer (HATCHER et YU, 2018) :

- (a) **Reconnaissance/ classement d'images et de vidéos** : en général, le plus grand domaine d'apprentissage profond, le traitement, la reconnaissance et la détection d'images et de vidéos ont connu une croissance explosive ces dernières années. Dans le traitement des images et de la vidéo, les structures profondes typiques sont les réseaux de neurones convolutifs, qui appellent d'abord plusieurs canaux d'images, puis regroupent les couches d'images, couche par couche réduisant la taille de l'image ou du champ de trame, avant de transmettre les résultats aux couches entièrement connectées.

Le traitement d'images et de vidéos a été appliqué à de nombreux domaines d'étude, notamment les systèmes autonomes, l'imagerie médicale, l'astrophysique, l'analyse biométrique, etc.

- (b) **Traitement audio** : le traitement du signal audio concerne généralement la réduction du bruit et la compression des données qui maintiennent la valeur maximale pour les auditeurs humains, ce qui est très

pertinent pour le domaine de l'audiologie. Les autos-encodeurs profonds sont très prometteur dans ce domaine, et la capacité à distinguer les voix, les langues et le bruit de fond d'une entrée de microphone multiple et singulier sont un avantage considérable que l'apprentissage en profondeur peut potentiellement atteindre.

En outre, le traitement du langage naturel concerne la détection, la compréhension et la traduction de la langue parlée.

- (c) **Systèmes autonomes et robotique** : le domaine de la robotique a connu des progrès incroyables dans la capacité de créer des robots de marche autonomes, ainsi que dans les domaines du vol, de la conduite et de la navigation, entre autres.

Dans le domaine de la manipulation robotique, l'apprentissage profond a permis des avancées significatives dans la formation rapide de bras robotiques pour des tâches de fabrication répétitives de diverses manières.

- (d) **Diagnostic médical** : fortement influencés par les progrès de l'analyse d'images, les diagnostics médicaux ont grandement bénéficié des améliorations rapides de l'apprentissage profond.

Des travaux considérables ont été réalisés pour améliorer la détection des maladies, des tumeurs et d'autres anomalies à partir d'images IRM, de CT-scanner (Computer Tomography, tomographie), etc.

- (e) **Biologie computationnelle** : bien que similaires aux applications pour le diagnostic médical, nous considérons néanmoins que l'apprentissage en profondeur appliqué aux sciences biologiques constitue une catégorie tout à fait différente.

Plus précisément, ces applications recourent le domaine de la chimie et des interactions moléculaires et étudient les processus inhérents à divers micros et macro-organismes.

- (f) **Sciences physiques** : l'apprentissage profond offre un potentiel considérable pour améliorer les sciences physiques via la réduction de la dimensionnalité et la capacité à réaliser des analyses détaillées avec des données volumineuses de manière généralisée.

En effet, les sciences astrophysiques, géologiques, environnementales et de la mécanique quantique en ont grandement bénéficié, pour n'en nommer que quelques-unes.

- (g) **Finance, économie, analyse de marché et autres** : en tant que mécanismes de prévision et d'analyse, les outils d'apprentissage profond ont la capacité d'apprendre à partir de données stochastiques et de reconnaître les tendances, de sorte que les systèmes basés sur l'apprentissage automatique ont été largement développés pour la prévision du marché.

En outre, la vérification et la validation des transactions monétaires bénéficient grandement des données potentielles générées par les utilisateurs et peuvent être utilisées pour détecter un comportement anormal.

5.4.3 Réseau de neurones convolutifs (CNN ou ConvNet) :

Un réseau de neurones convolutifs est un type de réseau de neurones artificiels acycliques (feed-forward), qui se distingue des réseaux de neurones traditionnels entièrement connectés en combinant un certain nombre de couches connectées localement destinées à la reconnaissance automatique des caractéristiques, suivies d'un nombre de couches entièrement connectées destinées à la classification (HATCHER et YU, 2018).

Bien qu'elles soient efficaces pour le traitement des images, les MLP ont du mal à gérer les grandes images en raison de la croissance exponentielle du nombre de connexions avec la taille de l'image, car chaque neurone est "entièrement connecté" aux neurones de la couche précédente et de la suivante.

Les réseaux de neurones convolutifs, dont le principe s'inspire de celui du cortex visuel des vertébrés, limitent le nombre de connexions entre un neurone et les neurones des couches adjacentes, ce qui diminue considérablement le nombre de paramètres à apprendre (GOODFELLOW, BENGIO et COURVILLE, 2016).

CNN est très efficace et robuste pour gérer les déformations d'image et autres transformations géométriques, telles que la translation, la rotation et la mise à l'échelle (OMARA IBRAHIM, 2017).

Les réseaux de neurones convolutifs ont de nombreuses applications en reconnaissance d'image et vidéo, systèmes de recommandation et traitement du langage naturel.

L'architecture CNN comprend plusieurs couches distinctes, notamment des ensembles de couches convolutives connectées localement (avec un nombre spécifique des différents noyaux apprenables dans chaque couche), des couches de sous-échantillonnage appelées couches de pooling et une ou plusieurs couches entièrement connectées.

Plus en détail, ces couches sont :

Couche convolutive : la convolution est la première couche permettant d'extraire des entités d'une image d'entrée. La convolution préserve la relation entre les pixels en apprenant les caractéristiques de l'image à l'aide de petits carrés de données d'entrée. C'est une opération mathématique qui prend deux entrées comme une matrice d'image et un filtre ou kernel (ALTENBERGER et LENZ, 2018).

Trois hyperparamètres permettent de dimensionner le volume de la couche de convolution : profondeur, pas et marge (O'SHEA et NASH, 2015).

- (a) Profondeur de la couche : nombre de noyaux de convolution (ou nombre de neurones associés au même champ récepteur).
- (b) Le pas : contrôle le chevauchement des champs récepteurs. Plus le pas est petit, plus les champs récepteurs se chevauchent et plus le volume de sortie est élevé.
- (c) La marge (à 0) ou zéro remplissage : est le processus simple de remplissage de la bordure de l'entrée, elle constitue une méthode efficace pour mieux contrôler la dimensionnalité des volumes de sortie.

Il est important de comprendre qu'en utilisant ces techniques, nous allons modifier la dimensionnalité spatiale de la sortie des couches convolutives. Pour calculer cela, vous pouvez utiliser la formule suivante :

$$\frac{(V - R) + 2Z}{s + 1} \quad (5.6)$$

Où V représente la taille du volume d'entrée, R représente la taille du champ de réception, Z la quantité de remplissage nul et S correspondant au pas. Si le résultat calculé à partir de cette équation n'est pas un entier, le pas n'a pas été correctement définie, car les neurones ne pourront pas s'adapter parfaitement à l'entrée donnée.

Couche de pooling (POOL) :

Un autre concept important des CNN est le pooling, qui est une forme de sous-échantillonnage de l'image. L'image d'entrée est divisée en une série de rectangles de n pixels latéraux qui ne se chevauchent pas (pooling)(GOODFELLOW, BENGIO et COURVILLE, 2016).

Chaque rectangle peut être vu comme une tuile. Le signal de sortie de la mosaïque est défini en fonction des valeurs prises par les différents pixels de la mosaïque.

Le regroupement réduit la taille d'une image intermédiaire, réduisant ainsi le nombre de paramètres et les calculs dans le réseau. Il est donc courant d'insérer périodiquement une couche de pooling entre deux couches convolutives successives d'une architecture de réseau neuronal convolutif afin de réduire le surapprentissage.

La forme la plus courante est une couche de regroupement avec des mosaïques de taille 2×2 (largeur / hauteur) et, en tant que valeur de sortie, la valeur d'entrée maximale. Dans ce cas, on parle de "Max-Pool 2×2 " (compression d'un facteur 4) (O'SHEA et NASH, 2015).

Il est possible d'utiliser d'autres fonctions de regroupement que le maximum. On peut utiliser un "pooling moyen" (la sortie est la moyenne des valeurs de patch d'entrée), "pooling standard L2".

En fait, même si au départ la mise en commun moyenne était souvent utilisée, il s'est avéré que le max-pooling était plus efficace car il augmentait de manière plus significative l'importance des activations fortes. Dans d'autres circonstances, le pooling stochastique peut être utilisée (ANTIPOV, 2017).

Couche entièrement connectée : la sortie de la dernière couche de convolution ou de pooling est acheminée vers une ou plusieurs couches entièrement connectées, comme dans un réseau de neurones traditionnel.

Dans ces couches, les sorties de tous les neurones de la couche $(l - 1)$ sont entièrement connectées à chaque neurone de la couche l . La sortie $y^{(l)}(j)$ du neurone j dans une couche entièrement connectée l est définie comme suit :

$$y^{(l)}(j) = f^{(l)} \left(\sum_{i=1}^{N^{l-1}} y^{(l-1)}(i) \cdot w^{(l)}(i, j) + b^{(l)} - j \right) \quad (5.7)$$

où $N^{(l-1)}$ est le nombre de neurones de la couche précédente ($l - 1$), $w^{(l)}(i, j)$ est le poids pour la connexion de neurone i dans la couche ($l - 1$) vers le neurone j dans la couche l , et $b^{(l)}$ est le biais du neurone j dans la couche l . $f^{(l)}$ représente la fonction d'activation de la couche l .

Architecture du réseau de neurones convolutifs :

Plusieurs architectures dans le domaine des réseaux de neurones convolutifs. Les plus courants sont :

LeNet : LeNet a été l'un des tout premiers réseaux de neurones convolutifs à propulser le champ de l'apprentissage profond. Ce travail pionnier de Yann LeCun a été nommé le Net5 après de nombreuses itérations réussies depuis l'année 1988 (LECUN et al., 1998). À cette époque, l'architecture LeNet était principalement utilisée pour les tâches de reconnaissance de caractères telles que la lecture de codes postaux, de chiffres, etc.

L'architecture LeNet5 était fondamentale, notamment en ce qui concernait le fait que les caractéristiques de l'image étaient réparties sur toute l'image. Les convolutions avec des paramètres pouvant être apprises étaient un moyen efficace d'extraire des caractéristiques similaires à plusieurs endroits avec peu de paramètres.

À l'époque, il n'y avait pas de GPU (Graphics Processing Unit) pour aider à l'entraînement, et même les processeurs étaient lents. Par conséquent, être capable de sauvegarder les paramètres et les calculs était un avantage clé.

Cela contraste avec l'utilisation de chaque pixel en tant qu'entrée distincte d'un grand réseau de neurones multicouches.

LeNet5 a expliqué que ceux-ci ne devraient pas être utilisés dans la première couche, car les images sont fortement corrélées spatialement et que l'utilisation de pixels individuels de l'image en tant qu'entités d'entrée distinctes ne tireraient pas parti de ces corrélations.

L'architecture du réseau est relativement simple, car elle se compose uniquement d'une couche d'entrée de taille 32×32 , d'une couche de sortie de taille 10, ainsi que de trois systèmes de convolution 5×5 , deux couches de pooling 2×2 et deux couches entièrement connectées entre les deux, ce qui en fait un total de six couches cachées (ALTENBERGER et LENZ, 2018).

Dans l'ensemble, ce réseau a été à l'origine de la plupart des architectures récentes et constitue une véritable source d'inspiration pour de nombreuses personnes sur le terrain.

AlexNet : en 2012, Alex Krizhevsky a publié AlexNet, une version plus profonde et beaucoup plus large du LeNet (KRIZHEVSKY, SUTSKEVER et E. HINTON, 2012).

AlexNet n'est pas très différent du LeNet, car il se compose également uniquement de la couche d'entrée, de quelques couches convolutives avec un pooling occasionnel, ainsi que de quelques couches entièrement connectées juste avant la couche de sortie (ALTENBERGER et LENZ, 2018).

Cependant, AlexNet a plus de couches et de neurones par couches et utilise également différents hyperparamètres.

VGGNet : le réseau VGG, introduit en 2014, offre une variante plus profonde mais plus simple des structures de convolution. Au moment de son introduction, ce modèle était considéré comme très profond (ALTENBERGER et LENZ, 2018).

Le VGGNet était très simple. Il ne s'agissait que de couches convolutives avec un champ récepteur 3×3 , ce qui représente la plus petite taille permettant de différencier les directions de base, ainsi que de 2×2 couches de pooling maximum, et de trois couches entièrement connectées à la fin.

GoogLeNet : en 2014, les chercheurs de Google ont présenté le réseau GoogLeNet, qui a pris la première place du concours ImageNet en 2014 pour les défis de classification et de détection.

Le modèle comprend une unité de base appelée "cellule de création" dans laquelle nous effectuons une série de convolutions à différentes échelles puis agrégées les résultats. Afin de sauvegarder le calcul, des convolutions de 1×1 sont utilisées pour réduire la profondeur du canal d'entrée. Pour chaque cellule, nous apprenons un ensemble de filtres 1×1 , 3×3 et 5×5 qui peuvent apprendre à extraire des entités à différentes échelles à partir de l'entrée. Le max-pooling est également utilisé, bien qu'avec un même remplissage pour préserver les dimensions afin que la sortie puisse être correctement concaténée.

Cette architecture est composée de 22 couches en profondeur. Il réduit le nombre de paramètres de 60 millions (AlexNet) à 4 millions (SZEGEDY et al., 2015).

ResNet : les réseaux résiduels profonds (ResNet) étaient une idée révolutionnaire qui permettait de développer des réseaux beaucoup plus profonds (des centaines de couches par opposition à des dizaines de couches) (ALTENBERGER et LENZ, 2018).

C'est un principe généralement accepté que les réseaux plus profonds sont capables d'apprendre des fonctions plus complexes et des représentations de l'intrant qui devraient conduire à de meilleures performances. Cependant, de nombreux chercheurs ont observé que l'ajout de couches supplémentaires finissait par avoir un effet négatif sur les performances finales.

Dans ResNets, les couches de convolution sont divisées en blocs résiduels et, pour chaque bloc, une connexion résiduelle est ajoutée, ce qui contourne le bloc correspondant. Ensuite, la sortie du bloc résiduel est fusionnée par sommation avec l'entrée d'origine transmise par la connexion résiduelle.

En ajoutant ces connexions résiduelles, le résultat d'une étape d'apprentissage peut être rétropropagé directement aux couches précédentes, sans aucune interférence des couches suivantes.

SqueezeNet : est une petite architecture de réseau de neurones à convolution profonde qui permet de réduire la taille du modèle de 50% par rapport à AlexNet (KRIZHEVSKY, SUTSKEVER et E. HINTON, 2012). Ils

ont également, avec les techniques de compression de modèle, compressé SqueezeNet à moins de 0,5 Mo, ou 510× plus petits que AlexNet (IANDOLA et al., 2016).

Dans l'article SqueezeNet (IANDOLA et al., 2016), l'auteur décrit trois stratégies principales :

- (a) Remplacer 3×3 noyaux par 1×1 pour les réseaux plus petits : pour réduire les paramètres du réseau, ils ont utilisé des filtres 1×1 au lieu de 3×3 . Etant donné que les noyaux 1×1 ont 9× moins de paramètres que les noyaux 3×3 .
- (b) Réduire le nombre de canaux d'entrée à 3×3 noyaux : utiliser les noyaux 1×1 comme couche de goulot d'étranglement appelée couche squeeze, afin de réduire la profondeur du modèle et le calcul des noyaux 3×3 suivants.
- (c) Un sous-échantillonnage plus tard dans le réseau pour conserver les grandes cartes d'activation afin de préserver les cartes de fonctions : l'intuition est que la grande carte d'activation est conservée en raison d'un sous-échantillonnage retardé dans le réseau, ce qui entraîne une précision de classification supérieure.

Cette architecture est composée de 68 couches en profondeur.

5.4.4 Créer et entraîner des modèles de Deep Learning

Voici les trois méthodes les plus populaires d'utilisation de Deep Learning :

L'entraînement à partir de zéro

Pour conduire un réseau profond à partir de zéro, vous devez collecter une très grande quantité de données étiquetées et concevoir une architecture réseau qui apprendra les fonctionnalités et le modèle. C'est une méthode adaptée aux nouvelles applications ou à celles qui obtiendront de nombreuses catégories de résultats.

Cette approche n'est pas très répandue car, en raison du volume important de données et du rythme d'apprentissage, l'entraînement des réseaux peut facilement être étalé sur plusieurs jours, voire plusieurs semaines.

Apprentissage par transfert (Transfer Learning)

La plupart des applications d'apprentissage profond utilisent l'apprentissage par transfert, consiste à mettre au point un modèle préentraîné. Le processus commence par un réseau existant, tel qu'AlexNet ou GoogLeNet, qui doit être enrichi de nouvelles données contenant des classes auparavant inconnues du réseau (PAN et YANG, 2010).

Une fois que vous avez effectué certains ajustements sur le réseau, vous pouvez effectuer une nouvelle tâche.

L'apprentissage par transfert a pour objectif d'améliorer l'apprentissage dans la tâche cible en exploitant les connaissances issues de la tâche source.

L'apprentissage par transfert nous permet de :

- Transférer les caractéristiques apprises d'un réseau pré-entraîné à un nouveau problème.
- Être plus rapide et plus facile que l'entraînement d'un nouveau réseau.
- Réduire le temps de l'entraînement et la taille de l'ensemble de données.
- Effectuer un apprentissage profond sans avoir besoin d'apprendre à créer un nouveau réseau.

L'apprentissage par transfert nécessite une interface pour accéder au contenu interne du réseau préexistant, procéder à des ajustements très précis et le perfectionner pour sa nouvelle tâche.

Les dernières couches ont tendance à être spécialisées sur les données que vous avez introduites dans le modèle (YOSINSKI et al., 2014).

D'autre part, les premières couches sont beaucoup plus génériques, il existe de nombreux motifs simples communs à une classe d'images beaucoup plus large.

L'apprentissage par transfert consiste à prendre un CNN formé sur un jeu de données, couper le ou les dernières couches, réentraîner les dernières couches du modèle sur un jeu de données différent.

Intuitivement, vous réutilisez le modèle pour qu'il reconnaisse différentes fonctionnalités de niveau supérieur. En conséquence, le temps de l'entraînement est considérablement réduit, de sorte que l'apprentissage par transfert est un outil utile lorsque vous ne disposez pas suffisamment des données ou lorsque l'entraînement nécessite trop de ressources (PAN et YANG, 2010).

Extraction de caractéristiques

Une approche moins fréquente de deep learning et plus spécialisée consiste à utiliser le réseau comme extracteur de caractéristiques.

Étant donné que toutes les couches sont responsables de l'apprentissage de certaines caractéristiques à partir d'images, nous pouvons récupérer ces caractéristiques du réseau à tout moment du processus d'entraînement. Il est alors possible d'utiliser ces caractéristiques en tant que données d'entrée pour un modèle de machine learning tel que SVM.

5.5 Le système biométrique multimodal proposé

5.5.1 Première stratégie

Un aperçu de cette approche est illustré à la figure 6.3.

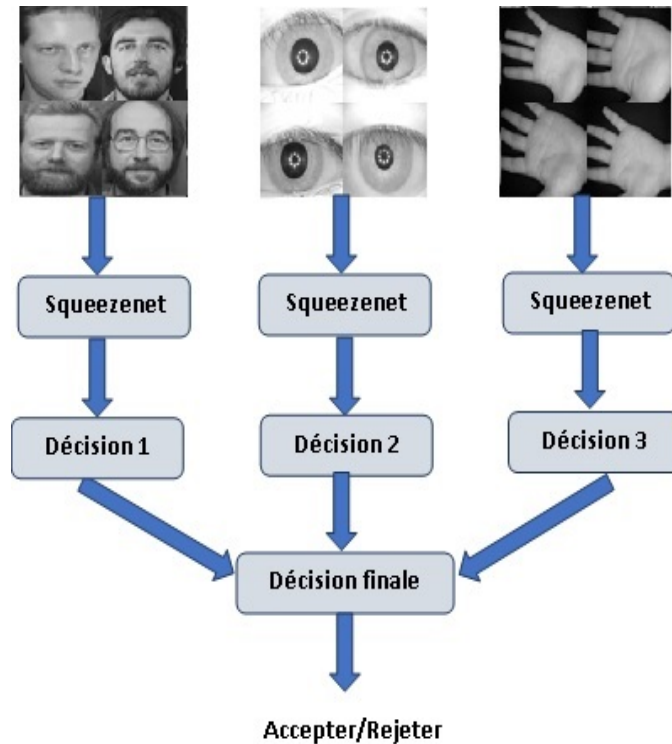


FIGURE 5.3 – Fusion tardive avec des canaux séparés.

Dans cette approche, nous proposons une classification des sujets en fonction de leurs caractéristiques distinctes et prenons donc une décision de classification finale basée sur les autres décisions.

En effet, nous explorons deux méthodes basées sur CNN en utilisant un modèle pré-entraînement téléchargé de la littérature.

En utilisant un modèle pré-entraîné, nous pouvons effectuer deux types d'apprentissage par transfert. Nous pouvons utiliser un élément du modèle, tel que l'extracteur de caractéristique, ou le fine-tuning en l'adaptant à notre tâche avec n classes au lieu de la classification de imagenet avec 1000 classes.

Les étapes suivantes résument l'extraction des caractéristiques profonde proposée :

- (a) Acquérir les images biométriques.
- (b) Redimensionnez les images biométriques en fonction des réseaux profonds. Par exemple, des images couleur 227×227 sont utilisées dans squeezeNet.
- (c) Les caractéristiques sont extraites à l'aide des couches des modèles d'apprentissage profond. Nous prenons les traits de la couche 66 qui est une couche de mise en commun (pool10). La taille du vecteur de caractéristique est 1000.
- (d) La classification est effectuée à l'aide des caractéristiques approfondies avec le classificateur SVM.

Les étapes suivantes résument le fine-tuning :

- (a) Acquérir les images biométriques.
- (b) Redimensionnez les images biométriques en fonction des réseaux profonds.
- (c) Prenons le modèle pré-entraîné et supprimer les cinq dernières couches et remplacer par une couche convolutionnelle, une couche entièrement connectée, une couche softmax et une couche de classification afin d'adopter les réseaux CNN pré-entraînés pour résoudre le problème.
- (d) La classification est effectuée à l'aide du modèle profond nouvellement créé.

Nous utilisons le squeezenet (IANDOLA et al., 2016) avec 68 couches et environ $X/10$ paramètres comparés à AlexNet avec la même précision, l'architecture du modèle est illustrée dans la figure 6.4.

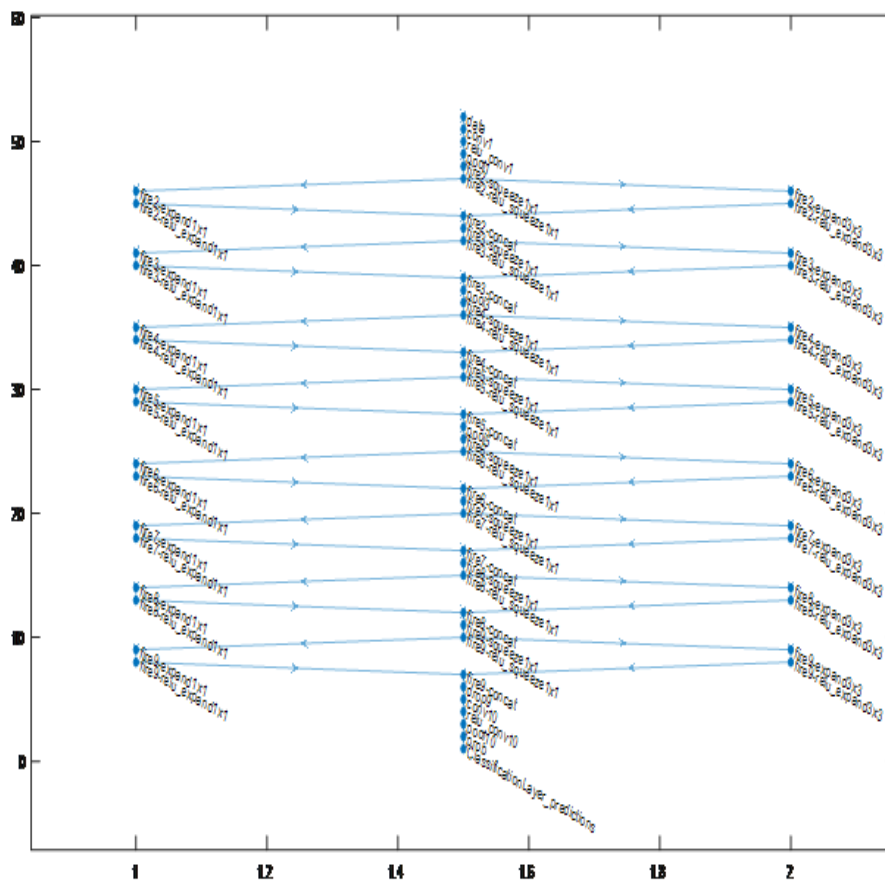


FIGURE 5.4 – Architecture de squeezenet.

5.5.2 La deuxième stratégie

Dans cette stratégie, les vecteurs de caractéristiques sont extraits séparément, en utilisant squeezenet pour extraire les caractéristiques profondes du visage, de l'iris et du palmprint, et nous prenons les caractéristiques de

la couche 66 qui est une couche de regroupement et nous combinons ces vecteurs pour former un vecteur de caractéristiques composites. Ensuite fait une classification à l'aide de SVM (voir figure 6.5).

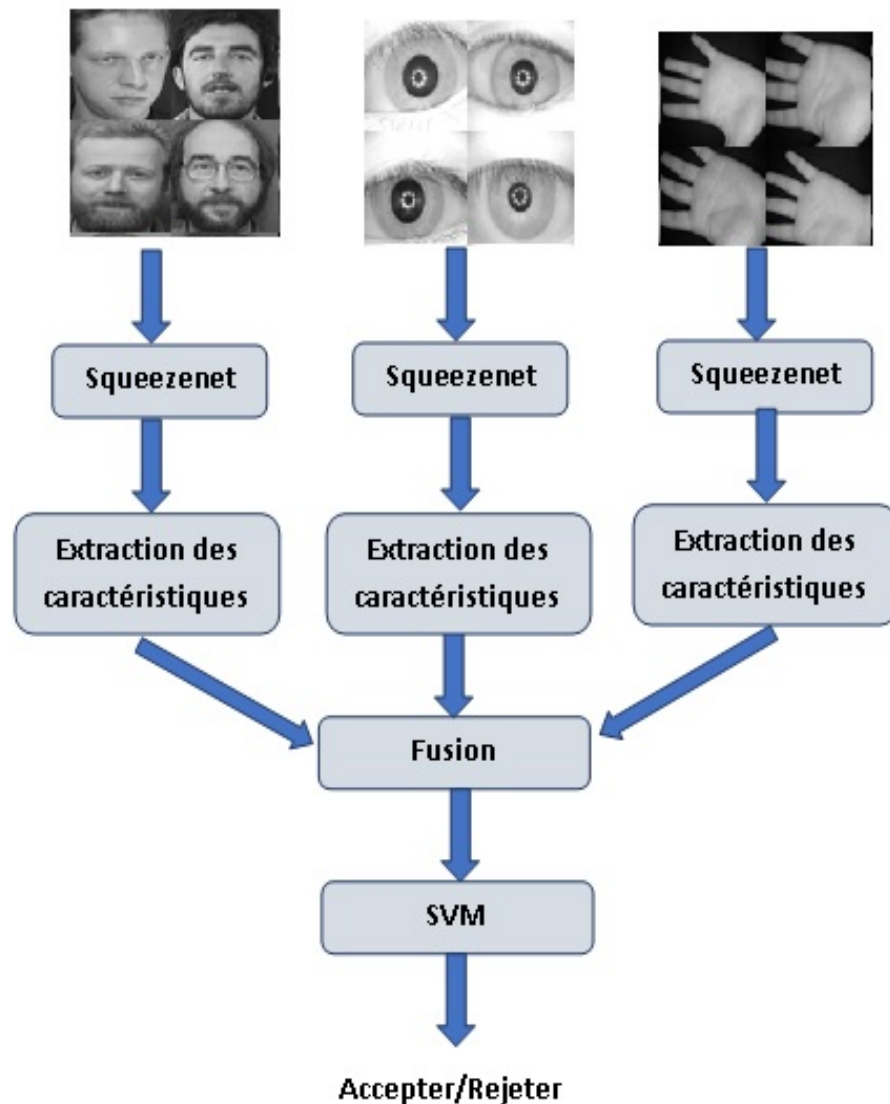


FIGURE 5.5 – Fusion au niveau des caractéristique.

5.6 Résultats expérimentaux et discussion

Dans cette section, nous discuterons des jeux de données utilisées et les résultats expérimentaux générés.

5.6.1 Base de données utilisées

Le jeu de données utilisées contient trois sous-ensembles, chacun concernant une caractéristique, comme le visage, les iris et palmprint, recueillis auprès d'une centaine de personnes.

Nous supposons que la même personne a des traits biométriques différents indépendants. Par conséquent, nous pouvons construire des bases de données multimodales virtuelles en combinant un sujet de la base de données face94 (*Face94 face database*) avec un sujet de la base de données IITD iris (*IITD Iris database*) et de la base de données CASIA palmprint (*CASIA Database*).

La base de données face94 contient 152 personnes ; chaque personne a 20 images avec une résolution de 180×200 pixels. L'éclairage est artificiel et certaines images sont capturées avec des lunettes et un mélange de tungstène et de lumière fluorescente.

La base de données IITD contient 224 personnes ; chaque personne a 5 images pour l'iris gauche et 5 images de l'iris droit.

La base de données CASIA Palmprint a été capturée par 312 personnes différentes à l'aide d'un dispositif d'imagerie spectrale multiple conçue par lui-même. Chaque personne a 8 images pour la main gauche et 8 images de la main droite.

De plus la méthode proposée est appliquée sur la base de donnée faciale ORL (*ORL database*), de l'oreille USTB I et USTB I (*Ear database*).

La base de données ORL a été prise entre avril 1992 et avril 1994. Elle contient 40 sujets et chaque sujet a 10 images. Les images ont été prises sous différents éclairages et expressions faciales.

Les bases de données USTB ont été collectées auprès des étudiants et des enseignants du département d'ingénierie de l'information de l'Université des sciences et technologies de Beijing (USTB). La base de données USTB I contient 60 personnes ; chaque personne a 3 images. Le jeu de données USTB I a été pris avec un éclairage standard. La base de données USTB II contient 77 sujets. Chaque sujet a 4 images. Il a été pris sous différents éclairages et rotations à 30 et -30 degrés.

5.6.2 Résultats expérimentaux

La mise en œuvre des différents algorithmes de notre système ainsi que le déroulement de tous les tests sont effectués sous Matlab 2018.

Première stratégie :

Extraction des caractéristiques : nous obtenons les résultats suivants en utilisant des caractéristiques distinctes ; nous pouvons faire une fusion tardive pour obtenir la décision finale en mettant en place un mécanisme de décision.

Ces vecteurs de caractéristique sont utilisées pour entraîner sur un SVM. Nous avons divisé la base de données à 50% pour l'entraînement et à 50% pour tester afin d'éviter le problème du surajustement.

Les résultats expérimentaux obtenus sont illustrés dans les figures 6.6, 6.7, 6.8 et 6.9.

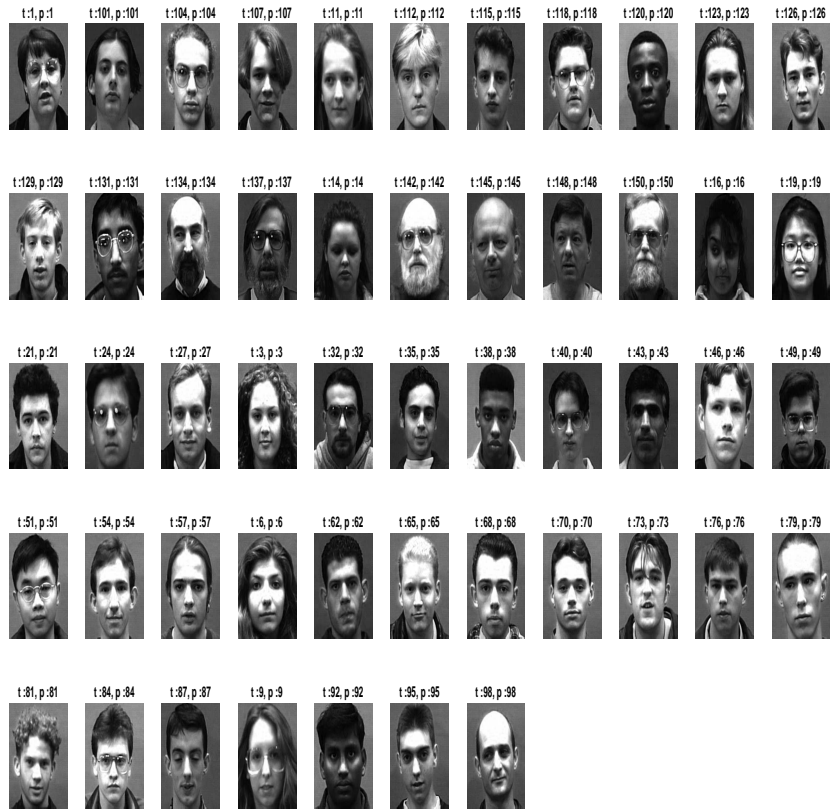


FIGURE 5.6 – Exemples de classes vraies et prédites

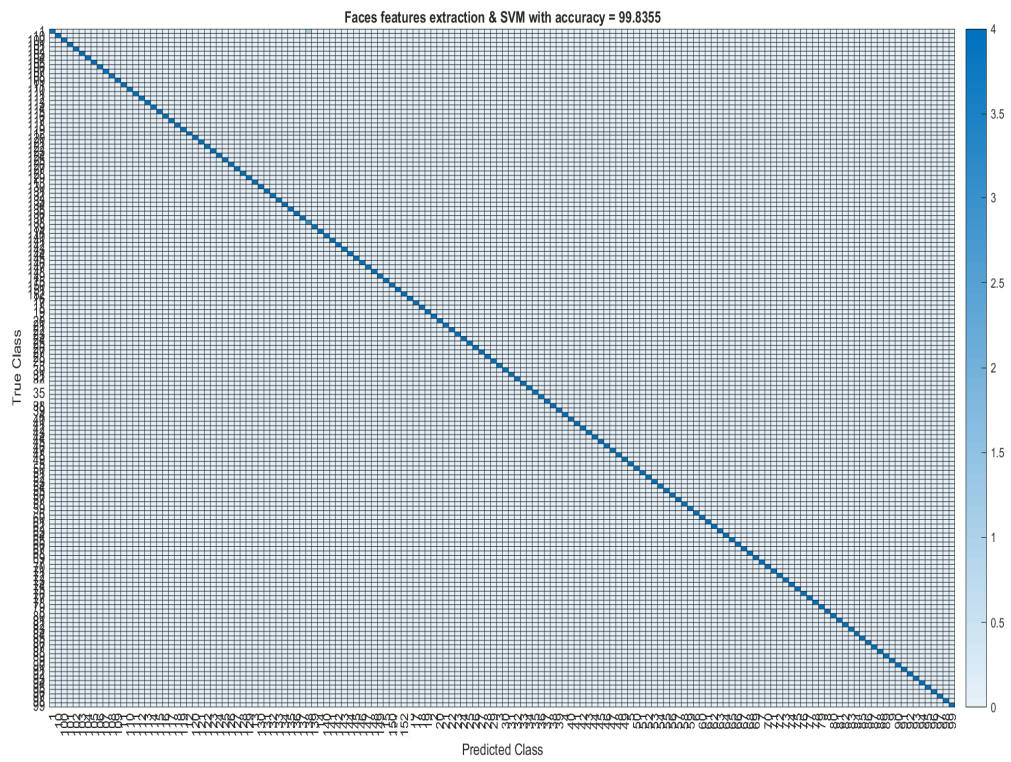


FIGURE 5.7 – Matrice de confusion et score de précision de visage.

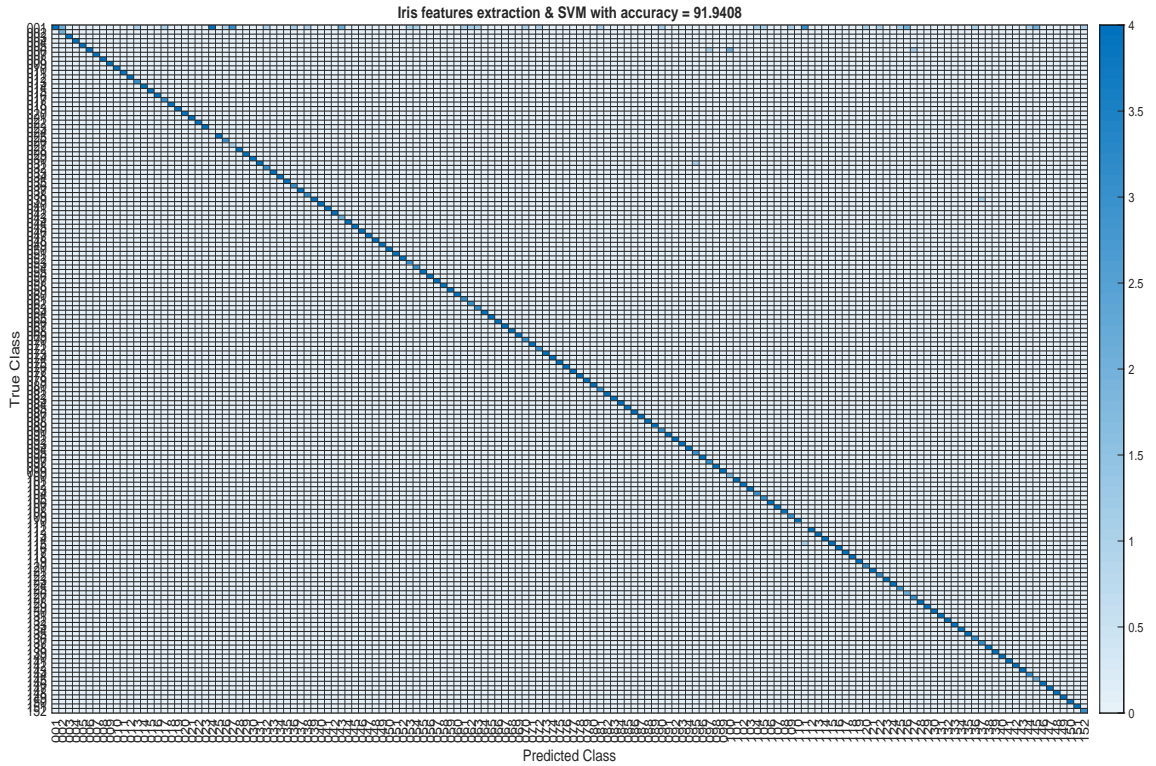


FIGURE 5.8 – Matrice de confusion et score de précision de l’iris.

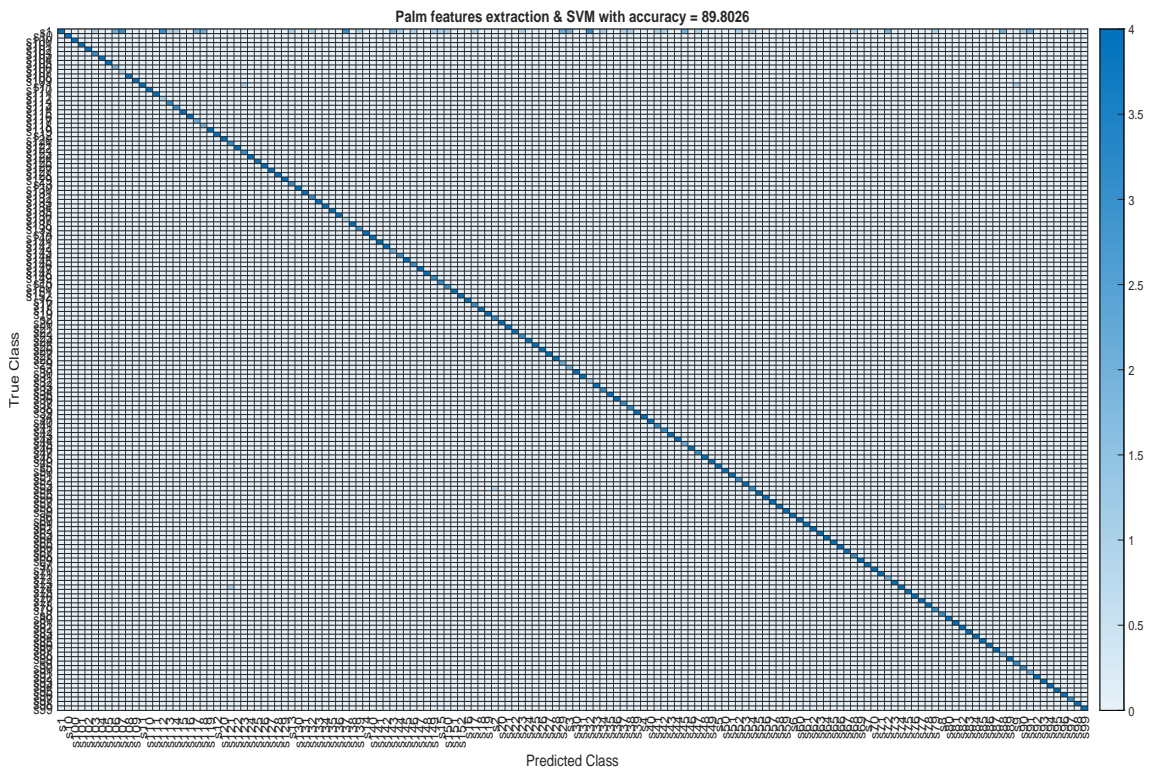


FIGURE 5.9 – Matrice de confusion et score de précision de palmprint.

Précision de fusion au niveau de décision avec OR= 100%.

Précision de fusion au niveau de décision avec AND-OR=99.5066%.

Fine-tuning : Nous testons les performances de la seconde méthode d'apprentissage par transfert, prenons le modèle pré entraîné et supprimer les cinq dernières couches; 'conv10', 'relu_conv10', 'pool10', 'prob', 'ClassificationLayer_predictions'.

Et nous les remplaçons par une couche convolutionnelle, une couche entièrement connectée, une couche softmax et une couche de classification.

Les figures 6.10, 6.11 et 6.12 illustrent le processus d'entraînement utilisant le visage, l'iris et palmprint.

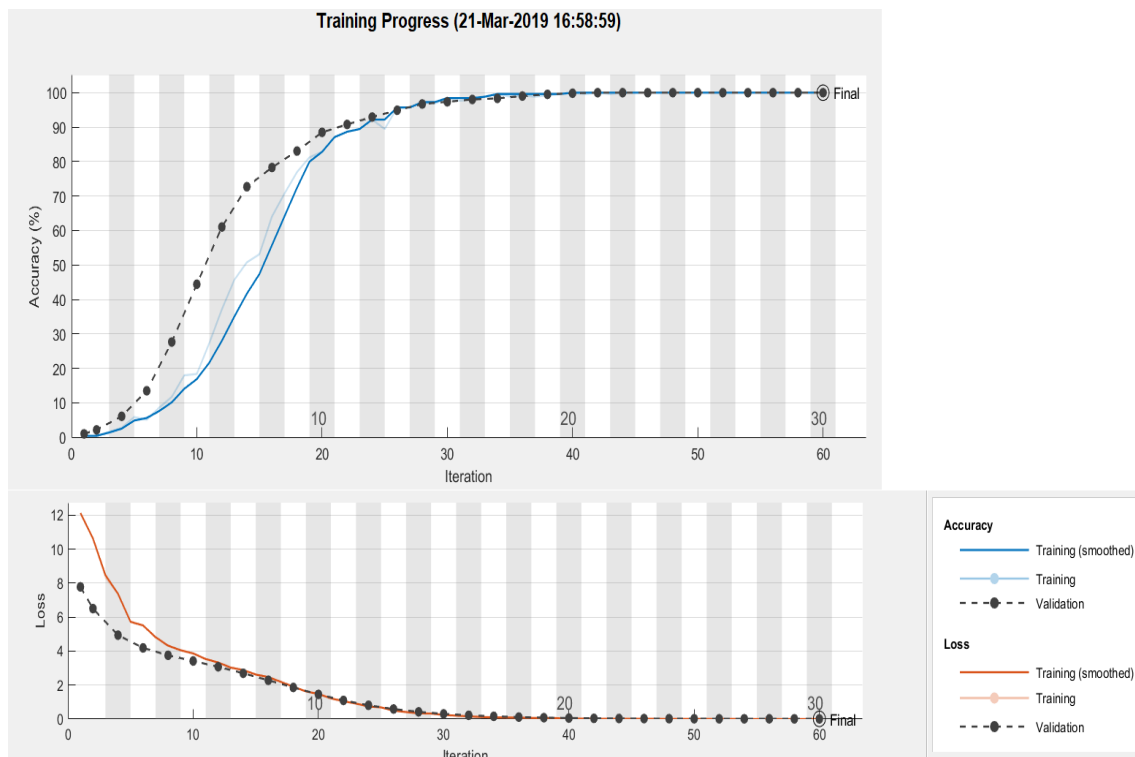


FIGURE 5.10 – Précision de l'entraînement de validation, et perte au fil des itérations de visage.

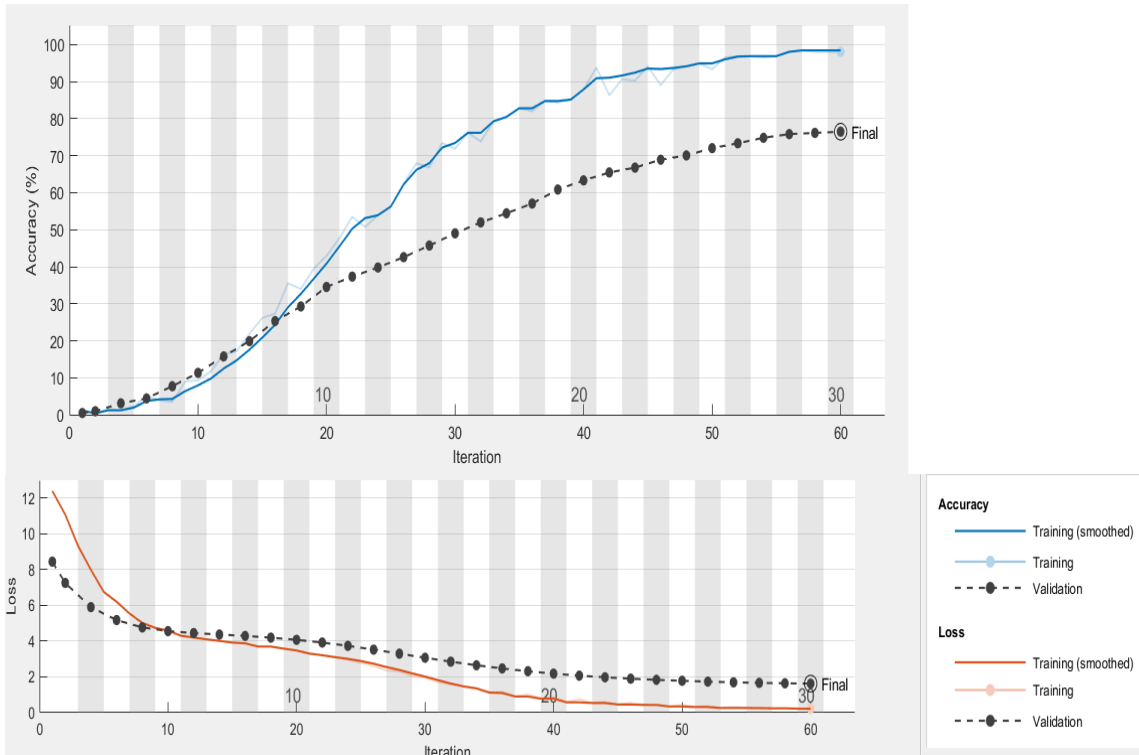


FIGURE 5.11 – Précision de l’entrainement de validation, et perte au fil des itérations de l’iris.

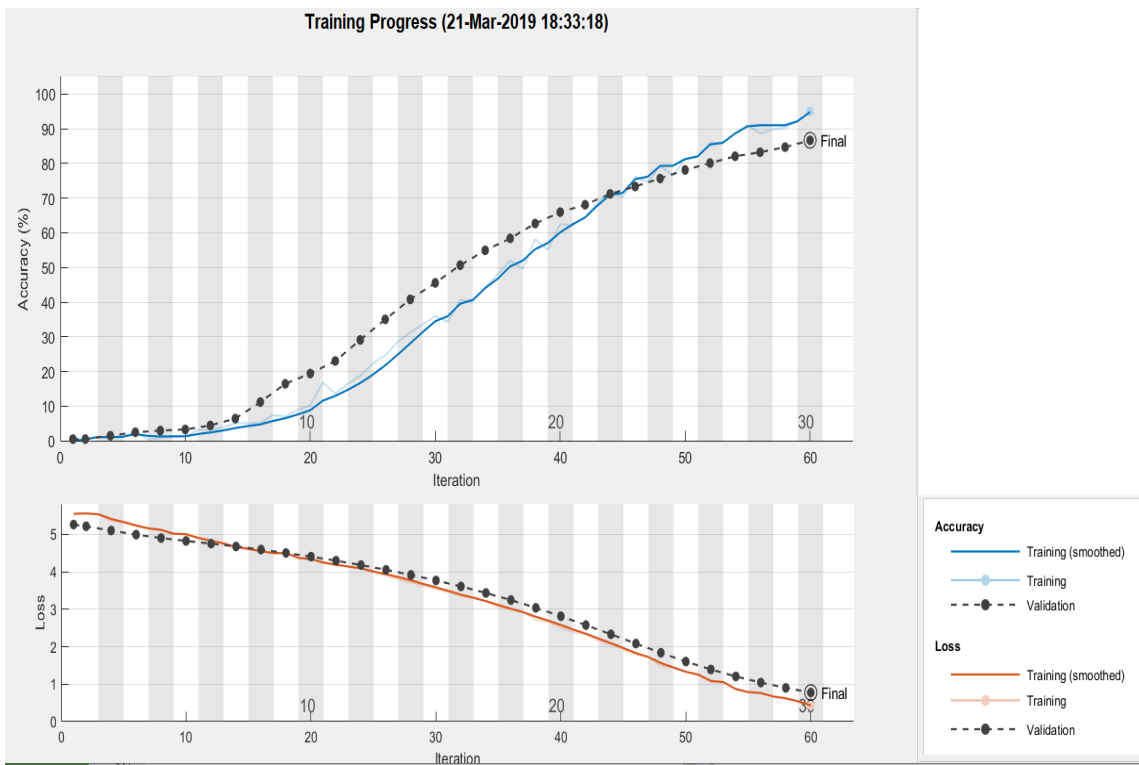


FIGURE 5.12 – Précision de l’entrainement de validation, et perte au fil des itérations de palmprint.

Et nous testons le modèle obtenu en utilisant le sous-ensemble de tests, nous obtenons les résultats suivants (voir figures 6.13, 6.14, 6.15 et 6.16) :

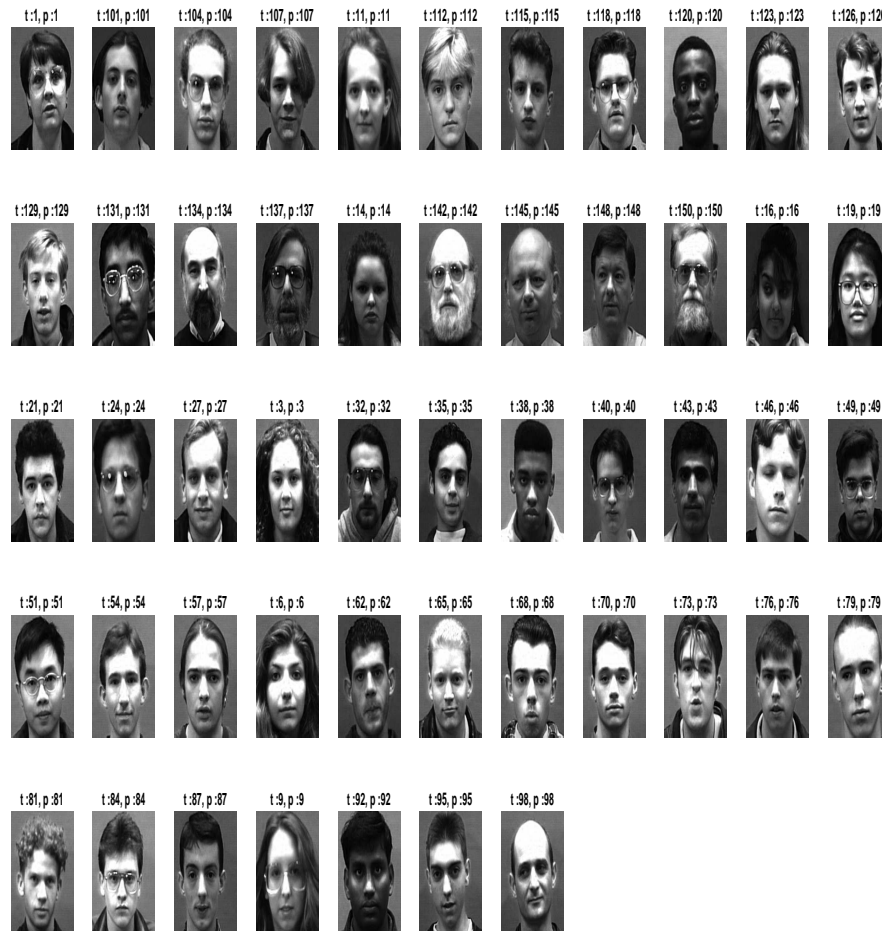


FIGURE 5.13 – Exemples de classes vraies et prédites.

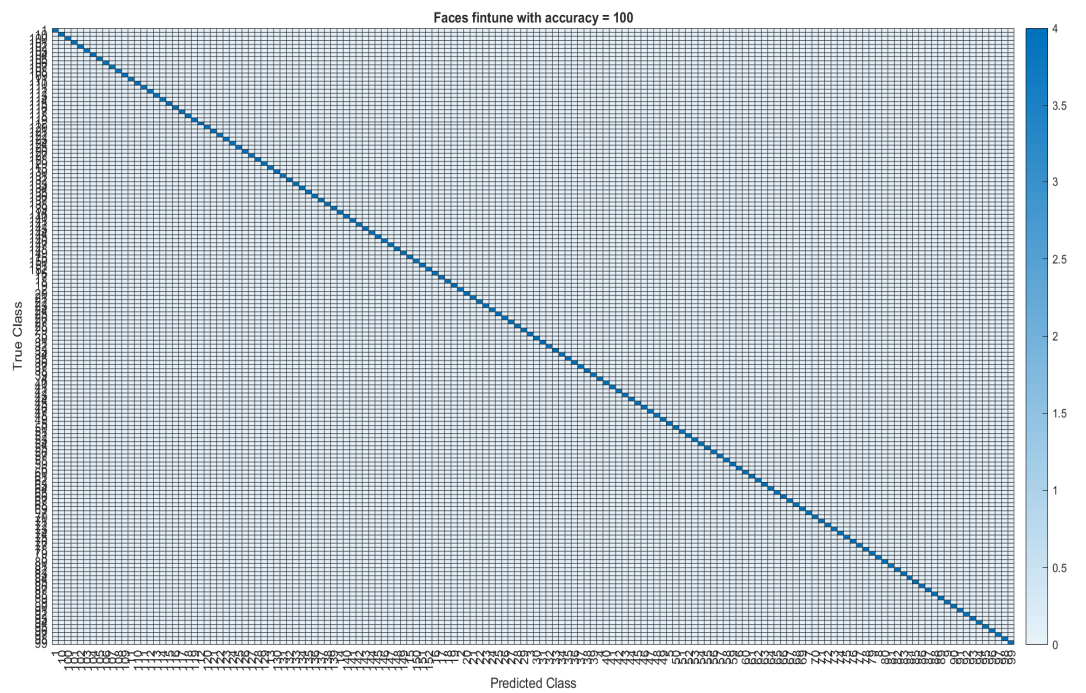


FIGURE 5.14 – Matrice de confusion et score de précision de visage.

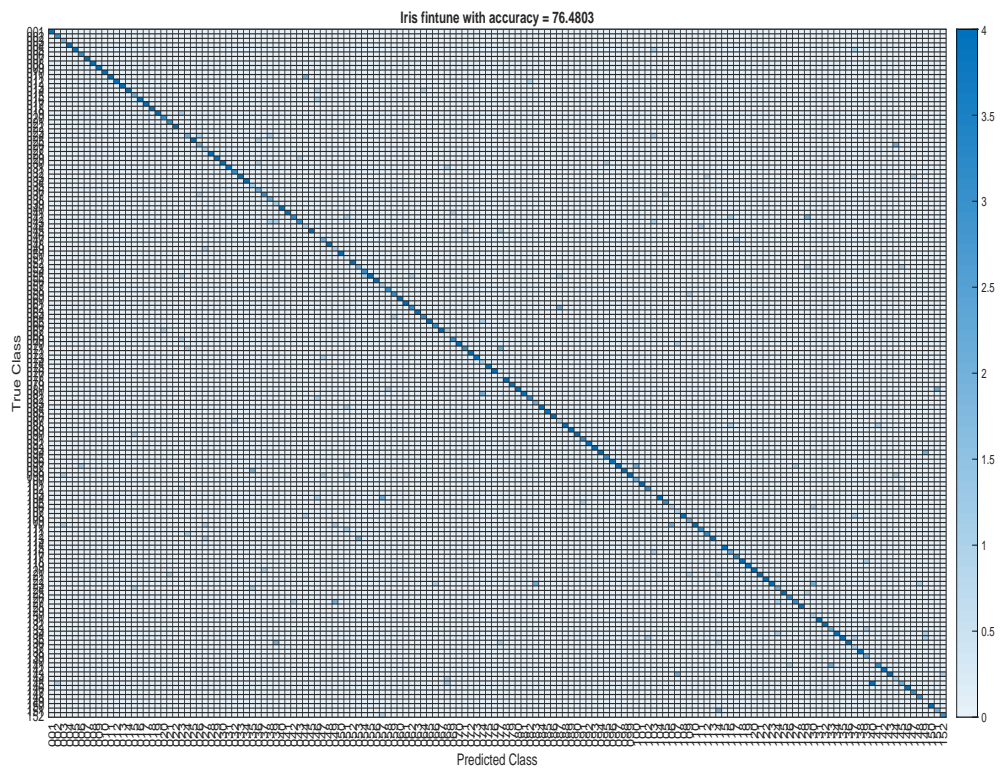


FIGURE 5.15 – Matrice de confusion et score de précision de l'iris.

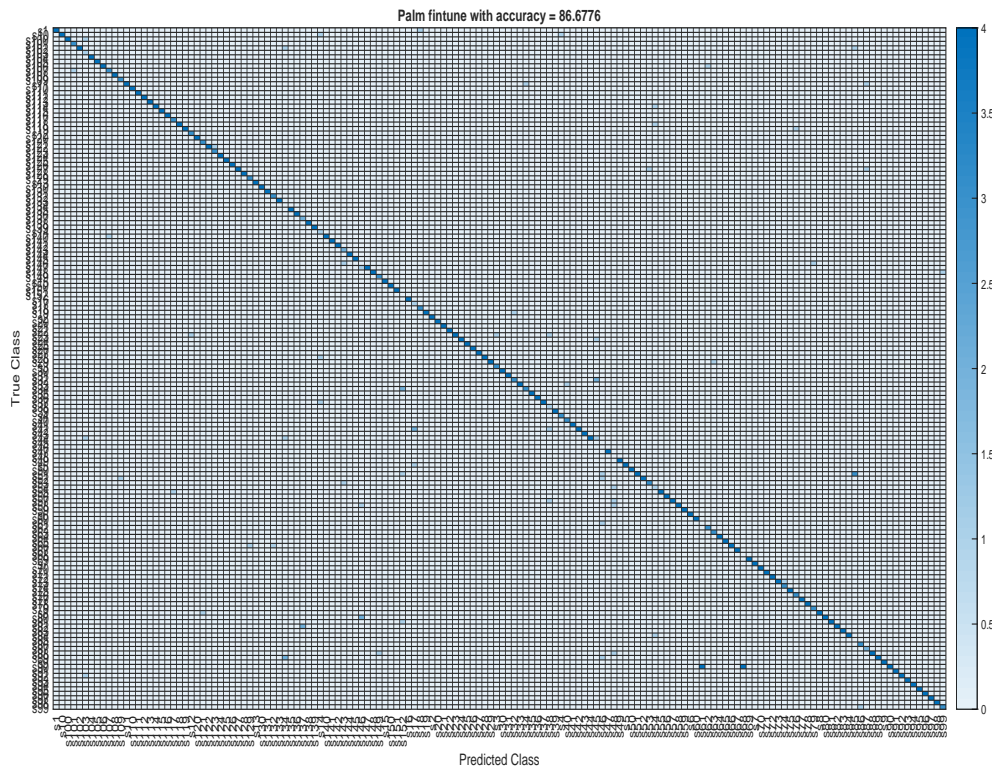


FIGURE 5.16 – Matrice de confusion et score de précision de palmprint.

Précision de fusion avec OR = 100%.

Précision de fusion avec AND-OR=96.5461%.

La deuxième stratégie :

Les résultats obtenus de cette stratégie sont résumés dans la figure 6.17.

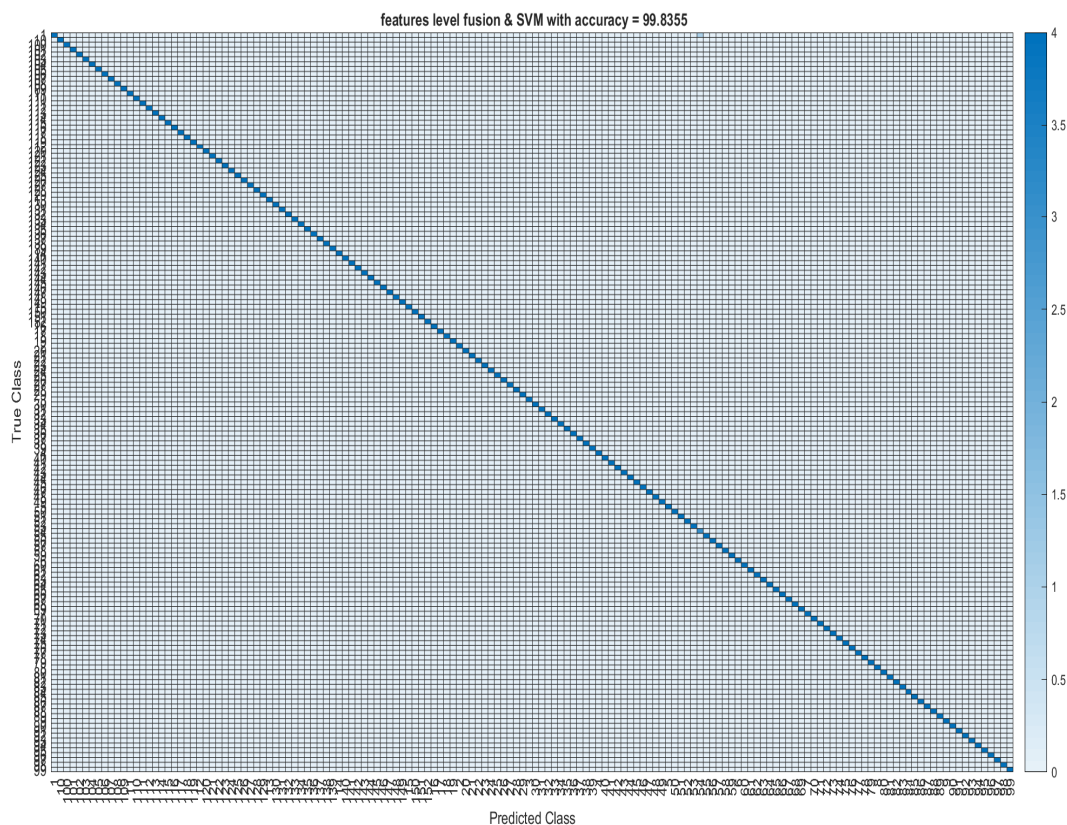


FIGURE 5.17 – Matrice de confusion et score de précision de la deuxième stratégie.

Toutes les expériences ont été répétées dix fois en utilisant des ensembles d'entraînement et de test choisis au hasard, et les résultats ont été moyennés (voir tableau 6.1).

Dans tableaux 6.1, 6.2, 6.3 stratégie 1 EC et stratégie 1 FT représente la première stratégie applique l'extraction des caractéristiques and la première stratégie applique le fine-tuning , respectivement.

TABLE 5.1 – Les résultats obtenus avec stratégie 1 et stratégie 2.

Système biométrique	Taux de reconnaissance %		
	Stratégie 1 FT	Stratégie 1 EC	Stratégie 2
Visage	99.80	99.90	–
Iris	78.67	72.67	–
Palmprint	90.46	87.77	–
Iris+Palmprint	98.14	97.10	94.70
Iris+visage	99.93	100	99.62
Visage+Palmprint	100	100	99.78
Visage+Iris+Palmprint	100	100	99.97

Comme montre le tableau 6.1 la fusion de l'iris, du visage et des palmprint conduit à une plus grande précision de reconnaissance comparée

au système unimodal et système bimodaux. Ces résultats montrent la robustesse et l'efficacité de notre système multimodal proposé.

5.6.3 Comparaisons de performance

Les résultats obtenus avec les deux approches proposer appliquer sur les images de visage et oreille sont comparés avec l'approche proposer dans (OMARA IBRAHIM, 2017)(voir les tableaux 6.2 et 6.3).

Pour construire MD1, nous prenons les trois premières images de visage de chaque sujet de la base de données ORL et les 40 premiers sujets de l'ensemble de données USTB I; par conséquent, chaque sujet de MD1 a trois images de face et d'oreille, séparément.

D'autre part, MD2 est construit à partir des quatre premières images de visage de chaque sujet de la base de données ORL et des 40 premiers sujets de l'ensemble de données USTB II.

Nous envisageons deux scénarios pour réaliser nos expériences et assurer une comparaison équitable entre notre approche proposée et l'approche proposer dans (OMARA IBRAHIM, 2017).

Dans le premier scénario, nous considérons deux images de chaque sujet pour l'entraînement (MD1 et MD2 avec une image et deux images a testé, respectivement), et dans le deuxième scénario, nous considérons trois images pour chaque sujet pour l'entraînement et une image pour les tests (MD2).

Comme le montrent les tableaux 6.2 et 6.3, nous constatons que nos résultats montrent systématiquement de meilleures performances de classification basées sur les caractéristiques profondes multimodales pour le visage et l'oreille que le système unimodal individuel visage ou oreille séparément.

TABLE 5.2 – Comparaison du taux de reconnaissance du visage, de l'oreille et de leur fusion dans le scénario 1.

Bases de données	Taux de reconnaissance %			
	(OMARA IBRAHIM, 2017)	Stratégie 1 EC	Stratégie 1 FT	Stratégie 2
Visage	87.5	97.63	90.65	–
Oreille USTB I	92.5	94.30	90.23	–
Oreille USTB II	81.3	96.25	82.21	–
Visage+oreille MD1	100	100	97.65	100
Visage+oreille MD2	100	100	95.74	99.75

TABLE 5.3 – Comparaison du taux de reconnaissance du visage, de l'oreille et de leur fusion dans le scénario 2.

Bases de données	Taux de reconnaissance %			
	(OMARA IBRAHIM, 2017)	Stratégie 1 EC	Stratégie 1 FT	Stratégie 2
Visage	92.5	97.5	92.5	–
Oreille USTB II	92.5	95	91.3	–
Visage+oreille MD2	100	100	98.87	100

5.7 Conclusions

Dans ce chapitre, nous avons proposé une approche originale pour l'identification multibiométrique basée sur l'apprentissage profond.

Nous avons essayé de classer les objets en fonction de leurs caractéristiques distinctes et prenons une décision finale de classification basée sur les autres décisions. Pour cela, nous avons exploré deux méthodes basées sur CNN en utilisant un modèle pré-entraînement téléchargé de la littérature (Squeezenet).

Nous avons utilisé un modèle squeezenet car nous ne pouvons pas créer un modèle à partir de rien en raison du petit nombre d'images contenues dans le jeu de données.

En utilisant un modèle pré-entraînement, nous pouvons effectuer deux types d'apprentissage par transfert. Nous pouvons utiliser un élément du modèle, tel que l'extracteur de caractéristiques, ou le fine-tuning en l'adaptant à notre tâche au lieu d'imagenet avec 1000 classes.

Les résultats expérimentaux des systèmes biométriques multimodaux proposés basés sur squeezenet fournissent un taux de reconnaissance prometteur.

Conclusion générale

Durant ce travail de thèse, d'abord nous nous sommes intéressés à la biométrie multimodale par l'iris, palmprint et le visage. Nous avons commencé par définir le contexte du sujet de la thèse en introduisant des notions de base sur les systèmes biométriques en général et les outils pour les évaluer.

Ensuite nous avons présenté les différents aspects des systèmes biométriques multimodaux par leur architecture, leurs sources d'information et leurs niveaux de fusion. Suivre des études bibliographiques sur les systèmes multimodaux.

Dans la première contribution, nous avons présenté une nouvelle méthode de reconnaissance basée sur la forme de la main, qui combine les moments invariants de Hu et les moments de Legendre. En fait, l'utilisation d'un seul vecteur de caractéristiques ne constitue pas une discrimination pour l'identification.

L'algorithme proposé présente certains avantages. Tout d'abord, l'orthogonalité et l'invariance géométrique permettent de minimiser la redondance de l'information et d'accroître le pouvoir de discrimination. Une plus grande précision peut être obtenue en utilisant des classificateurs plus complexes tels que les réseaux de neurones.

Dans la deuxième contribution, nous avons présenté une fusion au niveau du score du système de reconnaissance multimodal de l'iris, du visage et de palmprint.

Les transformations Gabor, DWT et CT sont utilisées pour obtenir une représentation, puis PCA, LDA et KFA sont utilisées pour réduire la taille des vecteurs de caractéristiques. Ensuite, les vecteurs de caractéristiques sont classés à l'aide de K-NN. Enfin, les résultats des trois classificateurs sont fusionnés avec la règle de somme pondérée et PSO.

Dans la troisième contribution, nous avons proposé une approche originale pour l'identification multibiométrique basée sur le deep learning.

Pour cela, nous avons exploré deux méthodes basées sur CNN en utilisant un modèle pré-entraînement téléchargé de la littérature (Squeezenet).

Nous avons effectué deux types d'apprentissage par transfert. Nous pouvons utiliser un élément du modèle, tel que l'extracteur de caractéristique, ou le fine-tuning en l'adaptant à notre tâche avec n classes au lieu d'imagenet avec 1000 classes.

Dans les perspectives, nous avons envisagé d'explorer d'autres signatures biométriques.

Des recherches supplémentaires seront nécessaires pour valider l'efficacité des approches proposées en utilisant des bases de données plus volumineuses contenant des images plus complexes.

Bibliographie

- ABDOLAH, Mohamad, Majid MOHAMADI et Mehdi JAFARI (2013). « Multimodal Biometric system Fusion Using Fingerprint and Iris with Fuzzy Logic ». In : *International Journal of Soft Computing and Engineering (IJSCE)* 2.
- AGBINYA, Johnson (déc. 2018). « New Human Hand Geometry Modeling for Biometric Security Systems ». In : International Conference on Cyber Security et Communication Systems (ICCSCS2018).
- AHMAD, M. I. et al. (2014). « Information fusion of face and palmprint multimodal biometrics ». In : *2014 IEEE REGION 10 SYMPOSIUM*, p. 635-639. DOI : [10.1109/TENCONSpring.2014.6863111](https://doi.org/10.1109/TENCONSpring.2014.6863111).
- AL-WAISY, A. S. et al. (2017). « A multimodal biometric system for personal identification based on deep learning approaches ». In : *2017 Seventh International Conference on Emerging Security Technologies (EST)*, p. 163-168. DOI : [10.1109/EST.2017.8090417](https://doi.org/10.1109/EST.2017.8090417).
- AL-WAISY, Alaa S. et al. (2018). « A multi-biometric iris recognition system based on a deep learning approach ». In : *Pattern Analysis and Applications* 21.3, p. 783-802. ISSN : 1433-755X. DOI : [10.1007/s10044-017-0656-1](https://doi.org/10.1007/s10044-017-0656-1). URL : <https://doi.org/10.1007/s10044-017-0656-1>.
- ALLANO, L. et al. (2009). *La biométrie multimodale : stratégies de fusion de scores et mesures de dépendance appliquées aux bases de personnes virtuelles*. URL : <https://books.google.dz/books?id=nkLoSAAACAAJ>.
- ALNASIRI, Thawar et al. (jan. 2009). « Object classification via geometrical, zernike and legendre moments ». In : *Journal of Theoretical and Applied Information Technology* 7, p. 031-037.
- ALSAADE, F. et al. (mar. 2008). « Enhancement of Multimodal Biometric Segregation Using Unconstrained Cohort Normalisation ». In : *Pattern Recogn.* 41.3, p. 814-820. ISSN : 0031-3203. DOI : [10.1016/j.patcog.2007.06.028](https://doi.org/10.1016/j.patcog.2007.06.028). URL : <http://dx.doi.org/10.1016/j.patcog.2007.06.028>.
- ALTENBERGER, Felix et Claus LENZ (2018). « A Non-Technical Survey on Deep Convolutional Neural Network Architectures ». In : *CoRR* abs/1803.02129. arXiv : [1803.02129](https://arxiv.org/abs/1803.02129). URL : <http://arxiv.org/abs/1803.02129>.
- AMAYEH, G. (2006). « Peg-free hand shape verification using high order zernike moments ». In : *Computer Vision and Pattern Recognition Workshop, CVPRW, Reno, USA*, p. 40-40.
- ANTAL, Margit, Laszlo SZABO et Tünde TORDAI (fév. 2018). « Online Signature Verification on MOBISIG Finger-Drawn Signature Corpus ». In : *Mobile Information Systems* 2018, p. 1-15.

- ANTIPOV, Grigory (déc. 2017). « Deep learning for semantic description of visual human traits ». Theses. Télécom ParisTech. URL : <https://pastel.archives-ouvertes.fr/tel-01725853>.
- ATREY, Pradeep K. et al. (nov. 2010). « Multimodal Fusion for Multimedia Analysis : A Survey ». In : *Multimedia Syst.* 16.6, p. 345-379. ISSN : 0942-4962. DOI : [10.1007/s00530-010-0182-0](https://doi.org/10.1007/s00530-010-0182-0). URL : <http://dx.doi.org/10.1007/s00530-010-0182-0>.
- ATTALLAH, Bilal, Amina SERIR et Youssef CHAHIR (2018). « Feature extraction in palmprint recognition using spiral of moment skewness and kurtosis algorithm ». In : *Pattern Analysis and Applications*, p. 1-9.
- BARBU, T., A. CIOBANU et M. LUCA (2015). « Multimodal biometric authentication based on voice, face and iris ». In : *2015 E-Health and Bioengineering Conference (EHB)*, p. 1-4. DOI : [10.1109/EHB.2015.7391373](https://doi.org/10.1109/EHB.2015.7391373).
- BAYRAM, Kadir Sercan et Bülent BOLAT (2018). « Multibiometric identification by using ear, face, and thermal face ». In : *EURASIP Journal on Image and Video Processing* 2018.1, p. 32. ISSN : 1687-5281. DOI : [10.1186/s13640-018-0274-x](https://doi.org/10.1186/s13640-018-0274-x). URL : <https://doi.org/10.1186/s13640-018-0274-x>.
- BEHERA, Mithuna et V.K. GOVINDAN (2014). « Palm print Authentication Using PCA Technique ». In : *International Journal of Computer Science and Information Technologies* 5, p. 3638-3640.
- BHARADI, V. A., B. PANDYA et B. NEMADE (2014). « Multimodal biometric recognition using iris and fingerprint : By texture feature extraction using hybrid wavelets ». In : *2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence)*, p. 697-702. DOI : [10.1109/CONFLUENCE.2014.6949309](https://doi.org/10.1109/CONFLUENCE.2014.6949309).
- BISWAS, Suparna et Jaya SIL (2015). « An efficient face recognition method using contourlet transform ». In : *2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICR-CICN)*, p. 27-31. DOI : [10.1109/ICRCICN.2015.7434204](https://doi.org/10.1109/ICRCICN.2015.7434204).
- BOLLE, Ruud et Sharath PANKANTI (1998). *Biometrics, Personal Identification in Networked Society : Personal Identification in Networked Society*. Sous la dir. d'Anil K. JAIN. Norwell, MA, USA : Kluwer Academic Publishers. ISBN : 0792383451.
- BORAH, Tripti Rani, Kandarpa Kumar SARMA et Pran Hari TALUKDAR (2015). « Retina recognition system using adaptive neuro fuzzy inference system ». In : *2015 International Conference on Computer, Communication and Control (IC4)*, p. 1-6. DOI : [10.1109/IC4.2015.7375663](https://doi.org/10.1109/IC4.2015.7375663).
- BOUZOUINA, Y. et L. HAMAMI (2017). « Multimodal biometric : Iris and face recognition based on feature selection of iris with GA and scores level fusion with SVM ». In : *2017 2nd International Conference on Bioengineering for Smart Technologies (BioSMART)*, p. 1-7. DOI : [10.1109/BIOSMART.2017.8095312](https://doi.org/10.1109/BIOSMART.2017.8095312).
- BULATOV, Yaroslav et al. (2004). « Hand Recognition Using Geometric Classifiers ». In : *Biometric Authentication*. Sous la dir. de David ZHANG et Anil K. JAIN. Berlin, Heidelberg : Springer Berlin Heidelberg, p. 753-759. ISBN : 978-3-540-25948-0.

- CAO, Kai et Anil K. JAIN (2015). « Learning Fingerprint Reconstruction : From Minutiae to Image ». In : *IEEE Transactions on Information Forensics and Security* 10.1, p. 104-117. ISSN : 1556-6013. DOI : [10.1109/TIFS.2014.2363951](https://doi.org/10.1109/TIFS.2014.2363951).
- CASIA Database. <http://biometrics.idealtest.org/>. Accessed : 30-10-2017.
- CASIA-MS Database. <http://biometrics.idealtest.org/>. Accessed : 30-10-2017.
- CHARFI, Nesrine (jan. 2017). « Biometric recognition based on hand shape and palmprint modalities ». Theses. Ecole nationale supérieure Mines-Télécom Atlantique. URL : <https://tel.archives-ouvertes.fr/tel-01781354>.
- CHERIFI, D., I. HAFNAOUI et A. NAIT-ALI (2015). « Multimodal Score-Level Fusion Using Hybrid GA-PSO for Multibiometric System ». In : *Informatika (Slovenia)* 39, 209—216.
- CONNOR, Patrick et Arun ROSS (fév. 2018). « Biometric Recognition By Gait : A Survey of Modalities and Features ». In : *Computer Vision and Image Understanding* 167.
- DAVIS, J. et G. BRADSKI (1999). « Real-time Motion Template Gradients using Intel CVLib ». In : *IEEE ICCV Workshop on Framerate Vision*.
- DUTA, N. (2009). « A survey of biometric technology based on hand shape ». In : *Pattern Recognition* 42.11, p. 2797-2806.
- DUTAGACI, Helin, Bülent SANKUR et Erdem YÖRÜK (2008). « Comparative analysis of global hand appearance-based person recognition ». In : *J. Electronic Imaging* 17, p. 1-19.
- Ear database. <http://www1.ustb.edu.cn/resb/en/index.html>. Accessed : 01-05-2019.
- EJARQUE, Pascual et al. (2009). « Eigenfeatures and Supervectors in Feature and Score Fusion for SVM Face and Speaker Verification ». In : *Biometric ID Management and Multimodal Communication*. Sous la dir. de Julian FIERREZ et al. Berlin, Heidelberg : Springer Berlin Heidelberg, p. 81-88.
- EKKA, B. K., N. B. PUHAN et R. PANDA (2015). « Retinal verification using point set matching ». In : *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, p. 159-163. DOI : [10.1109/SPIN.2015.7095402](https://doi.org/10.1109/SPIN.2015.7095402).
- EL-ABED, Mohamad (déc. 2011). « Evaluation of Biometric Systems ». Theses. Université de Caen. URL : <https://tel.archives-ouvertes.fr/tel-01007679>.
- ELREFAEI, Lamiaa A. et al. (2018). « Developing Iris Recognition System for Smartphone Security ». In : *Multimedia Tools and Applications* 77.12, p. 14579-14603. ISSN : 1573-7721. DOI : [10.1007/s11042-017-5049-3](https://doi.org/10.1007/s11042-017-5049-3). URL : <https://doi.org/10.1007/s11042-017-5049-3>.
- ESKANDARI, Maryam, Önsen TOYGAR et Hasan DEMIREL (2014). « Feature extractor selection for face-iris multimodal recognition ». In : *Signal, Image and Video Processing* 8.6, p. 1189-1198. DOI : [10.1007/s11760-014-0659-y](https://doi.org/10.1007/s11760-014-0659-y).

- Face94 face database. <http://cmp.felk.cvut.cz/~spacelib/faces/faces94.html>. Accessed : 19-05-2018.
- FINE, Terrence L. (1999). *Feedforward Neural Network Methodology*. Sous la dir. de S. L. LAURITZEN et al. 1st. Berlin, Heidelberg : Springer-Verlag. ISBN : 0387987452.
- FLUSSER, J. et T. SUK (1993). « Pattern Recognition by Affine Moment Invariants ». In : *Pattern Recognition* 26.1, p. 167-174.
- FLYNN JAIN ANIL Ross Patrick, Arun A. (jan. 2008). *Handbook of Biometrics*. ISBN : 978-0-387-71040-2.
- GHOUALMI, Lamis, Amer DRAA et Salim CHIKHI (mar. 2016). « An Ear Biometric System based on Artificial Bees and the Scale Invariant Feature Transform ». In : *Expert Systems with Applications* 57.
- GOODFELLOW, Ian, Yoshua BENGIO et Aaron COURVILLE (2016). *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press.
- GOWDA, H. D. Supreetha, G. Hemantha KUMAR et Mohammad IMRAN (2018). « Multi-modal biometric system on various levels of fusion using LPQ features ». In : *Journal of Information and Optimization Sciences* 39.1, p. 169-181. DOI : [10.1080/02522667.2017.1372918](https://doi.org/10.1080/02522667.2017.1372918). eprint : <https://doi.org/10.1080/02522667.2017.1372918>. URL : <https://doi.org/10.1080/02522667.2017.1372918>.
- GUERFI, Souhila (oct. 2008). « Face authentication by biometrical 2D/3D faces characteristics recognition ». Theses. Université d'Evry-Val d'Essonne. URL : <https://tel.archives-ouvertes.fr/tel-00623243>.
- GUO, Jing-Ming et al. (2012). « Contact-free hand geometry-based identification system ». In : *Expert Systems with Applications* 39.14, p. 11728-11736. ISSN : 0957-4174. DOI : <https://doi.org/10.1016/j.eswa.2012.04.081>. URL : <http://www.sciencedirect.com/science/article/pii/S0957417412006781>.
- HAGHIGHAT MOHAMMAD, Zonouz Saman et Abdel-Mottaleb MOHAMMED (2013). « Identification Using Encrypted Biometrics ». In : *Computer Analysis of Images and Patterns (CAIP), At York, UK*.
- HASSAN, Soliman, Mohamed Abdelnasser SABER et Atwan AHMED (2012). « Feature Level Fusion of Palm Veins and Signature Biometrics ». In : *International Journal of Video Image Processing and Network Security IJVIPNS-IJENS Vol : 12 No*, p. 28.
- HATCHER, W. G. et W. YU (2018). « A Survey of Deep Learning : Platforms, Applications and Emerging Research Trends ». In : *IEEE Access* 6, p. 24411-24432. ISSN : 2169-3536. DOI : [10.1109/ACCESS.2018.2830661](https://doi.org/10.1109/ACCESS.2018.2830661).
- HAWARAH, Lamis (oct. 2008). « A Probabilistic Approach to Classify Incomplete Objects in a Decision Tree ». Theses. Université Joseph-Fourier - Grenoble I. URL : <https://tel.archives-ouvertes.fr/tel-00335313>.
- HONG MIN, Tae, Ying LI HAN et Rae-Hong PARK (mai 2015). « Efficient iris localisation using a guided filter ». In : *IET Image Processing* 9, p. 405-412.

- HU, M. (1962). « Visual Pattern Recognition by Moment Invariants ». In : *RE Transaction on Information Theory*, 8.2, p. 179-187.
- HU, Rong-Xiang et al. (2012). « Hand shape recognition based on coherent distance shape contexts ». In : *Pattern Recognition* 45.9. Best Papers of Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'2011), p. 3348 -3359. ISSN : 0031-3203. DOI : <https://doi.org/10.1016/j.patcog.2012.02.018>. URL : <http://www.sciencedirect.com/science/article/pii/S0031320312000969>.
- IANDOLA, Forrest N. et al. (2016). « SqueezeNet : AlexNet-level accuracy with 50x fewer parameters and <1MB model size ». In : *CoRR* abs/1602.07360.
- IANNARELLI, A.V. (1964). *The Iannarelli system of ear identification*. Police science series. Foundation Press. URL : <https://books.google.dz/books?id=d0ppAAAAIAAJ>.
- IITD Iris database. https://www4.comp.polyu.edu.hk/csajaykr/IITD/Database_Iris.htm. Accessed : 30-10-2017.
- INBAVALLI, P. et G. Naga NANDHINI (2014). « Body Odor as a Biometric Authentication ». In : *International Journal of Computer Science and Information Technologies* 5, p. 6270-627.
- JAIN, A. K. et N. DUTA (1999). « Deformable matching of hand shapes for user verification ». In : *Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348)*. T. 2, 857-861 vol.2. DOI : [10.1109/ICIP.1999.823019](https://doi.org/10.1109/ICIP.1999.823019).
- JAIN, Anil K., Arun ROSS et Salil PRABHAKAR (fév. 2004). « An Introduction to Biometric Recognition ». In : *Circuits and Systems for Video Technology, IEEE Transactions on* 14, p. 4 -20. DOI : [10.1109/TCSVT.2003.818349](https://doi.org/10.1109/TCSVT.2003.818349).
- JOHAR, Tania et Pooja KAUSHIK (2015). « Iris Segmentation and Normalization using Daugma's Rubber Sheet Model ». In : *International Journal of Scientific and Technical Advancements* 1, 11—14.
- JUAN MANUEL RAMIREZ-CORTES Pilar Gomez-Gil, Gabriel Sanchez-Perez et David BAEZ-LOPEZ (2008). « A Feature Extraction Method Based on the Pattern Spectrum for Hand Shape Biometry ». In : *Proceedings of the World Congress on Engineering et Computer Science*.
- KALYAN, Veeramachaneni, Osadciw Lisa ANN et Varshney Pramod K. (2003). « Adaptive Multimodal Biometric Fusion Algorithm Using Particle Swarm ». In : *Proceedings of SPIE, Aerosense. Paper 146*.
- KEREN, D. (1994). « Using symbolic computation to find algebraic invariants ». In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.11, p. 1143-1149. ISSN : 0162-8828. DOI : [10.1109/34.334397](https://doi.org/10.1109/34.334397).
- KHIARI, Nefissa (2016). « Biométrie multimodale basée sur l'iris et le visage ». Thèse de doctorat dirigée par Lelandais Bonadè, SylvieHamrouni, Kamel et Montagne, Christophe Traitement du signal et des images Paris Saclay 2016. Thèse de doct. URL : <http://www.theses.fr/2016SACLE014>.

- KIM, Wan, Jong MIN SONG et Kang RYOUNG PARK (juil. 2018). « Multimodal Biometric Recognition Based on Convolutional Neural Network by the Fusion of Finger-Vein and Finger Shape Using Near-Infrared (NIR) Camera Sensor ». In : *Sensors* 18, p. 2296. DOI : [10.3390/s18072296](https://doi.org/10.3390/s18072296).
- KISKU, Dakshina Ranjan, Phalguni GUPTA et Jamuna Kanta SING (2011). « Multibiometrics Feature Level Fusion by Graph Clustering ». In : *International Journal of Security and Its Applications* 5.2, p. -.
- KRENKER, Andrej, Janez BESTER et Andrej KOS (2011). « Introduction to the Artificial Neural Networks ». In : *Artificial Neural Networks*. Sous la dir. de Kenji SUZUKI. Rijeka : IntechOpen. Chap. 1. DOI : [10.5772/15751](https://doi.org/10.5772/15751). URL : <https://doi.org/10.5772/15751>.
- KRIZHEVSKY, Alex, Ilya SUTSKEVER et Geoffrey E. HINTON (jan. 2012). « ImageNet Classification with Deep Convolutional Neural Networks ». In : *Neural Information Processing Systems* 25. DOI : [10.1145/3065386](https://doi.org/10.1145/3065386).
- KURBAN, Onur, Tülay YILDIRIM et Ahmet BILGIC (juil. 2017). « A multi-biometric recognition system based on deep features of face and gesture energy image ». In : *IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, p. 361-364. DOI : [10.1109/INISTA.2017.8001186](https://doi.org/10.1109/INISTA.2017.8001186).
- LECUN, Y. et al. (1998). « Gradient-based learning applied to document recognition ». In : *Proceedings of the IEEE* 86.11, p. 2278-2324. ISSN : 0018-9219. DOI : [10.1109/5.726791](https://doi.org/10.1109/5.726791).
- LECUN, Yann, Y BENGIO et Geoffrey HINTON (mai 2015). « Deep Learning ». In : *Nature* 521, p. 436-44. DOI : [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- LI, Stan Z. et Anil JAIN (2015). *Encyclopedia of Biometrics*. Springer Publishing Company, Incorporated. ISBN : 148997489X, 9781489974891.
- LIU, Qingshan et al. (2002). « Kernel Eigenfaces vs. Kernel Fisherfaces : Face recognition using kernel methods ». In : *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, p. 215-220.
- LUO, Yue-Tong et al. (2016). « Local line directional pattern for palm-print recognition ». In : *Pattern Recognition* 50, p. 26 -44. ISSN : 0031-3203. DOI : <https://doi.org/10.1016/j.patcog.2015.08.025>. URL : <http://www.sciencedirect.com/science/article/pii/S0031320315003131>.
- LUQUE-BAENA, Rafael M. et al. (2013). « Assessment of geometric features for individual identification and verification in biometric hand systems ». In : *Expert Systems with Applications* 40.9, p. 3580 -3594. ISSN : 0957-4174. DOI : <https://doi.org/10.1016/j.eswa.2012.12.065>. URL : <http://www.sciencedirect.com/science/article/pii/S0957417412013061>.
- MAITRA, S. (1979). « Moment Invariants ». In : in Proc. of the IEEE, p. 697-699.
- MALTONI, Davide et al. (2009). *Handbook of Fingerprint Recognition*. 2nd. Springer Publishing Company, Incorporated. ISBN : 1848822537.
- MANSFIELD, A.J. et al. (2002). *Best Practices in Testing and Reporting Performance of Biometric Devices*. NPL report. Centre for Mathematics et

- Scientific Computing, National Physical Laboratory. URL : <https://books.google.dz/books?id=4aAyMwEACAAJ>.
- MARINI, Federico et Beata WALCZAK (2015). « Particle swarm optimization (PSO). A tutorial ». In : *Chemometrics and Intelligent Laboratory Systems* 149, p. 153 -165. ISSN : 0169-7439. DOI : <https://doi.org/10.1016/j.chemolab.2015.08.020>. URL : <http://www.sciencedirect.com/science/article/pii/S0169743915002117>.
- MARYAM, E. et T. ONSSEN (2012). « Fusion of face and iris biometrics using local and global feature extraction methods ». In : *Signal, Image and Video Processing*, p. 1-12.
- MATHIEU-DUPAS, Eve (2010). « Algorithme des k plus proches voisins pondérés et application en diagnostic ». In : *42èmes Journées de Statistique*. Marseille, France, France. URL : <https://hal.inria.fr/inria-00494814>.
- MAZOUNI, R. et A. RAHMOUN (2011). « On Comparing Verification Performances of Multimodal Biometrics Fusion Techniques ». In : *International Journal of Computer Applications* 33.7, 0975—8887.
- MESHGINI S., Aghagolzadeh A. Seyedarabi H. (2013). « Face recognition using gabor based direct linear discriminant analysis and support vector machine ». In : *Computers Electrical Engineering* 39, p. 727-745.
- MEZAI, L., F. HACHOUF et M. BENGHERABI (2011). « Fusion of face and voice using the Dempster-Shafer Theory for person verification ». In : *International Workshop on Systems, Signal Processing and their Applications, WOSSPA*, p. 103-106. DOI : [10.1109/WOSSPA.2011.5931424](https://doi.org/10.1109/WOSSPA.2011.5931424).
- MITCHELL, Melanie (1998). *An Introduction to Genetic Algorithms*. Cambridge, MA, USA : MIT Press. ISBN : 0262631857.
- MODAK, Sandip Kumar Singh et Vijay Kumar JHA (2017). « Feature level Fusion of Face and Hand for Multibiometric Based Personal Identification ». In : *International Journal of Computer Science and Mobile Computing* 6, p. 237-242.
- MORIZET, Nicolas (mar. 2009). « Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris ». Theses. Télécom ParisTech. URL : <https://pastel.archives-ouvertes.fr/pastel-00005811>.
- MUROFUSHI, Toshiaki et Michio SUGENO (1989). « An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure ». In : *Fuzzy Sets and Systems* 29.2, p. 201 -227. ISSN : 0165-0114. DOI : [https://doi.org/10.1016/0165-0114\(89\)90194-2](https://doi.org/10.1016/0165-0114(89)90194-2). URL : <http://www.sciencedirect.com/science/article/pii/0165011489901942>.
- NANDAKUMAR, K. (2008). *Multibiometric Systems : Fusion Strategies and Template Security*. AD-a488 585. Michigan State University. Department of Computer Science et Engineering. URL : <https://books.google.dz/books?id=odwPNwQXag4C>.
- NIRGUDE, Manisha A. et Sachin R. GENGAJE (2015). « Wavelet based Iris Recognition System ». In : *International Journal of Engineering Research and Technology (IJERT)* 04, p. 1015-1017.

- NORDIN, Md Jan et A HUSSEIN (mar. 2014). « Palmprint verification using invariant moments based on wavelet transform ». In : *Journal of Computer Science* 10, p. 1389-1396. DOI : [10.3844/jcssp.2014.1389.1396](https://doi.org/10.3844/jcssp.2014.1389.1396).
- ODEN, Cenker, Aytul ERCIL et Burak BUKE (2003). « Combining implicit polynomials and geometric features for hand recognition ». In : *Pattern Recognition Letters* 24.13. Audio- and Video-based Biometric Person Authentication (AVBPA 2001), p. 2145 -2152. ISSN : 0167-8655. DOI : [https://doi.org/10.1016/S0167-8655\(03\)00087-4](https://doi.org/10.1016/S0167-8655(03)00087-4). URL : <http://www.sciencedirect.com/science/article/pii/S0167865503000874>.
- OH, Sung-Kwun, Sung-Hoon YOO et Witold PEDRYCZ (2013). « Design of face recognition algorithm using PCA -LDA combined for hybrid data pre-processing and polynomial-based RBF neural networks : Design and its application ». In : *Expert Systems with Applications* 40.5, p. 1451 -1466. ISSN : 0957-4174. DOI : <https://doi.org/10.1016/j.eswa.2012.08.046>. URL : <http://www.sciencedirect.com/science/article/pii/S095741741201007X>.
- OLA M. ALY Gouda I. Salama, Tarek A. Mahmoud et Hoda M. ONSI (2013). « An Adaptive Multimodal Biometrics System using PSO ». In : *International Journal of Advanced Computer Science and Applications* 4.7, p. -.
- OMARA IBRAHIM Xiao Gang, Amrani Moussa Yan Zifei Zuo Wangmeng (2017). *Deep features for efficient multi-biometric recognition with face and ear images*. DOI : [10.1117/12.2281694](https://doi.org/10.1117/12.2281694). URL : <https://doi.org/10.1117/12.2281694>.
- ORL database. [http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_faces.zip](http://www.cl.cam.ac.uk/Research/DTG/attarchive/pub/data/att_faces.zip). Accessed : 30-10-2017.
- O'SHEA, Keiron et Ryan NASH (2015). « An Introduction to Convolutional Neural Networks ». In : *CoRR* 1511-08458. arXiv : [1511.08458](https://arxiv.org/abs/1511.08458). URL : <http://arxiv.org/abs/1511.08458>.
- OTSU, N. (1978). « A Threshold Selection Method From Gray-scale Histogram ». In : *IEEE Transaction Syst., Man, Cybern* 8, p. 62-66.
- PAN, S. J. et Q. YANG (2010). « A Survey on Transfer Learning ». In : *IEEE Transactions on Knowledge and Data Engineering* 22.10, p. 1345-1359. ISSN : 1041-4347. DOI : [10.1109/TKDE.2009.191](https://doi.org/10.1109/TKDE.2009.191).
- PARKAVI, R., K. R. C. BABU et J. A. KUMAR (2017). « Multimodal Biometrics for user authentication ». In : *2017 11th International Conference on Intelligent Systems and Control (ISCO)*, p. 501-505. DOI : [10.1109/ISCO.2017.7856044](https://doi.org/10.1109/ISCO.2017.7856044).
- PATIL, V.H. et Sampada DHOLE (2016). « An Efficient Secure Multimodal Biometric Fusion Using Palm print and Face Image ». In : *International Journal of Applied Engineering Research* 11.10, p. 7147-7150.
- PERRONNIN, Florent et Jean-Luc DUGELAY (déc. 2002). « Introduction à la biométrie authentification des individus par traitement audio-vidéo ». In : *Traitement du signal, Volume 19, N4, 2002*. DOI : <http://hdl.handle.net/2042/2211>. URL : <http://www.eurecom.fr/publication/1120>.

- PRAVEEN KUMAR NAYAK, Devesh Narayan (2013). « Multimodal Biometric Face and Fingerprint Recognition Using Adaptive principal Component Analysis and Multilayer Perception ». In : *International Journal of Research in Computer and Communication Technology* 2. ISSN : 2320-5156.
- RAJAGOPAL, G. et S. K. MANOHARAN (2015). « Personal Authentication Using Multifeatures Multispectral Palm Print Traits ». In : *The Scientific World Journal* 11, —. DOI : [10.1155/2015/861629](https://doi.org/10.1155/2015/861629).
- RESHMI, M.P et V.J.Arul KARTHICK (2013). « Biometric Identification System using Lips ». In : *International Journal of Science and Research (IJSR)* 2.
- RIBARIC, S. et I. FRATRIC (2005). « A biometric identification system based on eigenpalm and eigenfinger features ». In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27.11, p. 1698-1709. ISSN : 0162-8828. DOI : [10.1109/TPAMI.2005.209](https://doi.org/10.1109/TPAMI.2005.209).
- ROSS, A. et A. JAIN (2003). « Information fusion in biometrics ». In : *Pattern Recognition Letters* 24.13, p. 2115-2125.
- SAINI, Preeti et Parneet KAUR (2013). « Automatic speech recognition : A review ». In : *International journal of Engineering Trends & Technology*, p. 132-136. URL : <http://www.ijettjournal.org/volume-4/issue-2/IJETT-V4I2P210.pdf>.
- SAXENA, Nidhi et al. (2012). « HAND GEOMETRY : A New Method for Biometric Recognition ». In : *International Journal of Soft Computing and Engineering (IJSCE)* 2. ISSN : 2231-2307.
- SCHMIDHUBER, Jürgen (2015). « Deep learning in neural networks : An overview ». In : *Neural Networks* 61, p. 85 -117. ISSN : 0893-6080. DOI : <https://doi.org/10.1016/j.neunet.2014.09.003>. URL : <http://www.sciencedirect.com/science/article/pii/S0893608014002135>.
- SHAMS, Mahmoud, Shahenda SARHAN et Ahmad TOLBA (mai 2017). « Adaptive Deep Learning Vector Quantisation for Multimodal Authentication ». In : *Journal of Information Hiding and Multimedia Signal Processing* 8, p. 702-722.
- SHASHOA, Nasar Aldian Ambark et al. (2016). « Classification depend on linear discriminant analysis using desired outputs ». In : *2016 17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, p. 328-332. DOI : [10.1109/STA.2016.7952041](https://doi.org/10.1109/STA.2016.7952041).
- SHUANG XU, Jifeng Ding (2012). « Palmprint Image Processing and Linear Discriminant Analysis Method ». In : *JOURNAL OF MULTIMEDIA* 7.
- SHUKLA, Anupam, Ritu TIWARI et Rahul KALA (2010). « Towards Hybrid and Adaptive Computing - A Perspective ». In : *Studies in Computational Intelligence*.
- SIREESHA, V. et K. SANDHYARANI (2016). « Decision Level Fusion Based Multimodal Biometric System ». In : *International Journal of Advanced Research in Computer Engineering Technology (IJARCET)* 5.

- SOLEYMANI, Sobhan et al. (2018). « Multi-Level Feature Abstraction from Convolutional Neural Networks for Multimodal Biometric Identification ». In : *CoRR* abs/1807.01332. arXiv : 1807.01332. URL : <http://arxiv.org/abs/1807.01332>.
- SRIDHA, D. et I.V. Murali KRISHNA (2012). « Combined Classifier for Face Recognition using Legendre Moments ». In : *Computer Engineering and Applications* 1.1.
- SUGENO, M. (1993). « FUZZY MEASURES AND FUZZY INTEGRALS—A SURVEY ». In : *Readings in Fuzzy Sets for Intelligent Systems*. Sous la dir. de Didier DUBOIS, Henri PRADE et Ronald R. YAGER. Morgan Kaufmann, p. 251 -257. ISBN : 978-1-4832-1450-4. DOI : <https://doi.org/10.1016/B978-1-4832-1450-4.50027-4>. URL : <http://www.sciencedirect.com/science/article/pii/B9781483214504500274>.
- SZEGEDY, C. et al. (2015). « Going deeper with convolutions ». In : *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 1-9. DOI : [10.1109/CVPR.2015.7298594](https://doi.org/10.1109/CVPR.2015.7298594).
- TALREJA, Veeru, Matthew C. VALENTI et Nasser M. NASRABADI (2017). « Multibiometric Secure System Based on Deep Learning ». In : *CoRR* abs/1708.02314. arXiv : 1708.02314. URL : <http://arxiv.org/abs/1708.02314>.
- TOUZET, Claude (1992). *LES RESEAUX DE NEURONES ARTIFICIELS, INTRODUCTION AU CONNEXIONNISME*. Collection de l'EERIE. EC2. URL : <https://hal-amu.archives-ouvertes.fr/hal-01338010>.
- TURK, M. et A. PENTLAND (1991). « Eigenfaces for Face Recognition ». In : *Journal of Cognitive Neuroscience* 3, p. 71-86.
- VAPNIK, Vladimir N. (1995). *The Nature of Statistical Learning Theory*. Berlin, Heidelberg : Springer-Verlag. ISBN : 0-387-94559-8.
- VINAY, A. et al. (2015). « Performance Study of LDA and KFA for Gabor Based Face Recognition System ». In : *Procedia Computer Science* 57. 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015), p. 960 -969. ISSN : 1877-0509. DOI : <https://doi.org/10.1016/j.procs.2015.07.493>. URL : <http://www.sciencedirect.com/science/article/pii/S1877050915020220>.
- WANG, Fenghua et Jiuqiang HAN (jan. 2008). « Robust multimodal biometric authentication integrating iris, face and palmprint ». In : *Information technology and control* 37.
- WANG, Kunxia et al. (jan. 2015). « Speech Emotion Recognition Using Fourier Parameters ». In : *Affective Computing, IEEE Transactions on* 6, p. 69-75.
- WANG, Xuan, Li LEI et Mingzhe WANG (2012). « Palmprint verification based on 2D – Gabor wavelet and pulse-coupled neural network ». In : *Knowledge-Based Systems* 27, p. 451 -455. ISSN : 0950-7051. DOI : <https://doi.org/10.1016/j.knosys.2011.10.008>. URL : <http://www.sciencedirect.com/science/article/pii/S0950705111002280>.
- WEI SHU et D. ZHANG (1998). « Palmprint verification : an implementation of biometric technology ». In : *Proceedings. Fourteenth International*

- Conference on Pattern Recognition (Cat. No.98EX170)*. T. 1, p. 219-221. DOI : [10.1109/ICPR.1998.711120](https://doi.org/10.1109/ICPR.1998.711120).
- WU, X. et Q. ZHAO (2015). « Deformed Palmprint Matching Based on Stable Regions ». In : *IEEE Transactions on Image Processing* 24.12, p. 4978-4989. ISSN : 1057-7149. DOI : [10.1109/TIP.2015.2478386](https://doi.org/10.1109/TIP.2015.2478386).
- XIONG, Wei et al. (2005). « Model-guided deformable hand shape recognition without positioning aids ». In : *Pattern Recognition* 38.10, p. 1651-1664. ISSN : 0031-3203. DOI : <https://doi.org/10.1016/j.patcog.2004.07.008>. URL : <http://www.sciencedirect.com/science/article/pii/S0031320304003528>.
- YANG, Jucheng et al. (2018). « A Novel Multimodal Biometrics Recognition Model Based on Stacked ELM and CCA Methods ». In : *Symmetry* 10.4.
- YORUK, Erdem et al. (août 2006). « Shape-based hand recognition ». In : *Image Processing, IEEE Transactions on* 15, p. 1803 -1815. DOI : [10.1109/TIP.2006.873439](https://doi.org/10.1109/TIP.2006.873439).
- YOSINSKI, Jason et al. (2014). « How transferable are features in deep neural networks? » In : *CoRR* abs/1411.1792. arXiv : 1411 . 1792. URL : <http://arxiv.org/abs/1411.1792>.
- YÖRÜK, Erdem, Helin DUTAĞACI et Bülent SANKUR (2006). « Hand biometrics ». In : *Image and Vision Computing* 24.5, p. 483 -497. ISSN : 0262-8856. DOI : <https://doi.org/10.1016/j.imavis.2006.01.020>. URL : <http://www.sciencedirect.com/science/article/pii/S0262885606000497>.
- ZADEH, Lotfi A. (1965). « Fuzzy Sets ». In : *Information Control* 8, p. 338-353.
- ZHU, Le qing et San yuan ZHANG (2010). « Multimodal biometric identification system based on finger geometry, knuckle print and palm print ». In : *Pattern Recognition Letters* 31.12. *Pattern Recognition of Non-Speech Audio*, p. 1641 -1649. ISSN : 0167-8655. DOI : <https://doi.org/10.1016/j.patrec.2010.05.010>. URL : <http://www.sciencedirect.com/science/article/pii/S0167865510001509>.
- ZUO, W. et al. (2010). « The multiscale competitive code via sparse representation for palmprint verification ». In : *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 2265-2272. DOI : [10.1109/CVPR.2010.5539909](https://doi.org/10.1109/CVPR.2010.5539909).