

وزارة التعليم العالي و البحث العلمي

BADJI MOKHTAR UNIVERSITY- ANNABA  
UNIVERSITE BADJI MOKHTAR - ANNABA



جامعة باجي مختار - عنابة

FACULTE DES SCIENCES DE L'INGENIORAT  
DEPARTEMENT D'INFORMATIQUE

Année: 2022

## THESE

Présentée en vue de l'obtention du diplôme de DOCTORAT en sciences

Thème

### Identification biométrique par l'iris

Option : Intelligence Artificielle

Par :

**Tounsi Souheila**

**Directeur de Thèse:**

Mme Boukari Karima - MCA- Université de Badji Mokhtar, Annaba.

**Devant Jury:**

**Président:**

Mme Souici-Meslati Labiba - Prof- Université de Badji Mokhtar, Annaba.

**Examineurs:**

Mr Ouchtati Salim- Prof- à l'Université du 20 Aout 1955 de Skikda.

Mr Moussaoui Abedlekrim-Prof-à l'Université du 8 Mai 1945 de Guelma.

Mr Messaoudi Kamel -Prof- à l'université Mohamed Chérif Messaadia de Souk Ahras.



## *Remerciements*

Je remercie, avant tout, Dieu, le Tout-Puissant, de m'avoir accordé parmi Ses innombrables Grâces, santé et courage pour accomplir ce travail.

Je remercie, en premier lieu, Madame **Boukari Karima** Maître de conférence A à l'Université Badji Mokhtar Annaba et directeur de cette thèse, pour avoir assuré le suivi de cette thèse. Son expérience et son aide scientifique m'ont été essentielles, la qualité de ses conseils et ses encouragements et son soutien permanent tant au niveau des connaissances qu'au niveau humain, sa disponibilité ainsi que l'attention qu'il a portée à ce travail ont été un atout précieux dans l'avancement de cette étude. **Merci beaucoup Madame Boukari Karima**

Je tiens à remercier également Madame **Souici-Meslati Labiba** professeur à l'Université Badji Mokhtar Annaba pour l'honneur qu'elle m'a fait en présidant le jury de ma soutenance, qui a été une des meilleurs enseignants préférées et qui a contribué de façon considérable à ma formation d'ingénieur et de magister.

Je tiens, également, à remercier les membres du jury qui m'ont fait l'honneur d'avoir accepté d'accorder de leur temps pour examiner et bien vouloir évaluer mon travail, plus précisément :

Monsieur **Ouchtati Salim** professeur à l'Université du 20 Aout 1955 de Skikda.

Monsieur **Moussaoui Abdelkrim** professeur à l'Université du 8 Mai 1945 de Guelma.

Monsieur **Messaoudi Kamel** Professeur à l'université Mohamed Chérif Messaadia de Souk Ahras.

Je voudrais aussi profiter de cette occasion pour remercier tous ceux qui ont contribué, de près ou de loin, à l'élaboration de ce travail.

*Je dédie cet événement marquant de ma vie,*

*A la mémoire de mon père.*

*Disparu trop tôt. Tes conseils ont toujours guidé mes pas vers la réussite. Ta patience sans fin, ta compréhension et ton encouragement sont pour moi le soutien indispensable que tu as toujours su m'apporter. J'espère que, du monde qui est sien maintenant, il apprécie cet humble geste comme preuve de reconnaissance de la part d'une fille qui a toujours prié pour le salut de son âme. Puisse Dieu, le tout puissant, l'avoir en sa sainte miséricorde.*

*À ma très chère mère.*

*Autant de phrases aussi expressives soient-elles ne sauraient montrer le degré d'amour et d'affection que j'éprouve pour toi. Tu m'as comblé avec ta tendresse et affection tout au long de mon parcours. Tu n'as cessé de me soutenir et de m'encourager durant toutes les années de mes études, tu as toujours été présente à mes côtés pour me consoler quand il fallait. Puisse le tout puissant te donner santé, bonheur et longue vie afin que je puisse te combler à mon tour.*

*A mon mari.*

*Mon compagnon de route, Pour son affection, sa patience et son encouragement.*

*A mes enfants.*

*LIWAA ERRACHID et MOHAMED RAID le symbole de la joie, l'ambiance et la tendresse. Que Dieu les gardent et leurs ouvre les chemins de la réussite.*

*A mes frères et mes sœurs.*

*Qui sont toujours à mes côtés, prêts à m'aider.*

*A tous ceux qui me sont chers... Je dédie les fruits de ce modeste travail*

# Résumé

La biométrie par l'iris représente une des méthodes les plus fiables et les plus performantes à cause de son faible taux d'erreurs. Ces performances ont été observées dans des situations contrôlées, qui imposent des contraintes lors de l'acquisition pour l'obtention d'images de bonne qualité. Relâcher ces contraintes, au moins partiellement, implique des dégradations de la qualité des images acquises et par conséquent une réduction des performances de des systèmes de reconnaissance de l'iris, pour contourner le problème, il est possible d'exploiter l'information la moins bruitée de l'iris à savoir la collerette. Une des principales solutions proposées dans la littérature pour remédier à ces limites est d'améliorer l'étape de segmentation de l'iris. L'objectif principal de ce travail de thèse a été de proposer des méthodes originales pour la segmentation des images dégradées de l'iris.

Suite au grand succès des réseaux de neurones convolutifs (CNN) dans la classification et la reconnaissance d'images, nous avons utilisé une méthode d'apprentissage profond pour étudier le problème de la classification des images de l'iris plus précisément la collerette basse.

En dépit des difficultés à obtenir une image de bonne qualité, dans cette thèse une reconnaissance de l'iris à travers la collerette a été testée et validée sur la base CASIA V4, la caractérisation a été réalisée avec les réseaux de neurones convolutionnels sur un modèle pré-entraîné suivie d'une classification avec les machines à vecteurs de support.

**Mots clés :** Reconnaissance d'iris, Collerette zigzag, CNN, CASIA-Iris-Lamp V4, biométrie, SVM

# ملخص

تعد القياسات الحيوية للقرحية واحدة من أكثر الطرق موثوقة وفعالية بسبب معدل الخطأ المنخفض فيها. تمت ملاحظة هذه العروض في المواقف الخاضعة للرقابة ، والتي تفرض قيودًا أثناء الاستحواذ للحصول على صور جيدة الجودة. إن تخفيف هذه القيود، جزئيًا على الأقل، يعني تدهور جودة الصور المكتسبة وبالتالي انخفاض أداء أنظمة التعرف على قرحية العين للتحايل على المشكلة، من الممكن استغلال أقل قدر من المعلومات المزعجة في القرحية، وهي القلقة. أحد الحلول الرئيسي المقترحة في الأدبيات للتغلب على هذه القيود هي تحسين خطوة تجزئة القرحية. كان الهدف الرئيسي من هذه الأطروحة هو اقتراح طرق أصلية لتجزئة الصور المتدهورة للقرحية.

بعد النجاح الكبير الذي حققته الشبكات العصبية التلافيفية (CNN) في تصنيف الصور والتعرف عليها ، استخدمنا طريقة تعلم عميق لدراسة مشكلة تصنيف صورة قرحية العين ، وبشكل أدق الطوق السفلي.

على الرغم من الصعوبات في الحصول على صورة ذات جودة جيدة ، في هذه الأطروحة تم اختبار التعرف على القرحية من خلال الطوق والتحقق من صحته في قاعدة بيانات CASIA V4 ، تم إجراء التوصيف باستخدام الشبكات العصبية التلافيفية على نموذج مدرب مسبقًا متبوعًا بالتصنيف مع آلات ناقلات الدعم.

**الكلمات المفتاحية:** التعرف على قرحية العين ، طوق متعرج ، CNN ، CASIA-Iris-Lamp V4 ، البيومترية ، SVM .

# Abstract

Iris biometrics is one of the most reliable and efficient methods because of its low error rate. These performances have been observed in controlled situations, which impose constraints during acquisition to obtain good quality images. Relaxing these constraints, at least partially, implies degradations of the quality of the acquired images and consequently a reduction of the performances of iris recognition systems. To circumvent the problem, it is possible to exploit the least noisy information of the iris, namely the collar. One of the main solutions proposed in the literature to overcome these limitations is to improve the iris segmentation step. The main objective of this thesis work was to propose original methods for the segmentation of degraded iris images.

Following the great success of convolutional neural networks (CNN) in image classification and recognition, we used a deep learning method to study the problem of classification of iris images, more precisely the low collar.

Despite the difficulties in obtaining a good quality image, in this thesis recognition of the iris through the collar has been tested and validated on the CASIA V4 database, the characterization was performed with convolutional neural networks on a pre-trained model followed by a classification with support vector machines.

**Keywords:** Iris recognition, Zigzag collar, CNN, CASIA-Iris-Lamp V4, biometric, SVM

## Table des matières

Introduction générale	1
Chapitre 1: Biométrie de l'iris dans des images dégradées	4
1. Introduction	4
2. Généralité biométrique	4
3. Comparaison des différentes modalités biométriques	6
4. Systèmes biométriques	7
5. Système de reconnaissance de l'iris	9
5.1. Biologie de l'iris	9
5.2. Historique des systèmes de reconnaissance de l'iris	11
5.3. Module d'acquisition d'images d'iris	16
5.3.1. Acquisition d'images dans des conditions contrôlées	16
5.3.2. Acquisition d'image dans des conditions non contrôlées	17
5.4. Spécificités d'un système de reconnaissance d'iris	21
6. Colletette zigzag (colletette irienne) – Etat de l'art	23
7. Conclusion	26
Chapitre 2: Les techniques de reconnaissance de l'iris	28
1. Introduction	28
2. les techniques de prétraitement	28
2.1. Méthode de la transformée de Hough	30
2.2. Histogramme et méthode de filtrage	30
2.3. Opérations morphologiques	32
3. Techniques de segmentation de l'iris	33
3.1. Segmentation de l'iris à l'aide de techniques traditionnelles	33
3.1.1. Techniques de segmentation basées sur la transformée de Hough	34
3.1.2. Techniques de segmentation basées sur les histogrammes et les contours	35
3.1.3. Autres techniques de segmentation.	36
4. Techniques de normalisation	38
5. Technique d'extraction des caractéristiques	39
5.1. Extraction de caractéristiques à l'aide de techniques traditionnelles	40
5.2. Extraction de caractéristiques à l'aide de techniques d'apprentissage profond	41
6. Conclusion	44

## Table des matières

Le chapitre 3: Les algorithmes de classification	45
1. Introduction	45
2. Machine Learning (apprentissage automatique)	45
3. Développement de l'apprentissage automatique	47
4. Les applications d'apprentissage automatique	50
5. L'apprentissage automatique et ses applications	51
5.1. Arbres de décision	51
5.2. Régression linéaire	52
5.3. Bayes naïfs	53
5.4. Régression logistique	54
5.5. Machine à vecteur de support	55
6. L'apprentissage profond vs l'apprentissage automatique	57
7. Vue d'ensemble de l'apprentissage profond	58
8. Les Applications de l'apprentissage profond	58
9. l'apprentissage profond (Deep Learning)	60
9.1. Les réseaux neuronaux artificiels ANN	60
9.2. Réseau neuronal convolutif (CNN :Convolutional neural network)	61
9.2.1. Blocs de construction	62
9.2.1.1. Couche convolutive	62
9.2.1.2. Fonction d'activation	63
9.2.1.3. Couche de mise en commun ( pooling layer)	65
9.2.1.4. Couches entièrement connectées (ou couches denses)	66
9.3. Apprentissage d'un réseau neuronal CNN	67
9.3.1. Pré- apprentissage (Pre -training)	67
9.3.1.1. Collecte et répartition des données	67
9.3.2.2. Apprentissage par transfert	67
9.3.2. L'étape d'apprentissage	68
9.3.3. Post-apprentissage (Post-training)	70
9.3.3.1. Métriques pour l'évaluation des performances du réseau	70
10. Etat de l'art des techniques de classification	74
10.1. Classification à l'aide de techniques traditionnelles	74

## *Table des matières*

10.2. Classification à l'aide de techniques d'apprentissage profond	76
11. Conclusion	77
Chapitre 4: Le système proposé	78
1. Introduction	78
2. Acquisition de l'image	79
3. Prétraitement des images	80
3.1. Élimination des points blancs et filtrage	80
3.2. Localisation de l'iris (collarete zigzag)	81
3.3. Normalisation de l'iris (collarete zigzag)	85
3.4. Post-traitement par CLAHE	86
4. Extraction de caractéristiques	88
5. La classification	93
6. Résultats expérimentaux et analyse	94
6.1. Première expérimentation (détection de la région de l'iris)	94
6.2. Deuxième expérimentation (détection de la région de la collerette zigzag)	96
7. Conclusion	100
Conclusion générale	102

## *Liste des abréviations*

**ACP** : Analyse en Composantes Principales.

**AHE**: Adaptive Histogram Equalization.

**ANN**: Artificial Neural Network.

**AUC**: The Area Under the Curve.

**BPDFHE**: Brightness Preserving Dynamic Fuzzy Histogram Equalization.

**CLAHE**: Contrast Limited Adaptive Histogram Equalization.

**CNN**: Convolutional Neural Network.

**CHT**: Circular Hough Transform.

**DCT**: Discrete Cosine Transform.

**DWT**: Discret Wavelet Transform.

**Db2**: Daubechies 2.

**DL**: Deep Learning.

**DLDA**: Direct Linear Discriminant Analysis.

**ED**: Euclidean Distance.

**FWT**: Fast Wavelet Transform.

**HT**: Hough Transform.

**IA**: Intelligence Artificielle.

**K-NN**: K-Nearest Neighbor.

**LBP**: Local Binary Pattern.

**LVQ**: Learning Vector Quantization Classifier.

**ML**: Machine Learning.

**NIR**: Near Infrared Imaging.

**ROC**: Receiver Operating Characteristic.

**SVM**: Support Vector Machine.

**Figure 1.1** – modalités biométriques [3].

**Figure 1.2** – quelques applications biométriques [3].

**Figure 1.3** – Acquisition de données avec des capteurs biométriques [3].

**Figure 1.4** – Module de reconnaissance d'un système biométrique [3].

**Figure 1.5**– Anatomie et différentes caractéristiques de l'iris humain [7].

**Figure 1.6** – image de l'iris de bonne qualité acquise dans des conditions contrôlées [40]

**Figure 1.7** – des images d'iris dégradées acquise en NIR [40].

**Figure 1.8** – des images d'iris dégradées acquise en lumière visible [3].

**Figure 1.9-(a)**[44] et **(b)**[45] collerette zigzag.

**Figure 2.1**-la structure d'un système de reconnaissance de l'iris [60].

**Figure 3.1**-Exemple d'arbre de décision [172]

**Figure 3.2**-Représentation visuelle de la régression linéaire [174].

**Figure 3.3**- Représentation visuelle de la fonction logistique [23].

**Figure 3.4**- Modèle SVM linéaire. Deux classes (rouge contre bleu) ont été classées [180].

**Figure 3.5**-la naissance de l'apprentissage profond

**Figure 3.6**-Modèle simple de la façon dont les neurones artificiels sont connectés entre les différentes couches d'un réseau neuronal à action direct [187]

**Figure 3.7** - Les cartes résultantes après le regroupement max et average sur une carte de caractéristiques [187].

**Figure 3.8** – Analyse des courbe ROC [42]

**Figure 4.1** - Architecture générale de notre système reconnaissance d'iris développé.

**Figure 4.2**-Élimination des points blancs et filtrage.

**Figure 4.3**-Détection de la pupille

**Figure 4.4**-Isolement de la zone de collerette en zigzag (a), et le demi-cercle inférieur de la zone de collerette en zigzag (b).

**Figure 4.5** - le modèle de Daugman pour transférer la région de l'iris des coordonnées cartésiennes aux coordonnées polaires[245].

**Figure 4.6** - (a) normalisation de l'iris (b) amélioration de l'iris avec la méthode CLAHE

**Figure 4.7**-Normalisation de l'iris : demi-cercle inférieur de la région de la collerette.

**Figure 4.8**- Une illustration de l'architecture des réseaux neuronaux convolutifs. Les carrés gris mentionnent les cartes de caractéristiques et les carrés verts mentionnent le filtre de

convolution. Les lignes croisées entre les deux dernières couches mentionnent les neurones entièrement connectés [245].

**Figure 4.9-** GoogLeNet module d'inception [252]

**Figure 4.10-** Architecture détaillée d'Alex-Net [233].

**Figure 4.11.** Matrice de confusion pour le premier test  
(Collarete : demi-cercle inférieur)

**Figure 4.12-** Matrice de confusion pour le deuxième test  
(Collarete : cercle complet)

**Figure 4.13-** courbe ROC avec (collerette : cercle complet)

**Figure 4.14-** courbe ROC avec (collerette : demi-cercle inférieur)

**Tableau 1.1-** Comparaison entre les modalités biométriques, Performance haute (H), Moyenne (M), Faible (F) [4].

**Tableau 2.1-** Les types de bruit et les méthodes de prétraitement correspondantes [60]

**Tableau 3.1-** Développement de la LM [157]

**Tableau 3.2** - Tableau présentant une matrice de confusion, utilisée comme représentation visuelle de la performance d'un algorithme.

**Tableau 4.1** - GoogLeNet : l'architecture Inception [252].

**Tableau 4.2-** La couche d'Alex-Net [233]

**Tableau 4.3-** comparaison de taux de reconnaissance (accuracy), avec et sans post-traitement en utilisant les deux modèles AlexNet et GoogleNet.

**Tableau 4.4-** comparaison de taux de reconnaissance (accuracy), avec et sans prétraitement en utilisant les deux modèles AlexNet et GoogleNet

**Tableau 4.5** - Accuracy : collaquette cercle complet et collaquette demi-cercle inférieur.

**Tableau 4.6-** Paramètres de performance

**Tableau 1.1-** Comparaison entre les modalités biométriques [4]. Performance haute (H), Moyenne (M), Faible (F).

**Tableau 2.1-** Les types de bruit et les méthodes de prétraitement correspondantes.

**Tableau 3.1** - Tableau présentant une matrice de confusion, utilisée comme représentation visuelle de la performance d'un algorithme par exemple.

**Tableau 4.1** - GoogLeNet : l'architecture Inception.

**Tableau 4.2** - La couche d'Alex-Net [17]

**Tableau 4.3** - comparaison de taux de reconnaissance (accuracy), avec et sans post-traitement en utilisant les deux modèles AlexNet et GoogleNet.

**Tableau 4.4** - comparaison de taux de reconnaissance (accuracy), avec et sans prétraitement en utilisant les deux modèles AlexNet et GoogleNet

**Tableau 4.5** - Accuracy : collarete cercle complet et collarete demi-cercle inférieur.

**Tableau 4.6** - Paramètres de performance.

# Introduction générale

Les systèmes de vision par ordinateur évoluent rapidement et occupent une place importante dans notre vie quotidienne. L'expansion de ces systèmes s'explique par l'accroissement permanent des applications dans différents domaines tels que : la télédétection, la médecine, ou encore la sécurité et notamment en biométrie, qui correspond au sujet de ce travail de thèse. Parmi les diverses modalités biométriques, l'iris lorsqu'il est acquis en conditions contrôlées, présente d'excellentes performances, mais ceci au prix de contraintes fortes imposées aux utilisateurs. Or certaines applications telles que la vidéosurveillance ou le contrôle d'accès dans les aéroports, exigent le relâchement de ces contraintes. Par exemple l'acquisition peut se faire à grande distance et en mouvement. La reconnaissance par l'iris est rendue difficile dans ces conditions à cause des différentes dégradations introduites sur l'image de l'œil. Ces dégradations ont des effets néfastes sur la précision de la segmentation de l'iris dans l'image, première étape de tout processus de reconnaissance qui a un impact très important sur les performances d'identification. Dans ce cadre complexe, il s'agit de concevoir un système capable de segmenter les images dégradées de l'iris en limitant au maximum le nombre d'erreurs possibles.

L'identification de personnes par l'iris pose plusieurs problèmes, parmi eux, nous pouvons citer, comment améliorer les performances de la reconnaissance de l'iris ? Comment localiser l'iris dans une image d'iris ? Dans ce travail de thèse, nous avons opté pour l'utilisation d'un algorithme traitant la zone la moins bruitée de l'iris qui est la collerette afin d'obtenir une meilleure performance.

Ces dernières années, l'apprentissage profond a été largement étudié et a obtenu des résultats remarquables dans de nombreux domaines de la vision par ordinateur, il s'est avéré plus performant que les algorithmes de Machine Learning classiques, face à divers problèmes d'intelligence artificielle. En effet, le principal avantage des algorithmes d'apprentissage en profondeur est qu'ils essaient d'apprendre progressivement des caractéristiques de haut niveau à partir des données, de sorte que les données peuvent être transmises directement au réseau pour obtenir un bon score dès le départ. Cela élimine complètement l'étape d'extraction manuelle des caractéristiques dans le processus d'apprentissage.

Les méthodes classiques pour la reconnaissance suivent un procédé typique consistant en deux étapes : l'extraction des caractéristiques d'image de l'iris en question, et ensuite, la classification à partir de ces descripteurs pour obtenir le résultat final. Dans ce schéma, l'étape d'extraction des caractéristiques a une importance primordiale car elle définit quelles informations sur l'iris en question sont données à l'entrée de la technique de classification, en employant les réseaux de neurones convolutifs CNNs (Convolutional Neural Network). Plus particulièrement, notre objectif est de trouver des méthodes optimales pour la conception, l'apprentissage de CNNs et la reconnaissance de l'iris. Un autre objectif majeur de ce travail est d'étudier l'utilité du transfert dans la l'apprentissage de réseaux de neurones convolutifs pour des tâches de reconnaissance. Nous proposons une nouvelle approche de reconnaissance de la collerette basée sur une architecture spécifique de CNN et la classification par les supports vecteurs machines (SVMs).

Le manuscrit est organisé sur quatre chapitres :

Le premier chapitre est consacré à la présentation générale de la biométrie. Il décrit tout d'abord le principe de fonctionnement des systèmes biométriques, puis définit les outils utilisés pour évaluer leurs performances. Nous abordons aussi la place de la reconnaissance de l'iris parmi les autres techniques biométriques. A travers ce chapitre, nous voulons positionner le problème de la reconnaissance de l'ris et présenter ses enjeux et intérêts par rapport aux autres techniques. Nous étudions également les principales composantes d'un système de reconnaissance, notamment la détection, l'extraction de caractéristiques et la

phase de reconnaissance. Enfin, nous mettons en lumière les difficultés rencontrées par les systèmes de reconnaissance de l'iris.

- Dans le second chapitre une analyse détaillée des différentes techniques développées au cours de ces dernières années dans les domaines de la segmentation de l'iris est présentée. Notre objectif est de classer ces méthodes de manière efficace afin de mettre en évidence les particularités ainsi que les avantages et les inconvénients de chacune d'elles.

Le troisième chapitre dresse un état de l'art des techniques de classification utilisées pour la reconnaissance de l'iris. Nous insistons notamment sur l'apprentissage profond et les réseaux de neurones convolutifs. Il est dédié à notre approche d'extraction des caractéristiques de l'iris et de classification basée sur les réseaux de neurones convolutifs.

Dans le quatrième chapitre, nous présentons les résultats expérimentaux obtenus par la méthode d'identification hybride dite CNN-SVM en analysant leurs performances, suivies d'une discussion avec interprétation des résultats.

Enfin, la conclusion générale résumera les résultats obtenus par les différentes approches et donnera quelques perspectives sur les travaux futurs.

# Chapitre 1

## Biométrie de l'iris dans des images dégradées

### 1. Introduction

Face à l'évolution logique des réglementations internationales, face aux menaces du vol d'identité, à la fraude documentaire au terrorisme ou de la cybercriminalité, de nouvelles solutions technologiques sont petit à petit mises en œuvre.

Parmi ces technologies, la biométrie devient rapidement la technologie la plus pertinente pour l'identification et l'authentification fiables et rapides des personnes basées sur une biométrie unique, les mots de passe traditionnels ont longtemps été le point faible des systèmes de sécurité. La biométrie vise à résoudre ce problème en associant l'identification à notre corps et à nos comportements.

### 2. Généralité biométrique

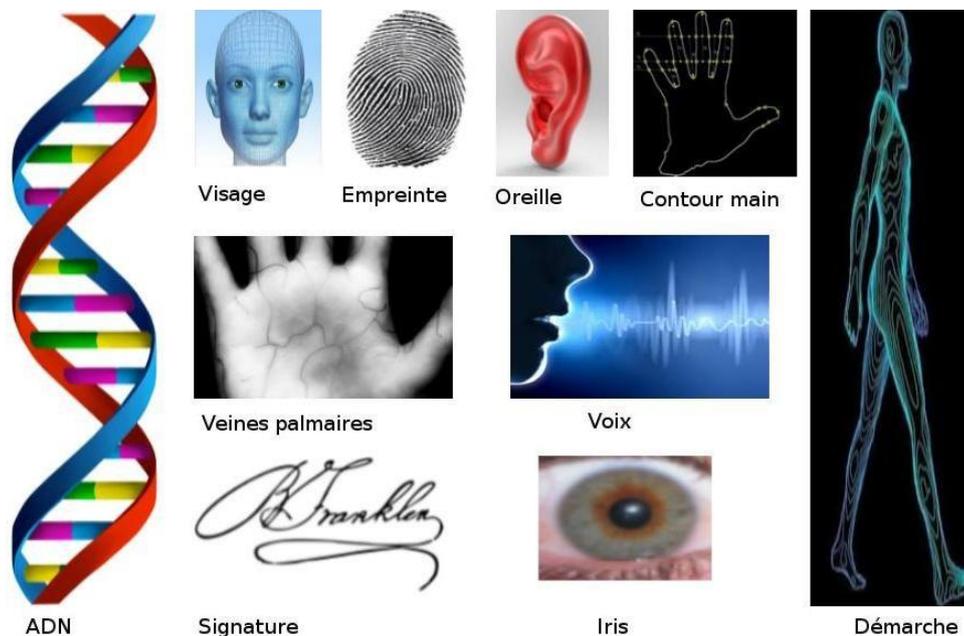
La technologie biométrique vise à établir une identification automatique basée sur des caractéristiques morphologiques humaines, telles que : l'iris, la rétine, l'empreinte digitale, le visage, le réseau veineux, la forme de la main, etc. Ou biologiques citant comme exemple : le sang, l'ADN, les cheveux, la salive, l'urine, etc. même comportemental, la dynamique de la signature, comme la voix, l'écriture et la frappe, etc. Tous ces attributs humains représentent des caractéristiques propres à chaque personne [1], [2].

Pour que ces attributs humains (biologiques, morphologiques ou comportementaux) soient considérés comme biométriques, ils doivent répondre aux différents critères suivants [1], [2] :

- **Unicité** : la modalité doit être unique pour chaque personne.
- **Universalité** : chaque personne doit avoir la modalité.
- **Permanence** : la modalité doit être suffisamment stable et ne doit pas être changée dans le temps.

- **Collectabilité** : la modalité doit être facile à mesurer (numériser) et à collecter dans tous les cas. Attention, la collecte doit être confidentielle.
- **Performance** : la modalité doit garantir une bonne précision de reconnaissance.
- **Acceptabilité** : En ce qui concerne les protocoles d'acquisition, les utilisateurs cibles doivent accepter leur biométrie dans leur vie quotidienne. .
- **Contournement** : la modalité doit être idéalement non falsifiable, pour garantir la sécurité du système biométrique [3].

La **figure 1.1** montre quelques exemples de différentes modalités biométriques. Il est important de préciser que différentes méthodes biométriques ne répondent pas entièrement aux différents critères décrits ci-dessus. Ces critères permettront de sélectionner le mode le plus avantageux en fonction de l'application visée.



**Figure 1.1** – modalités biométriques [3].

Avec la biométrie, une énorme révolution technologique est en cours. Les industriels et les chercheurs ont proposé aux utilisateurs plusieurs applications (**figure 1.2**), principalement dans trois domaines :

- **Gouvernemental**: de le but de contrôler les passages frontaliers et les documents d'identité, etc.
- **Légal**: comprend tous les problèmes liés à la justice (identification des criminels, des corps, , etc ).

- **Commercial** : permet la sécurisation des ordinateurs et des smartphones (accès aux informations personnelles, applications, identifiants, etc.), sites sensibles (bâtiments , laboratoires, industries, , etc), des paiements en ligne (e-commerce), des transactions bancaires et [3] .



Identification des crimes



Pièce d'identité



Transit international



Accès ordinateur



Accès smartphone



Accès physique



Transaction bancaire



Paiement en ligne

**Figure 1.2** – quelques applications biométriques [3].

### 3. Comparaison des différentes modalités biométriques

Le choix d'un système biométrique est étroitement lié à la qualité des informations fournies par le motif, qui dépend d'une part de sa nature et du niveau de sécurité qu'il apporte aux applications et du contexte dans lequel elles sont utilisées. Comme mentionné précédemment, les systèmes biométriques doivent répondre à un ensemble de critères tels que l'universalité, l'unicité, etc. Sur la base de ces critères, une comparaison des principales technologies biométriques est citée dans le **tableau 1.1**. Le tableau montre qu'aucune modalité biométrique ne répond à tous les critères et que chaque méthode présente des avantages et des inconvénients [4].

**Tableau 1.1**-Comparaison entre les modalités biométriques, Performance haute (H), Moyenne (M), Faible (F) [4].

Modalité biométrique	Universalité	Unicité	Stabilité	Mesurabilité	Performance	Acceptabilité	Circonvension
ADN	H	H	H	F	H	F	F
Oreille	M	M	H	M	M	H	M
Visage	H	F	M	H	F	H	H
Thermo Visage	H	H	F	H	M	H	H
Empreinte digitale	M	H	H	M	H	M	M
Marche	M	F	F	H	F	H	M
Géométrie de la main	M	M	M	H	M	M	M
Veines de la main	M	M	M	M	M	M	F
Iris	H	H	H	M	H	F	F
Dynamique de frappe	F	F	F	M	F	M	M
Odeur	H	H	H	F	F	M	F
Rétine	H	H	M	F	H	F	F
Signature	F	F	F	H	F	H	H
Voix	M	F	F	M	F	H	H

#### 4. Systèmes biométriques

Avec l'augmentation de la population mondiale, la sécurité de la vie humaine est devenue la priorité absolue de la plupart des pays développés. L'utilisation de systèmes d'identification biométrique automatique peut être considérée comme une solution fiable à divers problèmes liés à des domaines tels que la justice, les transports internationaux et les transactions bancaires.

Tous les systèmes biométriques, quelle que soit leur utilisation, comportent deux phases principales : la phase d'enregistrement et la phase de test ou de reconnaissance. La phase d'enregistrement consiste à acquérir des données de capteurs spécifiques (**figure 1.3**), permettant de convertir des modèles réels en

données numérisées sous forme d'images, de signaux, etc. Et enregistrez ces données numérisées sur un support physique. L'étape d'identification (**figure 1.4**) permet de traiter ces données et d'en extraire les traits discriminants et de générer des gabarits qui seront utilisés dans l'étape de comparaison. Ce dernier consiste à comparer le modèle avec différents modèles de référence enregistrés dans la base de données lors du processus d'enregistrement. Selon l'application visée, le système biométrique fonctionne soit en mode vérification (authentification) soit en mode identification [2].

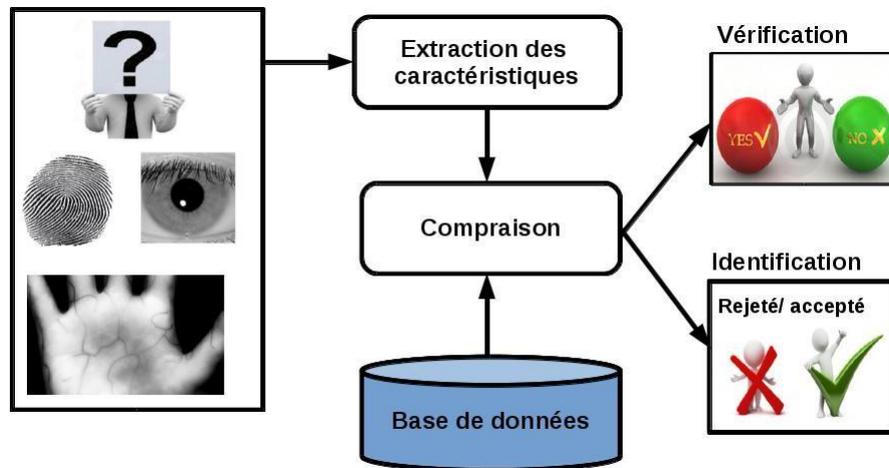


**Figure 1.3** – Acquisition de données avec des capteurs biométriques [3].

Le mode **de vérification** est conçu pour faire correspondre un modèle de test à l'un des modèles contenus dans la base de données et tenter de répondre à la question : Êtes-vous celui que vous prétendez être ? Dans cette opération, le système biométrique effectue un type de comparaison un à un dans lequel un modèle de test est comparé à un modèle de référence. La correspondance entre les modèles est quantifiée par le score de similarité (ou de dissemblance). Ce score de comparaison est comparé à un seuil prédéfini pour prendre une décision d'acceptation ou de rejet.

Le mode **d'identification** est conçu pour identifier l'identité des modèles de test afin de répondre aux questions suivantes : Qui êtes-vous ? Le système biométrique utilise une comparaison un à N (N est le nombre de modèles stockés), où le modèle de test est comparé à tous les modèles stockés de toutes les personnes. Dans ce cas, le

système fournit une liste des personnes les plus probables [3].



**Figure 1.4** – Module de reconnaissance d'un système biométrique [3].

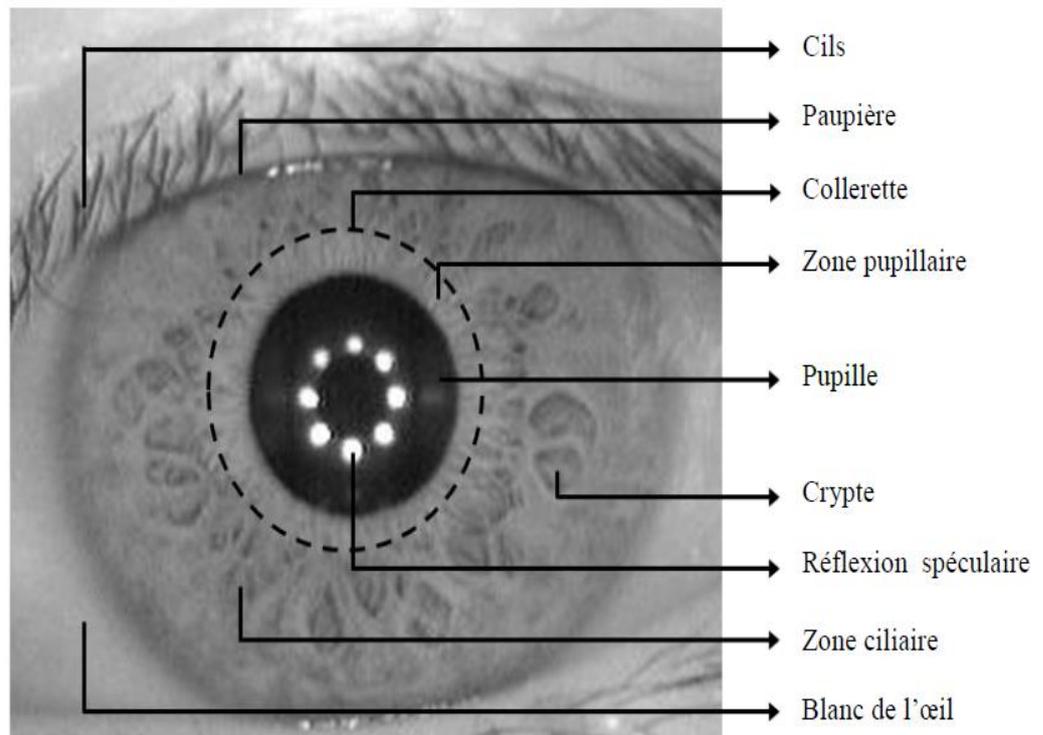
## 5. Système de reconnaissance de l'iris

### 5.1. Biologie de l'iris

L'existence d'un environnement mondial où les exigences de sécurité deviennent de plus en plus importantes explique le besoin continu de systèmes d'identifications avec des performances, une sécurité, une commodité et une facilité d'utilisation supérieures. L'identification d'une personne à partir d'une image d'iris est considérée comme l'une des meilleures techniques biométriques. Ceci est essentiellement lié aux caractéristiques et propriétés de l'iris [5]

L'iris, la partie colorée de l'œil humain, est un mince diaphragme circulaire resserré qui entoure la pupille, dont l'extrémité est entourée par le blanc de l'œil (sclérotique). Il est situé entre la cornée et le cristallin de l'œil humain, ce qui en fait un organe interne visible et bien protégé de l'extérieur. Son diamètre varie généralement entre 12 et 13 mm. Son rôle est de contrôler la quantité de lumière entrant dans la pupille. L'anatomie de l'image de l'iris ainsi que les différentes caractéristiques observables sont présentées à la **figure 1.5** [6].

L'iris est composé de plusieurs couches, la plus basse étant la couche de l'épithélium, qui contient des cellules pigmentaires denses, la mélanine, responsable de la couleur de l'iris. La couche stroma se situe au-dessus de la couche épithéliale et contient des vaisseaux sanguins, des cellules pigmentaires et les deux muscles de l'iris.



**Figure 1.5-**Anatomie et différentes caractéristiques de l'iris humain [7].

La surface externe multicouche de l'iris se compose de deux régions, généralement de couleurs différentes, la région pupillaire centrale et la région ciliaire externe. La collerette est la frontière séparant les deux régions, selon un motif en zigzag [8], [6], [9].

La formation de l'iris commence au 3ème mois de grossesse, et en plus de la pigmentation, le développement de l'iris et de sa texture visible s'achève au 8ème mois de grossesse. La pigmentation de stroma évolue jusqu'à l'âge de 12 mois, Alors restez stable pour la vie. La formation de motifs d'iris uniques est un phénomène aléatoire, indépendant des facteurs génétiques. La seule caractéristique qui dépend de l'hérédité est la pigmentation de l'iris qui détermine sa couleur. En raison des propriétés épigénétiques des motifs d'iris, même chez des jumeaux identiques, les yeux des individus contiennent des motifs d'iris complètement indépendants [8], [10], [11] ce qui conduit à une faible possibilité de similitude ( $\simeq 10^{-72}$ ) entre deux individus (même entre jumeaux) [12].

L'iris humain a une structure extraordinaire et offre une combinaison de nombreuses caractéristiques, telles que des cryptes, des arêtes, des arcs de ligaments, des silons,

des fibres, des filaments et des collerettes zigzag. Ces caractéristiques visibles, communément appelées la texture de l'iris, sont propres à chaque sujet. Les différences individuelles et la complexité physiologique de ces textures, telles que leur emplacement, Leur forme et l'intersection entre eux rendent la texture de l'iris unique et l'une des caractéristiques les plus abondantes du corps humain. [6], [13], [11]. Peu de particularités ou de maladies affectent l'apparence et la texture de l'iris. Cependant, dans ces cas, nous citons des cas d'albinisme dans lesquels l'iris est rouge en raison d'un manque de mélanine. Une autre maladie rare est l'absence complète de l'iris. De plus, la couleur de l'iris peut changer à mesure que la quantité de pigmentation de l'iris augmente pendant l'enfance. Cependant, l'apparence de l'iris est relativement stable pendant la majeure partie de la vie humaine [6], [9].

Enfin, nous pouvons en déduire que les caractéristiques de texture de l'iris sont stables dans le temps, avec seulement des changements mineurs. Ces caractéristiques sont uniques et universelles, de plus, l'iris est un organe interne et est protégé de l'environnement, c'est un organe visible de l'extérieur, qui permet de mesurer ou d'acquérir des images de la texture de l'iris de manière non invasive. Toutes ces caractéristiques de la texture de l'iris font de l'iris une modalité biométrique très importante pour l'identification et la reconnaissance humaines [14], [10].

## **5.2. Historique des systèmes de reconnaissance de l'iris**

L'iris est introduit pour la première fois comme un concept de reconnaissance de personnes en 1886 par Alphonse Bertillon qui ne s'intéressa cependant qu'à la couleur des yeux pour prendre une décision d'identification [15]. Par la suite l'ophtalmologiste Dr Frank Burch, qui a utilisé des modèles d'iris pour l'identification personnelle en 1936. Deux autres ophtalmologistes, Aran Safir et Leonard Flom Safir - qui considèrent que la texture de l'iris est un bon motif d'identification de personnes- [16], ont breveté cette idée en 1987 et ont demandé à John Daugman de créer des algorithmes applicables à la reconnaissance de l'iris en 1989 [17]. En 1994, John Daugman a concrétisé cette idée et a proposé un système automatique de reconnaissance de l'iris [18]. Il a détaillé les différents modules du système dans son travail [12].

Depuis lors, de nombreux travaux ont été réalisés dans le domaine de la reconnaissance de l'iris et, en général, les algorithmes de reconnaissance nécessitent une combinaison de diverses techniques. Plusieurs articles ont été publiés sur les systèmes, les techniques et les applications de la reconnaissance de l'iris au cours des

deux dernières décennies. Dans cette thèse, nous présentons un historique de toutes les étapes principales de la technique de reconnaissance de l'iris disponible dans la littérature en commençant par le travail initial de Daugman 1993 jusqu'à certains travaux récents.

**(Daugman, 1993)** a proposé la première méthodologie de travail sur la reconnaissance de l'iris, qui constitue la base de la plupart des activités de développement de la biométrie de l'iris à ce jour. Daugman a appliqué un opérateur intégral-différentiel pour localiser la région circulaire de l'iris et de la pupille, ainsi que les arcs des paupières supérieures et inférieures. L'opérateur recherche le chemin circulaire où la variation des valeurs des pixels est maximale, en faisant varier le rayon et la position  $x$  et  $y$  du centre du contour circulaire. L'opérateur est appliqué de manière itérative en réduisant progressivement le degré de lissage afin d'obtenir une localisation précise. Les paupières sont localisées de manière parallèle [12].

**(Boles et Boashash, 1993)** ont proposé un algorithme qui localise le centre de la pupille en utilisant une méthode de détection des bords, il enregistre les valeurs de niveau de gris sur des cercles concentriques virtuels, puis construit la représentation du passage à zéro sur ces cercles virtuels en se basant sur une transformée en ondelettes dyadique unidimensionnelle [19].

**(Wilde, 1997)** a introduit une nouvelle approche. Les limites intérieures et extérieures de l'iris sont calculées à l'aide d'une carte de bord binaire basée sur le gradient, suivie d'une transformation de Hough circulaire [20].

**(Boles et al., 1998)** ont appliqué la technique de détection des bords circulaires pour segmenter une image de l'iris et effectuer une analyse des cercles virtuels pour obtenir une signature normalisée de l'iris. Cette approche fournit un algorithme invariant en termes d'échelle, de rotation et de translation pour la reconnaissance de l'iris d'un œil humain [21].

**(Kong et Zhang, 2001)** ont développé un système qui se concentre principalement sur les perturbations du bruit, l'occlusion des cils et les réflexions spéculaires. La transformée de Hough a été utilisée pour isoler un iris, suivie de l'application de filtres de Gabor 1-d dans le domaine spatial et d'une fonction de seuil pour détecter l'occlusion des paupières et la réflexion spéculaire respectivement [22].

(**Daouk, 2002**) a proposé un schéma de reconnaissance de l'iris qui implique un mécanisme de fusion entre un schéma de détection des bords de Canny et une transformée de Hough circulaire. La limite de ce travail est que cette méthodologie n'est pas performante en cas de mauvais éclairage, d'occlusion par les paupières, de bruits ou de positionnement inapproprié des yeux [23].

(**Noh et al., 2003**) ont introduit une nouvelle technique d'extraction de caractéristiques où, au lieu d'utiliser la transformation en ondelettes, une méthode adaptative d'extraction de caractéristiques a été introduite dans laquelle deux types de caractéristiques globales et locales ont été extraites du coefficient d'ondelettes [24].

(**Daugman, 2004**) a proposé une approche qui constitue une amélioration de ses travaux antérieurs. En plus de l'opérateur Integro-différentiel décrit précédemment, un algorithme a été introduit pour détecter les paupières, qui impliquent des contours arqués avec un paramètre spline, au lieu des bords circulaires dans l'opérateur Integro différentiel [25].

(**Dorairaj et al., 2005**) ont développé un algorithme pour traiter les images d'iris hors axe en utilisant la technique de codage d'image PCA et ICA globale. Au cours de l'étape de prétraitement, ils ont utilisé la distance de Hamming pour calculer l'angle d'observation, puis l'opérateur différentiel intégral de Daugman pour segmenter l'iris [19].

(**Proenca et al., 2006**) ont introduit un nouvel et meilleur algorithme de segmentation pour les images d'iris en extrayant d'abord les caractéristiques de l'iris, puis en appliquant un algorithme de regroupement flou. Les résultats de cette expérience ont été comparés à ceux des techniques de segmentation de Daugman, Camus et Wilde, Wilde, Martin-Roche et al. et Tuceryan et se sont révélés plus performants [26].

(**Daugman, 2007**) a proposé un nouveau algorithme de traitement d'image à utiliser pendant l'étape de segmentation qui permet de traiter les images d'iris hors axe. La localisation de l'iris se fait en trois étapes (i) la segmentation de l'iris dans son ensemble, (ii) l'estimation du regard pour les images des yeux hors axe et (iii) l'exclusion des cils de la paupière supérieure [27].

(**Abiyev et al., 2008**) ont simulé un système de reconnaissance de l'iris en

utilisant des réseaux de neurones. La région de la pupille est détectée à l'aide d'une technique de zone rectangulaire 10x10 et aide à détecter le cercle intérieur de l'iris. La transformée de Hough linéaire est utilisée pour supprimer l'effet des paupières et une technique de seuillage est utilisée pour supprimer les cils [28].

(**Azizi et al., 2009**) ont proposé un travail qui traite de l'extraction de caractéristiques et de la sélection de sous-ensembles. Les caractéristiques de l'iris ont été extraites à l'aide de la transformée en contourlet. L'image de l'iris a ensuite été décomposée en sous-blocs qui contiennent toutes les informations de texture [29].

(**Narote et al., 2010**) ont évalué la performance de plusieurs ondelettes mères dans l'extraction des caractéristiques d'une image d'iris et ont déterminé une transformée en ondelettes optimale. Une image normalisée est décomposée en utilisant la décomposition de cinquième niveau et les caractéristiques sont extraites en utilisant différentes ondelettes mères : Haar, Daubechies, Coiflet, Symlet et Biorthogonale [30].

(**Hussain, 2010**) a proposé une technique pour extraire les caractéristiques des codes d'iris rectangulaires dans le domaine de l'espace propre. Différents nombres de vecteurs d'iris propres, comme 10, 7 et 4 ont été considérés pour évaluer la performance du système pour les codes d'iris avec et sans bruit [31].

(**Ali Alheeti, 2011**) a proposé une technique de reconnaissance de l'iris basée sur une technologie hybride, les transformées en ondelettes discrètes 2d avec des masques d'ondelettes comme les masques de transformées en ondelettes Haar et Db2 (Db2 : Daubechies 2) sont décomposées, suivies de l'application d'opérateurs de détection des contours comme canny, Prewitt, Roberts et Sobel pour reconnaître les caractéristiques [32].

(**Rashad et al., 2011**) ont proposé une approche de modèle statistique appelée modèle binaire local (LBP : Local Binary Pattern) avec des propriétés d'histogramme pour extraire les informations de texture de l'iris, puis pour concevoir un vecteur de caractéristiques. Cette caractéristique est introduite dans un classifieur basé sur un réseau de neurones appelé LVQ combiné (Combined Learning Vector Quantization Classifier) après avoir effectué une étude comparative [33].

(**Sathish, 2012**) a proposé un système de reconnaissance d'iris dans lequel

l'iris est segmenté en effectuant les étapes suivantes : une fonction de lissage gaussienne puis une égalisation d'histogramme sont appliquées pour améliorer le contraste de l'image de l'iris. Un détecteur de bord Canny suivi d'une transformation de Hough circulaire probabiliste est ensuite utilisé pour segmenter l'iris [34].

(**Jan, 2014**) a introduit un cadre de segmentation de l'iris à plusieurs étapes pour la localisation de l'iris, sans utiliser des approches traditionnelles, ils ont utilisé une approche itérative. En découplant la détection grossière du centre et la localisation fine des limites, il augmente la vitesse de traitement et la conception modulaire peut être réalisée. Cela diminue le temps de réponse pendant le processus de segmentation [35].

(**Abikoye et al., 2014**) ont utilisé Fast Wavelet Transform (FWT) pour extraire les caractéristiques de l'iris, les caractéristiques significatives de l'iris ont été codées et comparées entre les modèles, les caractéristiques codées ont été utilisées pour générer ses codes de caractéristiques de l'iris. L'algorithme est rapide et a un faible taux de complexité [36].

(**Oyeniran et al., 2019**) ont proposé une technique multi-algorithmique pour la reconnaissance des personnes à l'aide de l'iris en utilisant une approche de classificateurs multiples. Ils ont appliqué la transformée circulaire de Hough pour les techniques de localisation et de segmentation afin d'isoler un iris de l'image de l'œil entière et pour la détection du bruit [37].

(**Khanam et al., 2019**) ont mis en œuvre un réseau de neurones et une analyse discriminante de la méthode d'apprentissage automatique pour la reconnaissance de l'iris. La méthode proposée donne un meilleur taux de reconnaissance que la technique SVM avec une complexité de calcul moindre. Les méthodes de réseau de neurones et de l'analyse discriminante sont utilisées pour l'appariement et la recherche de la précision de reconnaissance [38].

(**Oyeniya et al., 2020**) ont proposé une extraction améliorée des caractéristiques de l'iris en utilisant la transformée en ondelettes continue. Cette méthode réduit considérablement le temps de calcul et améliore la précision par rapport au filtre de Gabor, à la transformée de Fourier et aux autres transformées en ondelettes [39].

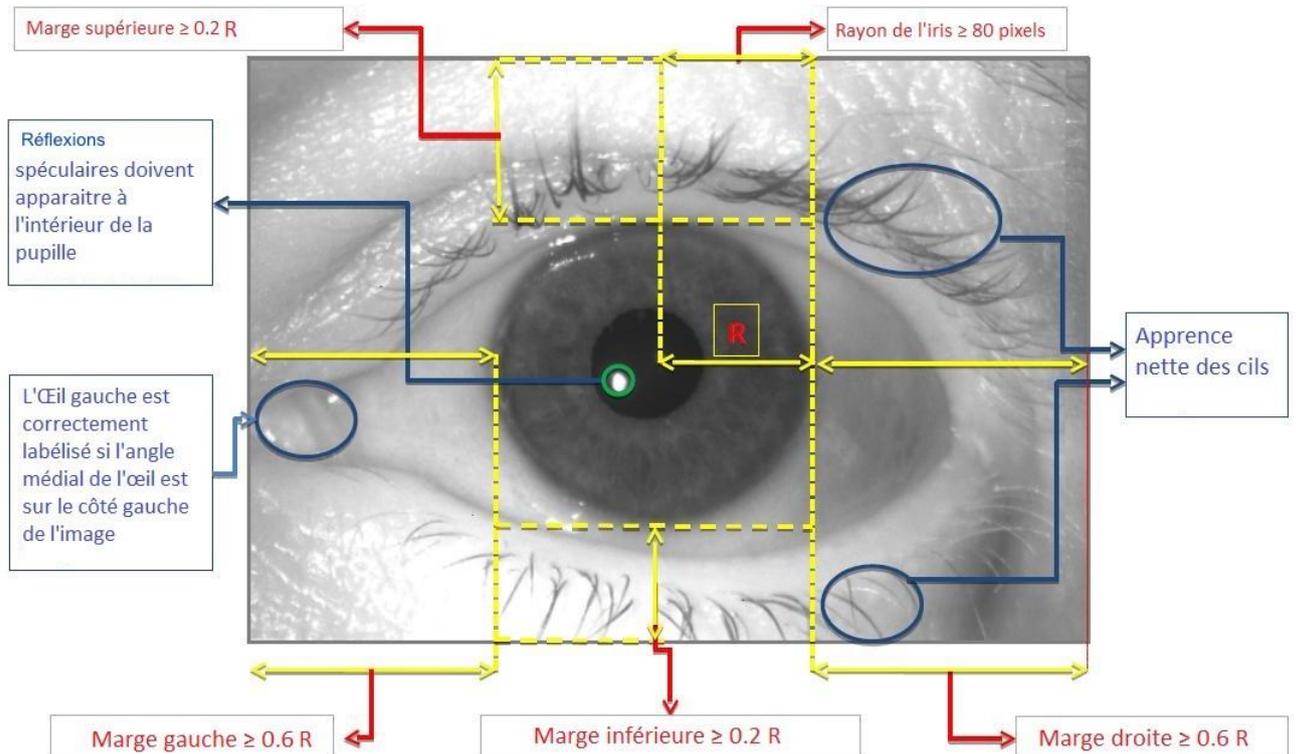
### 5.3. Module d'acquisition d'images d'iris

D'après la littérature et l'expérience, des images d'iris de haute qualité sont nécessaires lors de l'acquisition pour qu'un système de reconnaissance d'iris fonctionne correctement. De plus, l'iris est un petit objet difficile à localiser sur le visage, et nécessite donc l'utilisation d'instruments très puissants et très proches de l'individu lors de l'acquisition. Sa position derrière la cornée en fait un objet réfléchissant lorsqu'il est vu avec de la lumière, il est donc recouvert de reflets lumineux. Pour résoudre ce problème, une lumière proche infrarouge (NIR : Near Infrared Imaging ) est recommandée pour l'acquisition des images de l'iris. De plus, les acquisitions nécessitent une forte coopération de la personne. Par conséquent, l'acquisition est réalisée dans des conditions contrôlées. Ces dernières années, il y a eu une nouvelle tendance à faire des acquisitions dans des conditions incontrôlées pour assouplir les restrictions imposées et les rendre plus acceptables [3].

#### 5.3.1. Acquisition d'images dans des conditions contrôlées

La plupart des systèmes traditionnels de reconnaissance de l'iris nécessitent une grande coopération de la part des individus au cours du processus d'acquisition. Selon les auteurs du rapport IREX-5 [40], les bonnes images ont un diamètre d'iris égal à au moins 160 pixels.

Ils recommandent de prendre des photos avec une bonne illumination. Avant la capture, il faut s'assurer que l'individu est immobilisé à courte distance (distance entre 30 cm et 100 cm), en face et regarder directement la caméra. Les yeux doivent être ouverts. La caméra doit être devant le visage de la personne et dégagée de tout obstacle. Après capture, l'iris doit apparaître dans l'image, doit être centré, avoir un bon contraste, l'occlusion de la pupille par les paupières supérieures et est bien illuminé à une rotation maximale de  $15^\circ$  (liée au mouvement de la tête). La **figure 1.6** illustre un exemple d'une bonne image de l'œil obtenue avec toutes les recommandations ci-dessus.



**Figure 1.6** – image de l'iris de bonne qualité acquise dans des conditions contrôlées [40].

### 5.3.2. Acquisition d'image dans des conditions non contrôlées

Le relâchement des contraintes imposées lors de l'acquisition est devenu nécessaire pour rendre l'acquisition plus flexible et adaptable à différents environnements. En d'autres termes, il s'agit de prendre une image de l'œil d'un individu relativement en mouvement, à distance dans le NIR, ou d'utiliser la lumière visible au lieu de la lumière NIR, par conséquent, selon les degrés de liberté d'acquisition, plusieurs facteurs indésirables sont introduits, entraînant la dégradation de la qualité de l'image de l'iris [3].

- **Dégradations lors de l'acquisition en lumière NIR**

Dans [40], les facteurs défavorables (voir **figure 1.7**) sont regroupés en trois catégories. Le premier est lié à la personne effectuant l'acquisition, y compris le bruit du mouvement de la caméra, la rotation de l'iris dans l'image (probablement causée par le mouvement de la personne). La deuxième catégorie, liée au comportement individuel, comprend l'occlusion des paupières et des cils, les reflets lumineux liés

aux lunettes, la déclinaison et la forte dilatation ou contraction de la pupille. La dernière catégorie est liée à la caméra d'acquisition et renferme le bruit de focus, l'absence d'iris dans l'image, un éclairage irrégulier, un faible contraste pupille/iris ou iris/sclera.

- **Dégradations lors de l'acquisition en lumière visible**

Lors de l'acquisition d'images oculaires en NIR, des précautions doivent être prises pour éviter tout dommage possible, car l'œil humain n'est pas sensible à l'illumination NIR et ne peut pas utiliser ses mécanismes de défense naturels, tels que le clignement des yeux. En revanche, le spectre visible peut être capté par l'œil humain sans aucune limitation. Cela peut fournir un environnement de capture plus flexible. La qualité des images acquises dans ces environnements est dégradée par rapport aux images acquises dans des environnements utilisant la lumière NIR.

Les facteurs indésirables mentionnés dans les images NIR, tels que les changements d'éclairage, le faible contraste, l'occlusion par des lunettes, les reflets, semblent être plus importants et plus sévères dans les images couleur acquises en lumière visible. Ainsi, il y avait une détérioration considérable de la qualité de la texture de l'iris [3].

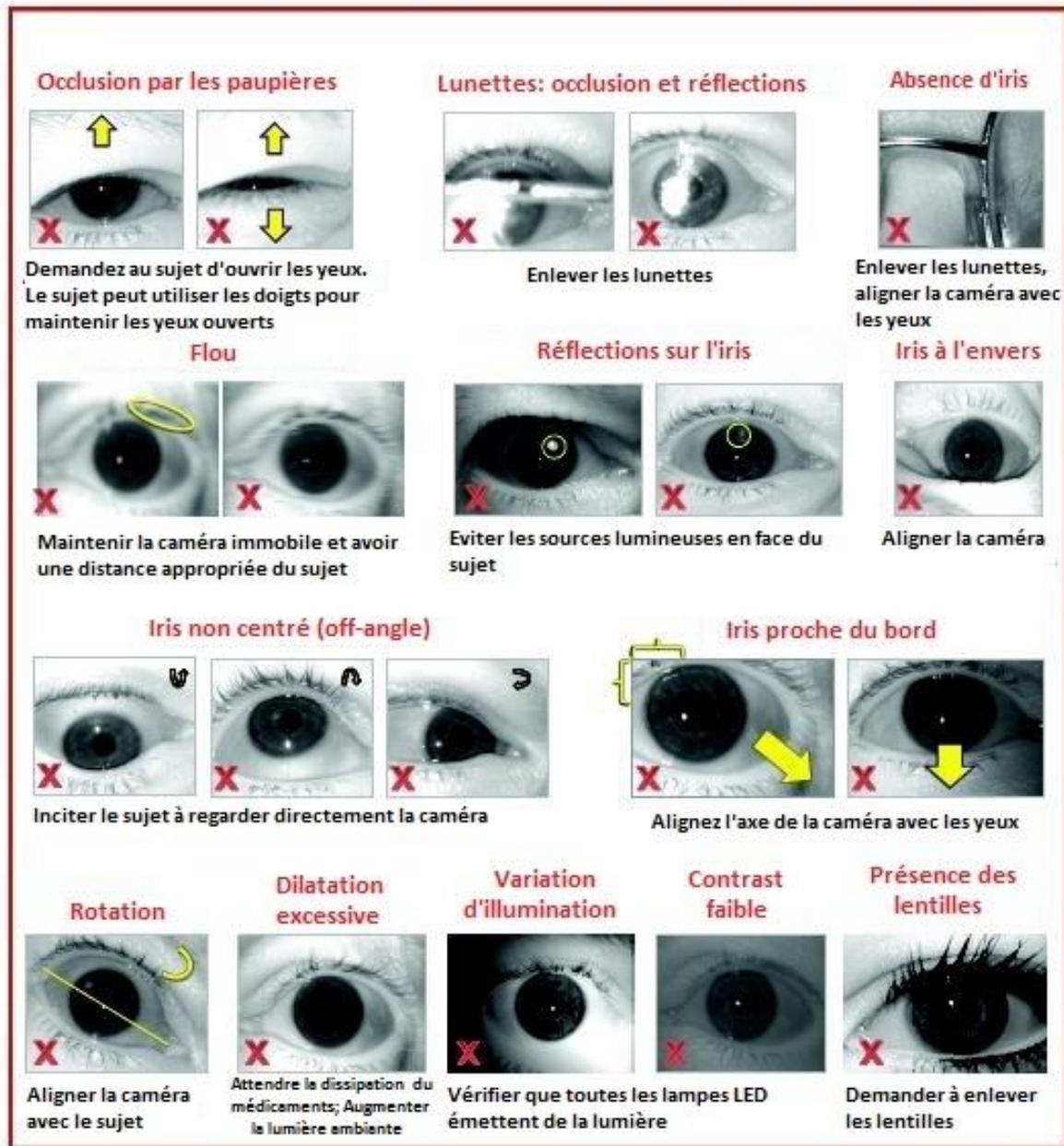


Figure 1.7 – des images d'iris dégradées acquise en NIR [40].

La figure 1.8 est un exemple d'acquisition à distance de plusieurs images dégradées à l'aide de la lumière visible dans un environnement plus flexible.



**Figure 1.8** – des images d'iris dégradées acquise en lumière visible [3].

La première détection des contours pupille/iris dans les images acquises en NIR est considérée comme une procédure naturelle pour toutes les méthodes de segmentation de l'iris décrites dans la littérature, puisque la pupille est la zone la plus sombre de l'image de l'œil. En revanche, dans le cas d'images en couleur, la pupille ne représente pas toujours la partie la plus sombre en raison des forts reflets. De plus,

avec les yeux foncés, le principal problème est que le contraste pupille/iris est très faible. La segmentation devient alors difficile, voire impossible.

Quelle que soit la méthode d'acquisition, des facteurs défavorables affecteront négativement la reconnaissance de l'iris, en particulier la segmentation de l'iris. Il a été rapporté dans [41] qu'une faible erreur dans l'estimation des paramètres du cercle (erreur de segmentation) modifie l'image normalisée de l'iris produite par le module de normalisation, ce qui réduit considérablement les performances de reconnaissance.

#### **5.4. Spécificités d'un système de reconnaissance d'iris**

La biométrie de l'iris présente plusieurs défis et difficultés que nous devons surmonter pour définir un système de reconnaissance fiable basé sur cette modalité. Comme nous l'avons souligné, l'iris est un organe interne du corps qui est visible de l'extérieur vers la cornée et devrait agir comme un miroir hautement réfléchissant. Le fonctionnement du système de reconnaissance de l'iris repose sur une séquence de plusieurs processus. Chaque sous-étape a ses propres difficultés et problèmes. Au niveau de l'acquisition des images de l'iris, les conditions d'acquisition nécessaires doivent toujours être contrôlées pour acquérir des images de bonne qualité qui limitent la variabilité liée à l'acquisition (distance, focus, luminosité). Il est très difficile d'acquérir des images d'iris à grande distance, de plus, ces systèmes nécessitent des conditions d'éclairage spécifiques et la coopération de sujet. Par contre, connaissant les propriétés réfléchissantes de l'iris, l'image acquise peut toujours contenir des réflexions spéculaires sur la pupille ou elles couvrent certaines zones importantes de l'iris qui peuvent compliquer la reconnaissance [7][42].

Après acquisition de l'image, l'iris doit être segmenté. Cette étape est considérée comme la plus difficile de toutes les étapes de traitement. Une segmentation précise de l'iris doit isoler l'iris de la pupille, du blanc de l'œil, de la paupière, des cils et de la réflexion spéculaire pendant un temps limité. La grande difficulté de segmentation de l'iris est liée à plusieurs facteurs. Premièrement, la détection des limites de l'iris est compliquée par un faible contraste entre l'iris et la pupille ou entre l'iris et le blanc de l'œil, selon l'éclairage de l'iris. La deuxième difficulté survient lorsque les paupières sont partiellement fermées, La texture de l'iris peut être couverte par ce que l'on appelle un bruit occultant. Ces bruits peuvent être dus à la présence de la paupière selon l'ouverture de l'œil et selon la population. Jusqu'à 60 % des textures d'iris

peuvent être occultée par ce genre de bruit. De plus, les cils qui recouvrent l'iris rendent plus difficile l'isolement de l'iris, notamment à cause du faible contraste entre l'iris et le bout des cils, en plus la présence des cils dont la forme irrégulière et la position aléatoire en font un 'bruit' très difficile à détecter. D'autre part, des réflexions spéculaires dues à l'illumination de l'iris peuvent apparaître dans la zone de l'iris. Ces réflexion en plus des réflexions dues au port de lunettes ou de lentilles de contact, ajoutent du bruit à l'image de l'iris et doivent être isolées. Il y a aussi du bruit de flou, l'iris est un petit objet qu'il faut capturer à distance, il faut donc utiliser une focale forte (et donc sensible). Typiquement, utiliser une telle distance focale signifie que la personne dont on a acquiert l'iris doit être placée à une distance quasi constante de l'objectif, sinon une image floue de l'iris peut être obtenue. En plus de ce flou d'acquisition, il y a aussi le flou de bougé de l'appareil photo, qui se produit lorsque le mouvement de l'objet est supérieur à la sensibilité de l'appareil photo, les images floues et non focalisées sont un véritable défi pour isoler correctement l'iris et nécessitent un traitement spécialisé [7][42].

Pour la normalisation de l'iris, le processus doit transformer l'iris segmenté en une matrice de normalisation universelle, en tenant compte des changements de taille et des rotations de l'iris. Le phénomène de l'iris se rencontre en cas de mouvements brusques de l'œil (l'ouverture/fermeture de la paupière) et la contraction ou la dilatation de la pupille.

Sachant que l'iris est un muscle qui se dilate ou se contracte en fonction de la luminosité dans l'environnement d'acquisition, ces mouvements du muscle de l'iris provoquent une déformation inverse de la texture de l'iris. Malheureusement, cette déformation est complètement non linéaire et difficilement imprévisible, donc un échantillonnage approprié est nécessaire pour tenir compte de la non-linéarité de la contraction de l'iris. D'autre part, la texture de l'iris n'est pas répartie de manière uniforme dans tout l'iris, et certaines zones de l'iris sont beaucoup plus riches que d'autres. En effet, les motifs spécifiques de certaines textures d'iris ne sont pas très riches, ce qui peut conduire à diverses erreurs de reconnaissance. Il semble donc difficile d'élaborer une norme qui tienne compte de tous ces facteurs tout en minimisant l'introduction de nouvelles erreurs [7][42].

Au niveau de l'encodage et de la création des profils d'iris, le défi est de trouver des transformations et des représentations qui extraient et expriment les informations les plus discriminantes de l'iris. Ces informations extraites doivent représenter la texture

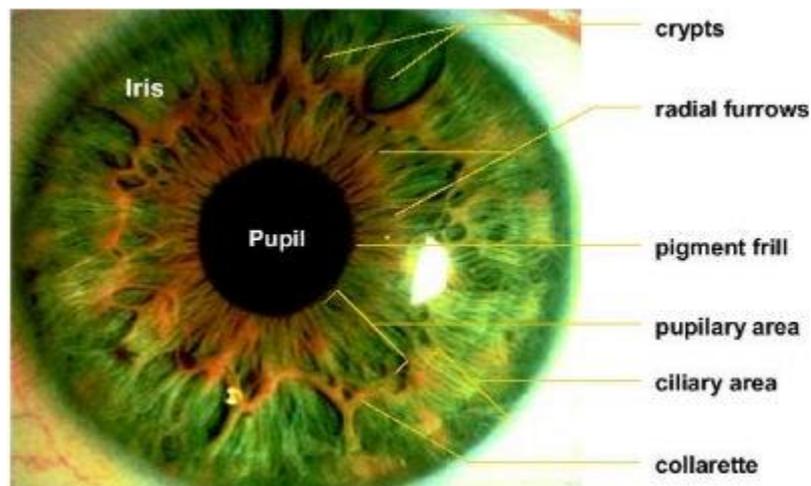
de l'iris et être aussi indépendantes que possible du bruit (changements d'éclairage, rotation et occlusion de l'iris, etc.). De plus, étant donné que la richesse de la texture de l'iris varie d'une personne à l'autre, certaines textures d'iris sont relativement uniformes, tandis que d'autres présentent de nombreux détails.

Enfin, la difficulté de la classification des images de l'iris réside dans le développement des méthodes d'identification performante ou de méthodes de vérification de l'iris efficaces, selon le motif requis.

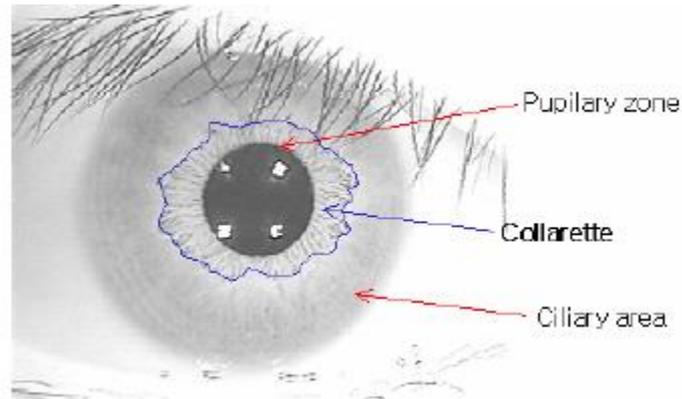
Nous recherchons donc principalement des méthodes de segmentation et de normalisation de l'iris pour mieux représenter la texture de l'iris. En conséquence, un codage et une classification plus efficaces peuvent être obtenus, conduisant finalement à de meilleures performances pour l'ensemble du système [7][42].

### 6. Collerette zigzag (collerette irienne) – Etat de l'art

L'iris est bien visible, surtout au microscope biologique, il présente un relief très irrégulier avec deux régions, une pupille interne et une externe périphérique ou ciliaire, séparées par la collerette irienne **figure 1.9**, la collerette irienne est située à l'union du tiers interne et des deux tiers externes de la face antérieure et apparaît comme une ligne irrégulière, saillante. Elle correspond à la limite de réabsorption de la membrane pupillaire, dont des reliquats sont généralement présents. Plus ou moins évident, généralement bien visible, souvent avec des dépôts pigmentés [43].



(a)



(b)

Figure 1.9-(a)[44] et (b)[45] collerette zigzag.

Dans la section 4.2 de ce chapitre, un survol des travaux qui ont été présentés dans le domaine de la détection de l'iris.

Dans les travaux mentionnés précédemment, les auteurs se sont concentrés sur la région de l'iris. Malgré l'efficacité des méthodes mentionnées ci-dessus pour les images moins bruitées, elles subissent souvent une sérieuse baisse de performance lorsqu'elles sont confrontées à des masques très irréguliers ou mal segmentés. Une solution à ce problème est donc nécessaire pour une segmentation robuste, complète et plus précise de l'iris.

Motivé par les travaux précédents, cette thèse se concentre sur la zone **collerette zigzag**, et plus précisément sur le **demi-cercle inférieur** du zigzag de la collerette. Le choix de la zone à étudier le demi-cercle inférieur de collerette zigzag - n'est pas arbitraire car il s'agit d'une zone de l'iris à texture riche, qui n'est affectée ni par les paupières ni par les cils. La principale contribution de notre idée est de sélectionner la région la plus importante de l'iris, un motif plus complexe, avec moins de bruit, pour éviter un autre traitement de suppression du bruit qui réduit la qualité de l'image dans la plupart des cas.

En fait, la région de la collerette zigzag est l'une des parties les plus importantes du motif de l'iris en raison de sa texture riche, de son insensibilité à la dilatation de la pupille, et parce qu'elle n'est pas affectée par les cils ou la paupière. Dans [46], les auteurs ont constaté empiriquement que la collerette en zigzag est généralement concentrique à la pupille et que son rayon est limité dans un certain intervalle.

L'utilisation de la zone de la collerette dans la reconnaissance de l'iris a donné lieu à de nombreuses avancées au cours de la dernière décennie :

- Dans **Sung et al. (2004)** [45], la détection de la limite de la collerette est utilisée pour améliorer le taux de reconnaissance. L'égalisation de l'histogramme et un filtre passe-haut sont appliqués, après avoir utilisé une TFD unidimensionnelle, les auteurs ont utilisé les informations statistiques de l'image pour détecter la limite de la collerette.
- Dans **He et Shi. (2005)** [46], l'auteur propose une nouvelle méthode de segmentation de l'iris basée sur le théorème de la corde croisée (crossed chord theorem) et la zone de la collerette en zigzag, ainsi, le coût de calcul est réduit et le cercle peut être localisé par la méthode de la corde croisée dont les paramètres peuvent être calculés par la méthode des triples points.
- Dans **Roy et Bhattacharya. (2006)** [55], les auteurs ont utilisé la méthode du code de chaîne et la zone de collerette en zigzag avec une machine à vecteur de support (SVM) pour améliorer la méthode de reconnaissance de l'iris pour l'identification des personnes.
- Dans **Roy et Bhattacharya. (2007)** [47], une technique efficace de reconnaissance de l'iris est présentée, basée sur la localisation de la zone de la collerette en zigzag et sur une machine à vecteurs de support asymétrique.
- Dans **Jeng et al. (2012)** [48], les auteurs ont présenté une nouvelle méthode de normalisation. Ils ont proposé une méthode de localisation de l'iris et une méthode de localisation de la collerette. Pour l'extraction des caractéristiques, ils ont proposé le filtre à ondelettes de Gabor pour extraire les caractéristiques des images d'iris.
- Dans **Joshi et al. (2012)** [51], l'auteur propose une nouvelle technique d'extraction de caractéristiques utilisant des ondelettes [52] combinées à la DLDA (Direct Linear Discriminant Analysis [53] pour extraire des vecteurs de caractéristiques discriminants de faible dimension de la région de la collerette.
- Dans **Rai et Yadav. (2014)** [49], l'auteur a utilisé l'ondelette de Haar pour localiser la collerette en zigzag et a utilisé un filtre Log Gabor 1D pour l'extraction des caractéristiques.

- Dans **Punyani et al. (2016)** [50], une combinaison de machine à vecteur de support (SVM : Support Vector Machine), de réseaux neurones artificiels (ANN : Artificial Neural Network) et de zone de collerette en zigzag est utilisée pour effectuer l'extraction de caractéristiques pour le système de reconnaissance de l'iris.
- Dans **Faundra et Sulistyaningrum. (2017)** [54] la segmentation et la normalisation de l'iris, des algorithmes basés sur la collerette en zigzag sont présentés. Les auteurs ont utilisé la détection des bords de canny et les transformées de Hough pour localiser la pupille près de la zone en zigzag. Ensuite, l'utilisation d'un modèle de rubber sheet de Daugman représente une collerette en zigzag isolée.
- Dans **Jauro et Yadav. (2020)** [56], l'auteur a utilisé une méthode de segmentation à deux niveaux pour segmenter l'image. Dans la section de segmentation de la limite intérieure, ils ont utilisé des méthodes telles que la pyramide de Gauss, la diffusion anisotrope, le seuillage, etc. Dans la section de segmentation de la limite extérieure, les auteurs ont effectué un processus de collerette en zigzag en utilisant la méthodologie de la collerette en zigzag. Enfin, la segmentation de la limite intérieure a été soustraite de la segmentation de la limite extérieure pour obtenir l'iris segmenté.

## 7. Conclusion

Bien que la biométrie ait été développée pour identifier les criminels, elle est aujourd'hui utilisée pour établir une reconnaissance de personne dans un grand nombre d'applications civiles, commerciales et gouvernementales [57], [58]. Parmi les différentes modalités biométriques, la reconnaissance d'image de l'iris a connu un véritable essor depuis 1993, et est aujourd'hui considérée comme l'une des meilleures modalités biométriques. Cependant, des performances supérieures sont toujours recherchées en raison des exigences de niveaux de sécurité, de certification, d'identification et de vérification [7]. Compte tenu des cinq tâches de la technologie du système de reconnaissance de l'iris, la segmentation est considérée comme la plus difficile. Le développement de ce processus est nécessaire pour améliorer les performances du système. La deuxième tâche tout aussi importante est la

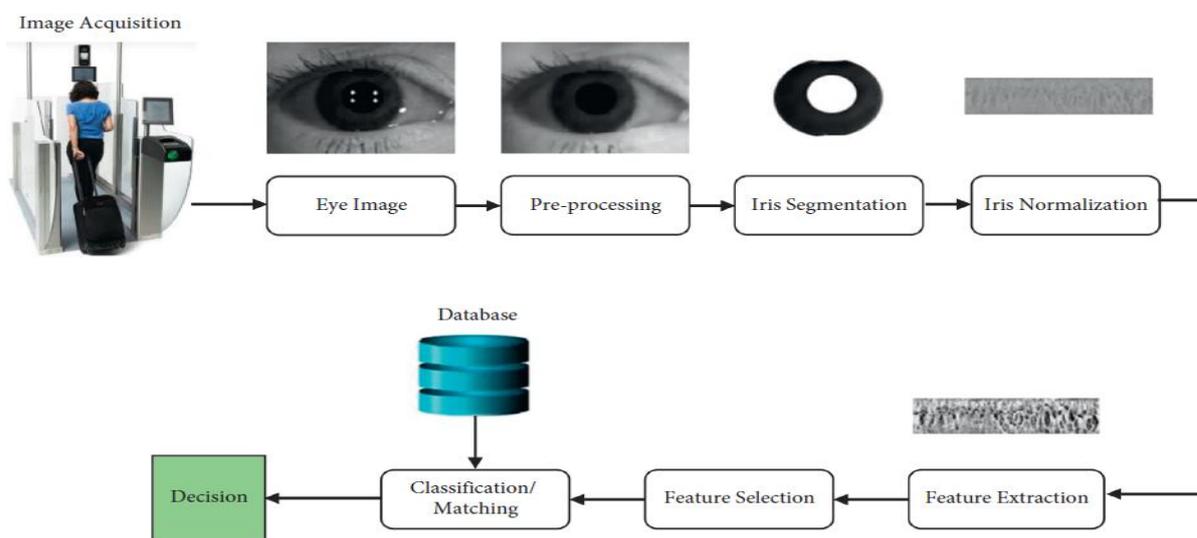
normalisation de l'iris, qui doit garantir un échantillonnage optimal des informations de l'iris en introduisant un minimum d'erreurs.

# Chapitre 2

## Les techniques de reconnaissance de l'iris

### 1. Introduction :

La demande croissante de systèmes d'authentification efficaces dans le domaine de la sécurité a suscité le développement de systèmes d'authentification plus sûrs et plus efficaces. Les approches traditionnelles de l'identification, telles que l'utilisation d'une clé ou d'un mot de passe, ne sont pas satisfaisantes dans plusieurs domaines d'application, car ces méthodes peuvent être facilement oubliées, volées ou piratées. Pour surmonter ces faiblesses, la science moderne s'intéresse à l'automatisation des systèmes de reconnaissance en utilisant techniques biométriques [59]. La **figure 2.1** montre la structure d'un système de reconnaissance de l'iris



**Figure 2.1**-la structure d'un système de reconnaissance de l'iris [60].

### 2. les techniques de prétraitement

Il est recommandé de mettre en œuvre une méthode de prétraitement dans la première

phase pour réaliser l'une des techniques ou l'un des algorithmes suggérés du système de reconnaissance. Ces techniques peuvent aider à éliminer les différents types de bruit qui se sont produit pendant le processus d'obtention de l'image de l'iris. Typiquement, les images d'iris obtenues souffrent de différents types de bruits tels que l'occlusion causée par les paupières ou les cils, l'éclairage ou la lumière, le bruit dû aux lunettes comme le flou, les réflexions spéculaires, les rayures et autres facteurs, distorsion, et les yeux hors axe. L'élimination de ces types de bruit par la phase de prétraitement améliorera la précision des performances de système de reconnaissance de l'iris [60].

**Le tableau 2.1**, montre les différents types de méthodes de prétraitement utilisées pour éliminer les différents types de bruit.

**Tableau 2.1-** Les types de bruit et les méthodes de prétraitement correspondantes [60]

La référence	Type de bruit	La méthode propose
Nixon and Aguado [61]	Bruit	Filtre median
Huang et al. [62]	Flou	Estimation par noyau du flou
Lee et al. [63]	Reflexions	Profil d'intensité de ligne
Santos and Hoyle [64]	Faible éclairage (faible illumination)	La méthode CLAHE
Dehkordi and Abu-Bakar [65]	Bruit tel que les paupières, les cils, les réflexions de la lumière et les pixels des pupilles.	Méthode de seuillage multiple
Liu et al. [66]	Faible précision des performances de la reconnaissance de l'iris à faible résolution	Algorithme d'apprentissage métrique hétérogène
Raffei et al. [67]	Faible contraste	Égalisation adaptative de l'histogramme
Bakshi et al. [68]	Les occlusions dues aux cils et aux paupières	Les filtres gaussiens et le détecteur de lignes de Hough
Kumar et al. [69]	La différence entre le pixel le plus clair et le plus foncé.	Filtre haut de forme et bas de forme (Top-hat and bottom-hat)
Baqar et al. [70]	Accentuation spéculaire due à la reflection.	Interpolation linéaire 2D
Djoumessi [71]	L'occlusions dues aux paupières	Détecteur de lignes de Hough
Gangwar et al. [72]	La réflexion spéculaire	Seuil
Radman et al. [73]	Réflexions	Méthode morphologique-retinex
Ribeiro et al. [74]	Faible résolution et faible qualité des images d'iris	La technique de l'auto-codeur empilé (SAE : Stacked Auto-Encoders) et la technique du réseau neuronal (CNN)
Arsenovic et al. [75]	L'ombre sur l'oeil	Normalisation de l'histogramme
Gad et al. [86]	La réflexion spéculaire	Opérations morphologiques
Susitha and Subban [77]	Faible qualité en raison d'un faible contraste	La méthode CLAHE
Das and Derakhshani[78]	Contraste et luminosité médiocres	La méthode BPDFHE (Brightness Preserving Dynamic Fuzzy Histogram Equalization)

### 2.1. Méthode de la transformée de Hough :

Pendant la phase de prétraitement de certains travaux précédents en utilisant la méthode de la transformée de Hough (HT : Hough Transform), Raffei et al. ont utilisé la méthode de la transformée de Hough linéaire pour éliminer les occlusions dues aux paupières [79]. Malgré cela, il y avait des limitations, par exemple, les images de l'iris obtenues en lumière visible avaient plus de limitations telles que le bruit de réflexion et les occlusions dues aux cils et aux paupières et aussi la forme de l'iris n'était pas idéale [80], donc Chai et al. ont introduit une méthode pour localiser les paupières ainsi que pour identifier et éliminer la réflexion [80]. Dans la localisation de la paupière, il a été supposé que la position du centre de l'iris peut être estimée en utilisant une transformée de hough circulaire (CHT : Circular Hough Transform).[81] ou le détecteur en cascade de Haar [82]. Cette méthode localisée a été proposée afin de diminuer l'impact du bruit, qui génère deux zones de taille d'intérêt  $N \times N$  dans les parties supérieure et inférieure de la région de l'iris. L'intensité moyenne du pixel de la région de l'iris pour chaque fenêtre est calculée pour découvrir les pixels de la paupière. Dans l'identification et l'élimination de la réflexion, ils ont proposé d'identifier d'abord les régions impactées par les différents types de réflexions, en appliquant une méthode de seuillage [83]. Ensuite, en utilisant une opération morphologique de dilatation, des pixels sont encore ajoutés près des frontières de la région importante de l'iris pour agrandir les régions sélectionnées. Il est à noter que l'occlusion de la région de l'iris due aux cils et aux paupières était encore fréquente, ce qui a conduit Yahiaoui et al. à développer davantage la méthode en appliquant un CHT et en localisant la région de la pupille à l'aide de contour de Canny [84]. Afin de préserver les bords robustes et de surmonter la sensibilité de l'algorithme de détection des bords, ils ont appliqué un filtre anisotrope suivi d'un seuil adaptatif. Avec ces limitations telles que les réflexions dues aux lunettes, aux cils, aux sourcils et aux verres correcteurs ou de mode avec l'image obtenue, toutes ces méthodes ont trouvé qu'il était difficile de localiser la région de l'iris [85], [86].

### 2.2. Histogramme et méthode de filtrage :

Pour augmenter le taux de reconnaissance de l'iris et la précision de la segmentation de la région de l'iris à partir du reste de l'arrière-plan de l'image, Santos et Hoyle ont utilisé

une méthode d'égalisation d'histogramme (HE : Histogram Equalization) pour améliorer une image de l'iris acquise à faible éclairage afin d'identifier les limites de la région de la pupille. Ces techniques sont utilisées pour redistribuer les intensités des pixels [64]. Pour améliorer les performances de reconnaissance de l'iris dans des conditions variables de qualité d'image et également le contraste des images d'iris acquises en utilisant un faible contraste ou un faible éclairage, Alvarez- Betancourt et Garcia-Silvente ont proposé la méthode d'égalisation d'histogramme adaptative limitée par le contraste (CLAHE : Contrast Limited Adaptive Histogram Equalization) [87], une version modifiée de la méthode d'égalisation d'histogramme adaptative (AHE : Adaptive Histogram Equalization) qui a été introduite par Zuiderveld [88]. Cette méthode permet d'améliorer le contraste des images en niveaux de gris. Maheshan et al. ont également travaillé sur les mêmes limitations et ont amélioré l'image en appliquant deux techniques : HE et CLAHE. L'objectif de HE est de déterminer la fréquence des couleurs sombres qui s'étendent généralement dans la plage de 0 à 50 pixels. Le CLAHE, quant à lui, fixe une limite de contraste qui donne une balance des blancs proportionnelle pour l'image sélectionnée [67].

Kumar et al. ont introduit une méthode d'amélioration du contraste en utilisant des filtres à haut de forme et bas de forme (top hat and bottom hat) pour améliorer la différence entre les pixels de la région la plus claire et ceux de la région la plus sombre. La technique proposée comporte quatre étapes [69]. La première est HE [61] qui a été appliquée pour améliorer le contraste. La deuxième étape est le filtre médian qui a été appliqué par Mark [61] pour supprimer le bruit. La troisième étape est la correction d'intensité gamma (GIC : Gamma Intensity Correction) qui, selon Saeedi et al, est utilisée pour la variation d'illumination [89]. Dans la quatrième étape, l'image de l'iris est exposée à une séquence de filtres haut de forme et bas de forme, comme suggéré par Bai [90].

Zhang et al [91] ont en revanche appliqué la méthode de filtrage homomorphique telle que suggérée par Zhang et Shen [92] pour que l'image de l'iris ait un contraste robuste et pour surmonter l'impact de l'occlusion [92].

Ensuite, cette démarche est suivie d'une approche suggérée par Chang et al [93, 94] qui ont appliqué la méthode de prétraitement sur l'image de l'iris d'entrée en trois phases pour supprimer les reflets spéculaires de l'image de l'iris. Au cours de la première étape, La technique du filtrage gaussien a été appliquée avec une valeur sigma qui était égale à

0.9. La deuxième étape consistait à transformer les images des yeux de gris en binaire, en utilisant une valeur de seuil qui était égale à 0.18. La dernière étape consistait à appliquer à nouveau le filtre gaussien sur l'image binaire des yeux avec une valeur sigma qui était égale à 2, après cela le filtre médian a été utilisé pour lisser l'image des yeux [93, 94].

De même, Hofbauer et al [100] ont également appliqué le filtrage médian pour lisser l'image des yeux. Le filtrage médian permet de rendre l'image plus régulière et élimine également l'impact des cils afin de ne pas réduire la netteté de la région de l'iris. Malgré cela, d'autres défis se sont présentés, et les images obtenues présentaient souvent des différences significatives au niveau de l'arrière-plan et un faible contraste dans certaines régions [95]. L'une des principales raisons de cette variation est la mise en évidence de la région de la pupille. Pour surmonter ces limitations, Khan et al [96] ont également utilisé la technique du filtre gaussien mais avec un écart-type  $\sigma$ . Ceci a pour but d'améliorer les résultats non adaptés, ainsi que pour créer des images de l'œil plus appropriées pour le seuillage. La valeur choisie de  $\sigma$  était basée sur le volume de la région de la pupille et de la région de l'iris. Elle doit être inférieure au rayon de la pupille mais supérieure à la largeur des cils. Pour les images de l'iris de Khan et al. [96], la valeur de  $\sigma = 5$  était plus appropriée que les autres valeurs de  $\sigma$  pour apprécier le l'arrière-plan local.

### 2.3. Opérations morphologiques :

Pour éliminer les petits composants, Umer et al. [97] ont proposé une méthode de prétraitement. Cette méthode commence par l'analyse de l'histogramme de l'image originale de l'œil en utilisant le seuillage qui a converti l'image d'entrée originale en une image binaire. Ensuite, l'image binaire de l'iris a été améliorée en éliminant les bruits à l'aide d'une opération morphologique pour obtenir une image améliorée de l'iris. Cependant, la précision des performances était encore affectée par des facteurs de bruit dans des environnements sans contraintes (non idéaux) tels que la réflexion, les paupières et les cils [97]. Selon Djoumessi [71], des techniques morphologiques pourraient être utilisées pour supprimer le bruit de réflexion des images d'iris, puis un seuillage pourrait être appliqué pour supprimer les cils de l'image. Cependant, d'autres limitations telles que les lunettes, la déviation due à un regard sur le côté, et l'occlusion due aux cheveux et aux paupières sont apparues. Radman et al. ont donc proposé une méthode rétine morphologique pour surmonter ces limites. Ils ont appliqué des

opérations morphologiques pour améliorer le contraste et supprimer le bruit de réflexion des images d'iris dans des environnements sans contraintes (non idéaux). Ensuite, la méthode de la moyenne quadratique (RMS : Root Mean Square) a été appliquée pour décider si l'amélioration du contraste de l'image de l'iris était nécessaire ou non. Enfin, ils ont appliqué un algorithme retinex étendu pour améliorer le contraste des images de l'iris [73]. De même, Gad et al [76] ont également travaillé sur les mêmes limitations et ont appliqué une série d'opérations morphologiques, visant également à obtenir une performance précise et rapide de la réflexion de la cornée. Ensuite, une méthode de seuil local adaptatif (ALT : Adaptive Local Threshold) basée sur le filtre moyen a été appliquée pour filtrer l'image de l'iris de sortie des pixels de la région brillante. Enfin, en utilisant la technique de masquage (MT : Masking Technique) [98], la valeur de tous les pixels situés au-dessus et au-dessous du diamètre de la pupille a été convertie en zéro pour supprimer les cils et les paupières [76].

### 3. Techniques de segmentation de l'iris

La région d'intérêt (ROI) est extraite immédiatement après la phase de prétraitement [85]. L'extraction de la région d'intérêt est appelée segmentation ou localisation [99]. La phase de segmentation vise principalement à détacher la partie de l'iris des parties résiduelles non utiles, à savoir les parties autour de la région de l'iris (sclérotique, paupières et peau) et la partie de la pupille [100]. Les techniques de reconnaissance de l'iris dépendent de la qualité des caractéristiques de la région qui a été segmentée. Ainsi, la précision des performances du système de reconnaissance est principalement due à la précision de la phase de segmentation [99]. La raison principale de la plupart des échecs des systèmes de reconnaissance est due à une segmentation inappropriée [101].

#### 3.1. Segmentation de l'iris à l'aide de techniques traditionnelles :

La plupart de ces techniques exécutent la segmentation par l'application de méthodes traditionnelles de traitement d'images, comme la méthode de Daugman, le clustering et l'algorithme de segmentation sémantique. Dans cette section, certaines techniques de segmentation utilisées dans des travaux antérieurs pour segmenter la région de l'iris sont expliquées. En utilisant des techniques traditionnelles. Bien que toutes ces techniques montrent une précision de haute performance pour les techniques de segmentation de l'image de l'iris, leurs performances sont toujours déterminées par différents facteurs environnementaux, tels que la longue distance d'éloignement, le hors axe, le flou, les

reflets, l'occlusion, la petite taille de l'iris dans des conditions sans contraintes (non idéales). Par conséquent, l'amélioration récente des techniques basées sur l'approche d'apprentissage profond a conduit à l'introduction de caractéristiques d'apprentissage profond pour la segmentation de l'iris.

### 3.1.1. Techniques de segmentation basées sur la transformée de Hough :

La segmentation de l'iris à l'aide de techniques traditionnelles a commencé par la méthode de la transformée de Hough (HT). Bakshi et al. [68] et Raffei et al. [102] ont appliqué une phase de segmentation de l'iris qui consiste en trois sous-phases, à savoir la localisation de la région de l'iris, l'occlusion causée par la paupière et l'occlusion causée par un cil. Dans la localisation de la région de l'iris, la méthode de la transformation de Hough circulaire (CHT) est mise en œuvre pour obtenir les limites de la région de l'iris [103]. Comme la pupille se trouve dans la partie de l'iris, cette opération commence par la limite externe (la limite entre l'iris et la sclérotique) plutôt que par l'image entière de l'œil. Ils ont appliqué une méthode de transformation de Hough linéaire (HT) pour supprimer les occlusions des paupières. Ensuite, ils ont utilisé la méthode de seuillage pour supprimer les occlusions des cils [104, 68, 102]. Malgré cela, il y avait des limitations dues aux environnements à distance et moins contraints, comme le bruit de réflexion et les occlusions dues aux paupières, aux cils et à l'ombre. Ensuite, la localisation de l'iris à l'aide de la méthode CHT qui a été introduite par Davies [105] a été utilisée pour détecter les demi-cercles inférieurs afin de déterminer le centre du cercle et d'éviter les interférences des cils ou des sourcils. Ensuite, la région de la moitié inférieure des pixels de l'iris a été identifiée par une analyse locale du niveau de gris à l'aide d'un seuillage qui est basé sur l'histogramme pour déterminer les pixels bruyants, ce qui a aidé à déterminer les seuils de la région de la moitié supérieure des pixels de l'iris pour la précision de la segmentation. En outre, le taux d'erreur du processus de segmentation de l'iris est très élevé car il prend beaucoup de temps et la qualité des images de l'iris est mauvaise [106]. Pour réduire la zone de recherche ainsi que le temps de calcul de la CHT, Umer et al. [107] ont introduit une technique de transformation de Hough circulaire restreinte (RCHT : Restricted Circular Hough Transform) pour localiser les bords extérieurs (iris-sclérotique) et intérieurs (pupille-iris) de la région de l'iris. Tout d'abord, ils ont utilisé un filtre morphologique pour éliminer les petits bruits afin d'améliorer l'image. Ensuite, le RCHT a été utilisé sur les points de bord de l'image binaire de l'œil pour obtenir le centre et le rayon du bord

interne circulaire (entre l'iris et la pupille) de la région de l'iris. Pour obtenir le bord extérieur circulaire (entre l'iris et la sclérotique) de la région de l'iris, la différence de haute intensité entre les pixels de l'iris pour les directions verticale et horizontale à partir du centre du bord intérieur a été détectée. Ici, ce système n'est pas proposé pour supprimer la réflexion, les paupières, les cils et le bruit [107].

Salve et Narote [108] ont appliqué une méthode de fusion basée sur le détecteur de bords de Canny ainsi que la méthode HT [109] pour segmenter la partie iris. Les performances de la segmentation de l'iris sont faibles en raison de la mauvaise qualité des images dans certaines bases de données telles que CASIA-V4-Iris Twins [108]. Pour augmenter la précision des performances de la segmentation de la région de l'iris avec certaines limitations telles que la paupière et la réflexion, Ahuja et al. [110] ont utilisé l'algorithme de segmentation qui était basé sur la méthode de la transformée de Hough généralisée (GHT : Generalized Hough Transform) utilisée par Ballard [111] qui a ensuite été développé à l'aide d'un algorithme intégral-différentiel modifié de Daugman [112] pour ségréguer le pixel de la région de l'iris dans l'image de l'œil comme décrit dans Haindl et Krupic`ka [113]. Cet algorithme a fourni les pixels segmentés et normalisés de la région de l'iris ainsi que le masque binaire pour éliminer le bruit de l'image de l'iris [110].

### 3.1.2. Techniques de segmentation basées sur les histogrammes et les contours :

En outre, pour effectuer la segmentation, Othman et al. [114] ont présenté une description de trois versions du système de reconnaissance de l'iris à code source ouvert (OSIRIS), à savoir OSIRISV2 ([https://svnext.it-sudparis.eu/svnview2-eph/ref\\_syst/Iris\\_Osiris/](https://svnext.it-sudparis.eu/svnview2-eph/ref_syst/Iris_Osiris/)), OSIRISV4 ([https://svnext.it-sudparis.eu/svnview2-eph/ref\\_syst/Iris\\_Osiris\\_v4.1/](https://svnext.it-sudparis.eu/svnview2-eph/ref_syst/Iris_Osiris_v4.1/)) et OSIRISV4.1. OSIRISV2 est la première version open-source depuis l'année 2009. La segmentation d'OSIRISV2 consiste en deux étapes :

(1) localisation de l'iris en utilisant la méthode de la transformée de Hough circulaire (CHT) et (2) initialisation des contours de l'iris en utilisant l'approche suggérée par Xu et Prince [115] pour initialiser un contour actif et ensuite initialiser une génération de masque de fenêtre qui référence uniquement la texture de la région de l'iris en exploitant les contours raffinés. En pratique, l'évaluation des versions OSIRIS.V2 [115] a montré les faiblesses de la phase de segmentation [114]. Ces faiblesses ont motivé les chercheurs à augmenter la précision des performances de la technique de segmentation dans toutes les versions d'OSIRIS [114]. Les versions OSIRIS.V4 et OSIRIS.V4.1 ont

utilisé l'algorithme de Viterbi pour surmonter les limites de la segmentation de l'iris dans OSIRIS.V2 [116]. Cependant, les images de faible qualité présentent généralement un flou, un faible contraste, des changements d'illumination, des réflexions spéculaires nettes et des occlusions dues aux paupières et aux cils, ainsi que des réflexions dues aux lunettes, et peuvent par conséquent poser un défi lors de la segmentation en utilisant les algorithmes habituels basés sur l'ajustement de contour [114].

Pour développer la performance de la segmentation de l'iris du modèle open-source OSIRISV4 pour ces défis, Yahiaoui et al. [84] ont introduit une méthode en étendant l'algorithme de Viterbi [116] en ajoutant des techniques statistiques pour la segmentation de l'iris basée sur des approches non supervisées, et ils se sont concentrés particulièrement sur la méthode de la chaîne de Markov cachée comme suggéré dans [84, 117, 118]. Afin d'augmenter le taux de performance de la phase de segmentation des images de l'iris, certaines limitations étaient encore présentes telles que la réflexion par les lunettes et les scanners différents ainsi que la réflexion occultée par les paupières et les cils, ce qui a conduit He et al [119] et Liu et al [120] à développer la méthode de la transformée de Hough (HT) basée sur la méthode de détection des bords qui a été appliquée par Canny [121] pour segmenter la région de l'iris. La méthode de la transformée de Hough a été utilisée pour détecter les frontières en exploitant la duplicité entre les points sur les frontières et les paramètres qui étaient les coordonnées du centre et du rayon de ces frontières [111], et ces paramètres ont été déterminés sur la base du poids de la matrice après que le seuil du rayon ait été donné [120, 119].

### 3.1.3. Autres techniques de segmentation

Osorio-Roig et al.[122] ont mis en œuvre une technique de segmentation sémantique appelée HMRF-PyrSeg [123] qui a été essentiellement conçue pour segmenter les images publiques. La technique proposée a été utilisée pour identifier la région de l'iris dans les images des yeux capturées dans le spectre des longueurs d'onde visibles. Osorio-Roig et al. [122] ont également fourni des informations sémantiques sur diverses régions de l'image de l'œil, telles que les sourcils, l'iris, la sclérotique et la pupille, afin de segmenter la région de l'iris. Pour améliorer la précision de la performance de cette technique, les recherches futures pourraient vouloir appliquer des techniques hybrides qui permettent de combiner des régions de différents niveaux de segmentation et également de déterminer automatiquement le niveau de segmentation le plus approprié pour chaque image [122].

Djoumessi [71] a proposé une amélioration des techniques réalisées comme suggéré par Masek [124] pour la segmentation de l'iris. La méthode proposée utilise l'opérateur de Daugman integro-différentiel pour déterminer les frontières des régions de la pupille et de l'iris. Ensuite, ils ont utilisé la méthode de la transformée de Hough linéaire (LHT) pour supprimer les paupières, et le seuil a été utilisé pour supprimer les autres occlusions telles que les cils de l'iris segmenté [75]. Pour augmenter la précision de performance de la méthode de segmentation de l'iris en éliminant les cils et les paupières, Nguyen et al. [125] ont appliqué l'opérateur intégro-différentiel basé sur le filtre gaussien flou pour détecter les bords circulaires qui étaient les bords internes et externes de la région de l'iris via l'exploration itérative des réponses les plus élevées d'un contour déterminé par les paramètres. Lorsqu'il y a une occlusion de l'iris due aux paupières et aux cils, le meilleur candidat est sélectionnée en utilisant l'approche itérative de la technique de seuillage localisant les paupières avec l'intégration du contour. L'application de l'opérateur intégro-différentiel transforme l'iris en un arc plutôt qu'en un cercle. Les pixels de l'iris sont différenciés des pixels non iris (tels que les cils et les paupières) dans une image d'entrée par la méthode du masque de bruit. Les masques de bruit, qui correspondent à chaque image d'entrée, sont créés au cours de la phase de segmentation et appliqués dans les phases de reconnaissance de l'iris [125].

Pour améliorer la vitesse et la performance des différents défis tels que les réflexions et les occlusions, Gad et al [76] ont appliqué la technique de masquage modifiée (MMT) pour la segmentation de l'iris. L'exécution de cet algorithme comprend trois étapes, à savoir : (1) la génération du masque pour supprimer les cils, (2) détection de l'iris par des opérations de convolution du masque pour détecter la frontière de la région de l'iris et la détection des paupières en utilisant un filtre à seuil local adaptatif (ALT), et (3) convolution avec un vecteur unitaire (M) pour extraire la région de l'iris en appliquant trois points dérivés du masque de l'iris avec trois points de distance. Ces points ne sont pas égaux en raison des facteurs de détection des frontières et de l'occlusion et des défis tels que l'utilisation de lentilles ou de lunettes, la distance, le faible contraste, les réflexions spéculaires, les paupières et les cils prévisibles, hors axe [76].

Cela a conduit Khan et al [96] à introduire la méthode de segmentation de l'iris basée sur un réseau de portes programmables (FPGA : Field Programmable Gate Array), qui est considéré comme le plus approprié pour les méthodes en temps réel. La méthode proposée a utilisé l'image prise après soustraction de l'arrière-plan, ainsi que des processus morphologiques pour localiser la pupille. La méthode non itérative suggérée a

été exécutée sur un réseau de portes programmables (FPGA) pour obtenir des performances et une vitesse élevées avec une mémoire plus petite [96]. Cependant, la méthode proposée n'a pas réussi à localiser le bord de la pupille en présence de cils et de sourcils épais, et elle ne fonctionnera pas avec précision sur les images du spectre visible en raison du contraste plus faible entre les régions de l'iris et de la pupille dans le spectre visible.

Reddy et al. [126] ont proposé un algorithme de segmentation de l'iris sur des images capturées sur mobile qui repose sur l'intégration des algorithmes intégréo-différentiels de Daugman et du clustering K-means. L'algorithme proposé peut être divisé en trois étapes principales. Il commence par l'élimination de la réflexion spéculaire, s'il y en a, en plus de l'amélioration du contraste des images de l'œil en appliquant CLAHE [127]. Elle est ensuite suivie par la mise en œuvre du clustering K-means pour extraire la région de l'iris. Enfin, l'algorithme intégréo-différentiel de Daugman est appliqué pour détecter le rayon et le centre des régions de la pupille et de l'iris [126]. Les études futures devraient se concentrer sur de nouvelles méthodes de compensation de l'éclairage conçues pour améliorer l'extraction des caractéristiques des images oculaires. Un intérêt particulier devrait être porté aux techniques conçues pour les images d'iris collectées avec du bruit dû aux lunettes et à une grande distance entre la caméra et le sujet [128].

#### **4. Techniques de normalisation**

La normalisation est l'opération qui consiste à convertir le motif circulaire de la région de l'iris segmentée en un motif rectangulaire. La phase de normalisation est appliquée après la phase de segmentation pour faciliter l'extraction des caractéristiques.

Pour normaliser la région de l'iris, Umer et al. [97], Yahiaoui et al. [84], et Al-Waisy et al. [129] ont utilisé un modèle de rubber sheet de Daugman [112] pour transformer l'iris segmenté d'une forme circulaire en une forme rectangulaire. Le modèle Daugman's rubber-sheet model transforme le modèle linéaire de la zone de l'iris de coordonnées cartésiennes en coordonnées polaires. Daugman [130] a amélioré modèle de rubber sheet pour normaliser la partie segmentée de l'iris, ce qui permet de réduire les incohérences de l'image lors de la prise de vue de l'iris. Des chercheurs comme Raffei et al. ont appliqué le modèle de rubber sheet suggéré par Daugman [130] pour convertir les dimensions de la région de l'iris segmentée de la forme circulaire en une forme rectangulaire, avec des dimensions de 20×240 pixels [104, 102].

De même, le modèle de rubber sheet de Daugman a été utilisé par Bakshi et al. [68], Djoumessi [71], Gad et al. [76], Salve et Narote [108], Nguyen et al. [125], et Kaur et al. [131] pour normaliser la région de l'iris. Ceci a été accompli en changeant la région de l'iris segmentée en un rectangle fixé. Par conséquent, un choix populaire serait la méthode de rubber sheet de Daugman qui peut être utilisée à plusieurs reprises. Cependant, le défi de cette méthode est qu'elle est plus sensible à la rotation. En raison de la différence du rayon de la pupille et de l'iris, certains des pixels sont répétés par le processus de normalisation pour remplir l'espace vide de l'image de l'iris. Plusieurs chercheurs passés et présents ont tenté d'améliorer continuellement les méthodes de normalisation afin de surmonter cette limitation, comme Shamsi et Rasouli, qui ont introduit une approche trapézoïdale [132]. Dans cette approche, la région de l'iris est divisée en un certain nombre de cercles concentriques, puis déroulé. Il n'est pas nécessaire de procéder à une itération car il s'agit d'un mappage un à un. Ainsi, cette technique réduit les fausses acceptations. La faiblesse de cette approche est qu'elle a une structure fixée [132]. À ce titre, Hilal et al [133] ont introduit une méthode flexible pour les images non circulaires de la région de la pupille. Cette méthode a utilisé des disques flexibles plutôt qu'un motif circulaire fixé pour tracer la nature incurvée de la région de l'iris. Le défi de cette approche est de détecter le centre de la région de la pupille. L'extraction de cette caractéristique serait difficile car toute erreur commise dans la localisation du centre peut avoir un impact sur les contours élastiques.

### **5. Technique d'extraction des caractéristiques :**

Après avoir obtenu les limites de l'iris et isoler cette région, des méthodes d'extraction de caractéristiques doivent être utilisées pour obtenir les caractéristiques spéciales de la texture de la région de l'iris qui diffèrent d'un individu à l'autre. La phase d'extraction des caractéristiques est un processus très important pour obtenir une grande précision dans l'identification d'une personne car les caractéristiques uniques de chaque personne doivent être extraites. Les techniques d'extraction de caractéristiques peuvent être divisées en deux sections qui sont l'extraction de caractéristiques à l'aide de techniques traditionnelles et l'extraction de caractéristiques à l'aide de techniques d'apprentissage profond.

### 5.1. Extraction de caractéristiques à l'aide de techniques traditionnelles.

Cette section explique l'application des techniques traditionnelles pour l'extraction de caractéristiques. Une de ces techniques pour extraire les caractéristiques des images d'iris normalisées serait les filtres 1D log-Gabor. Gong et al. [134], Raffei et al. [104, 67, 102], Bakshi et al. [68], et Salve et Narote [108] ont appliqué des filtres 1D log-Gabor pour l'extraction de caractéristiques d'images d'iris normalisées. Les filtres 1D log-Gabor sont une extension des filtres de Gabor qui n'étaient pas adaptés pour traiter les informations spectrales avec une localisation spatiale maximale. L'inconvénient du filtre de Gabor est que chaque fois que la largeur de bande dépasse une octave, même le filtre symétrique aura une composante continue (CC), alors que les filtres log-Gabor n'ont aucune composante de CC pour n'importe quelle largeur de bande [108]. En ayant des extrémités étendues, les filtres de log-Gabor sont plus efficaces de coder les images originales que les filtres de Gabor traditionnels, qui surreprésenteraient les segments de basse fréquence et sous-représenteraient les segments de haute fréquence dans tout codage. Les défis des filtres de Gabor sont abordés avec l'application des filtres de log-Gabor.

D'autres chercheurs tels que Yahiaoui et al. [84], Othman et al. [114], Jan et al. [135, 136], Jan et Usman [137], Raja et al. [138] et, plus récemment, Nazmdeh et al. [139] ont proposé l'utilisation de filtres de Gabor 2D, qui sont très populaires et ont été mis en œuvre avec succès dans des applications pratiques. Les filtres de Gabor 2D sont généralement utilisés pour obtenir les informations de textures globales dans les régions de l'iris pour la méthode d'extraction de caractéristiques.

Kumar et al. [69] ont appliqué la combinaison de la transformée en ondelettes discrètes (DWT : Discret Wavelet Transform) introduite par Jang et al. [140] et la transformée en cosinus discrète (DCT : Discrete Cosine Transform) pour extraire les caractéristiques de la région de l'iris. La méthode proposée présente encore des limites telles que la mise en œuvre dans des environnements non restreints et les applications en temps réel [69]. Ainsi, Barpanda et al. [141] ont proposé d'utiliser l'ondelette dérivée de la banque de filtres 9/7 de Cohen-Daubechies-Feauveau (CDF) pour extraire les caractéristiques de l'iris en trois étapes, à savoir : (1) le calcul de l'ondelette à deux niveaux à partir du gabarit de l'iris, (2) le calcul du spectre de puissance de l'image transformée par ondelette et (3) la composition de la transformée en cosinus discrète pour obtenir les coefficients du spectre [141]. Récemment, Barpanda et al. [142] ont proposé d'appliquer

le mel-cepstrum d'ondelettes afin d'extraire la caractéristique de l'iris. La technique suggérée est basée sur l'ondelette dérivée de la banque de filtres orthogonaux publics Cohen-Daubechies-Feauveau (CDF) 9/7. Dans la méthode proposée, les coefficients de cepstre de fréquence mélangée (MFCC) ont été appliqués pour distinguer les tissus de l'iris. La technique proposée est utilisée pour extraire les caractéristiques de l'iris en trois étapes. La première étape consiste à calculer les coefficients de détail à deux niveaux du tissu de l'iris. Ensuite, la méthode logarithmique est appliquée pour diviser ces coefficients détaillés en carrés irréguliers. Cela permet de réduire les dimensions des coefficients d'ondelettes, puis de désigner des poids irréguliers aux différents segments de fréquence. Ensuite, la méthode de la transformée en cosinus séparée est utilisée pour calculer les mêmes caractéristiques, grâce auxquelles les caractéristiques sont extraites [142]. La faiblesse de la méthode suggérée est qu'avec la diminution de la taille des caractéristiques, il y a une augmentation subtile de l'imprécision. Grâce à l'analyse des résultats, il est évident que les caractéristiques du cepstre reposent sur l'ondelette pour obtenir des performances équivalentes à celles de la banque de filtres CDF 9/7. Ainsi, on pense que les recherches futures devraient se concentrer sur le développement d'un système de reconnaissance de l'iris avec une précision plus élevée d'une manière efficace en termes de temps pour les images collectées sous le NIR et le spectre visible en utilisant les informations de couleurs.

## **5.2. Extraction de caractéristiques à l'aide de techniques d'apprentissage profond.**

Minaee et al. [143] ont appliqué l'architecture du réseau VGG16 suggérée par Krizhevsky et al. [144] et Simonyan et Zisserman [145] pour obtenir une extraction profonde des caractéristiques des images de l'iris et dessiner une carte profonde des caractéristiques en exécutant une séquence de convolution. Ensuite, la méthode d'analyse en composantes principales (ACP) a été appliquée par Abdi et Williams [146] pour réduire la dimensionnalité des caractéristiques. Le résultat de cette méthode pourrait être amélioré en entraînant un réseau profond spécialement pour la reconnaissance des images d'iris [143]. Pour découvrir les images des yeux générées par l'ordinateur en explorant les différences dans la région de l'œil, Carvalho et al. [147] ont introduit l'architecture VGG avec 19 couches (VGG19) au lieu de l'architecture VGG avec 16 couches (VGG16) qui a été initialement suggérée par Simonyan et Zisserman [145] en se basant sur les réseaux neuronaux profonds (DNN) comme base de l'ImageNet. Dans la méthode proposée, les blocs convolutifs fonctionnent comme des

couches d'extraction de caractéristiques. Les futures directions de recherche devraient approfondir l'amélioration des techniques d'extraction de caractéristiques en introduisant la fusion des caractéristiques extraites de différents modèles de CNN profonds, car le système de fusion est censé atteindre des performances plus élevées [147, 125].

Afin d'améliorer les performances tout en réduisant la complexité de calcul et l'espace mémoire requis, ainsi que la pleine participation à la conception du réseau profond de système de reconnaissance de l'iris pour la conception automatique des caractéristiques et la conception de la structure du réseau en même temps, Nguyen et al [148] ont introduit une méthode pour restructurer la conception du réseau de reconnaissance de l'iris qui tient compte des contraintes de calcul et de mémoire. La méthode proposée a été utilisée pour étudier les performances de l'IrisCode traditionnel et les comparer aux techniques de réseaux profonds. Elle a également tenté d'améliorer les réseaux profonds actuels de reconnaissance de l'iris pour obtenir une efficacité similaire ou supérieure avec le même niveau de mémoire et de complexité de calcul. Plus important encore, la technique proposée automatise entièrement le processus de conception du réseau pour rechercher la meilleure architecture de réseau de reconnaissance de l'iris confiné à la compacité du modèle et aux contraintes de calcul. Selon cet algorithme proposé, on peut conclure que la performance des réseaux profonds est affectée lorsque le calcul est limité de manière robuste, tandis que cela améliore la performance des techniques traditionnelles IrisCode à la fois dans le calcul et la précision de la reconnaissance de l'iris [148].

Ahuja et al. [110] ont proposé deux méthodes basées sur les CNNs pour vérifier les images de la partie périoculaire de l'œil, qui comprend la région de l'iris. Dans la première méthode, ils ont proposé une technique non supervisée basée sur les CNNs, fusionnant le résultat des vecteurs de caractéristiques externes et la technique de base RootSIFT (Scale Invariant Feature Transform), comme suggéré par Arandjelovic' et Zisserman [149]. Dans la deuxième méthode, ils ont proposé la technique supervisée basée sur CNN et ont utilisé une similarité cosinus comme suggéré par Nguyen et Bai [150]. pour obtenir la similarité entre deux images. Les deux méthodes sont basées sur le CNN profond et utilisent une architecture à couches superposées [110]. Afin d'améliorer la précision de la reconnaissance biométrique oculaire, Reddy et al. [151] ont introduit OcularNet basé sur une architecture CNN pour la biométrie oculaire mobile. Dans le modèle OcularNet proposé, ils ont extrait six patches inscrits dans les régions périoculaire et oculaire et ont formé un petit CNN pour chaque patch appelé

PatchCNN pour extraire et générer des descripteurs de caractéristiques qui sont différenciables entre les individus. La recommandation était que les études futures augmentent l'analyse des architectures différentes optimisées pour les mobiles comme MobileNet ainsi que Mobile-Net-v2 pour diminuer davantage le nombre de paramètres [151].

Pour obtenir des caractéristiques de l'iris avec une distinction plus claire ainsi que pour résoudre les problèmes de détermination des seuils de classification pour les tâches de reconnaissance de l'iris en augmentant les échantillons d'iris, Chen et al. [152] ont proposé une fonction de perte à centre étroit (T-center :Tight Center) pour extraire les caractéristiques de la région de l'iris avec une distinction plus proéminente. La fonction de perte T-center proposée a été utilisée pour résoudre le problème de distinction insuffisante avec la fonction de perte softmax conventionnelle. La capacité discriminative des caractéristiques profondes basées sur le modèle CNN peut être améliorée par la fusion linéaire des fonctions de perte softmax et T-center [152].

D'autres chercheurs comme Zhang et al. [91] et Al-Waisy et al. [129] ont introduit une technique basée sur une approche CNN pour extraire les caractéristiques de la région de l'iris. Dans la technique proposée, les caractéristiques de la région de l'iris ont été extraites automatiquement en appliquant une convolution avec des couches de mise en commun. Dans les utilisations pratiques, la prise d'images de l'iris est un défi lorsqu'il y a la réflexion due aux lunettes, aux environnements lumineux différents, etc [91, 129]. De plus, des recherches supplémentaires sont encore nécessaires pour améliorer la précision de la performance aléatoire en lisant aléatoirement l'image de l'iris.

Pour augmenter la précision de la performance de l'extraction des caractéristiques, Marra et al. [153] ont suggéré une technique basée sur les techniques CNN pour adapter et diminuer le nombre de paramètres ainsi que la complexité de calcul (à la fois dans les étapes d'apprentissage et de test) de l'AlexNet qui a été suggéré par Krizhevsky et al. [144]. La technique AlexNet proposée consiste en trois couches convolutionnelles pour extraire la caractéristique. Dans ces trois couches, une unité linéaire rectifiée (ReLU) est implémentée comme une tâche non linéaire après la convolution, qui est suivie par des couches de max-pooling [153]. Zhao et Kumar [154] ont proposé un cadre basé sur l'apprentissage profond, nommé UniNet.v2, pour la détection, la segmentation et la reconnaissance de l'iris.

Lee et al [155] ont introduit un système de reconnaissance oculaire qui définit le contour de l'iris sans mettre en œuvre la segmentation de la région de l'iris. Pour

résoudre le problème d'efficacité réduite de la reconnaissance de l'iris qui reposait sur la détermination grossière, une région oculaire légèrement plus grande que la région de l'iris a été utilisée pour mettre en œuvre la reconnaissance de l'iris, et un réseau résiduel profond (ResNet) a été appliqué pour résoudre le problème d'un taux de reconnaissance réduit en raison d'un désalignement et les images de vérification entre les images enregistrées [155]. Il y avait encore des limitations aux méthodes proposées, telles que la rotation sévère, les cils épais, l'hors axe, le flou de mouvement et le cadre de lunettes dans l'image de l'iris. D'autres limitations incluent les images d'iris avec une mauvaise résolution, par exemple lorsque l'image a été prise sur des sujets non coopératifs et lorsqu'il y avait une longue distance entre la caméra et le sujet. Pour résoudre ces problèmes, les chercheurs devraient chercher des moyens d'améliorer les performances du système de reconnaissance en utilisant les informations de reconnaissance des deux yeux et en collectant un ensemble de données d'images de l'iris, comme celles de jumeaux, en raison de leur similarité. En outre, l'accent devrait également être mis sur l'étude de techniques qui peuvent augmenter le domaine réceptif de l'architecture CNN et extraire des caractéristiques pour résoudre le problème de la réduction des performances de la reconnaissance de l'iris en raison du désalignement entre les images de l'iris enregistrées et celles de la reconnaissance, causé par la détection approximative de la région oculaire.

## **6. Conclusion :**

Une variété d'approches du système de reconnaissance d'image d'iris selon différentes phases basées sur deux approches qui sont l'apprentissage traditionnel et l'apprentissage profond ont été présentées dans ce chapitre avec les contributions des chercheurs ainsi que les forces et les faiblesses de certaines techniques de système de reconnaissance de l'iris, ce chapitre montre que l'avenir de système de reconnaissance de l'iris est prometteur et encourage les chercheurs à poursuivre leurs études afin de résoudre les défis mentionnés ci-dessus et à y apporter des solutions appropriées.

## **Le chapitre 3**

# **Les algorithmes de classification**

### **1. Introduction**

L'intelligence artificielle (IA) consiste à rendre les machines aussi intelligentes que le cerveau humain. En informatique, l'IA désigne l'étude des "agents intelligents" : tout dispositif qui perçoit son environnement et entreprend des actions qui maximisent ses chances de réussite pour atteindre ses objectifs [156]. La tendance technologique actuelle ouvre la voie à un avenir dans lequel les machines intelligentes, les processus en réseau et le big data sont réunis. Ce monde virtuel a généré une grande quantité de données qui accélère l'adoption de solutions et de pratiques d'apprentissage automatique [157]. De manière informelle, le terme "intelligence artificielle" est appliqué lorsqu'une machine est capable d'exécuter des fonctions que les humains associent à d'autres esprits humains, comme "l'apprentissage" et "la résolution de problèmes". L'apprentissage est un aspect essentiel des machines. Par conséquent, l'apprentissage automatique est un sous-domaine de l'IA [156]. L'apprentissage automatique permet aux ordinateurs d'imiter et d'adapter un comportement de type humain. Grâce à l'apprentissage automatique, chaque interaction, chaque action effectuée, devient quelque chose que le système peut apprendre et utiliser comme expérience pour la prochaine fois [157].

### **2. Apprentissage automatique (Machine Learning)**

Depuis leur évolution, les humains utilisent de nombreux types d'outils pour accomplir diverses tâches de manière plus simple. La créativité du cerveau humain a conduit à l'invention de différentes machines. Ces machines ont facilité la vie de l'homme en lui permettant de répondre à divers besoins, notamment les voyages, l'industrie et l'informatique, et l'apprentissage automatique est l'une d'entre elles.

Le terme a été inventé par Arthur Samuel en 1959, qui a défini la ML (Machine Learning) comme un domaine d'étude qui fournit une capacité d'apprentissage aux ordinateurs sans qu'ils soient explicitement programmés [158], Arthur Samuel était célèbre pour son programme de jeu de dames. L'apprentissage machine ML est utilisé pour apprendre aux machines à traiter les données de manière plus efficace. Parfois, après avoir consulté les données, nous ne pouvons pas interpréter ou extraire des informations de ces données. Dans ce cas, nous appliquons l'apprentissage automatique. Avec l'abondance des ensembles de données disponibles, la demande d'apprentissage automatique est en hausse. De nombreuses industries appliquent l'apprentissage automatique pour extraire des données pertinentes. L'objectif de l'apprentissage automatique est d'apprendre à partir des données. De nombreuses études ont été réalisées sur la manière de faire apprendre les machines par elles-mêmes sans être explicitement programmées, pour atteindre une plus grande précision (accuracy), la précision étant mesurée en fonction du nombre de fois où les actions choisies sont correctes. De nombreux mathématiciens et programmeurs appliquent plusieurs approches pour trouver la solution à ce problème qui comporte d'énormes ensembles de données [159].

Les problèmes du monde réel sont très complexes, ce qui en fait d'excellents candidats pour l'application de l'apprentissage automatique. L'apprentissage automatique peut être appliqué à divers domaines de l'informatique afin de concevoir et de programmer des algorithmes explicites offrant un rendement élevé, par exemple, le filtrage des spams, la détection des fraudes sur les réseaux sociaux, le commerce des actions en ligne, la détection des visages et des formes, le diagnostic médical, la prédiction du trafic, la reconnaissance des caractères et la recommandation de produits, entre autres. Les voitures Google à conduite autonome, Netflix qui présente les films et les émissions qu'une personne pourrait aimer, les moteurs de recommandation en ligne - comme les suggestions d'amis sur Facebook, et la détection de la fraude par carte de crédit - sont tous des exemples concrets d'application de l'apprentissage automatique [157].

Les algorithmes d'apprentissage automatique sont organisés dans les groupes suivants :

- **Apprentissage supervisé** : les différents algorithmes génèrent une fonction qui associe les entrées aux sorties souhaitées. Une formulation standard de la tâche d'apprentissage supervisé

est le problème de la classification : l'apprenant doit apprendre (pour approcher le comportement de) une fonction qui fait correspondre un vecteur à l'une de plusieurs classes en examinant plusieurs exemples d'entrée-sortie de la fonction.

- **Apprentissage non supervisé** : modélisation d'un ensemble d'entrées : les exemples étiquetés ne sont pas disponibles.
- **Apprentissage semi-supervisé** : combine des exemples étiquetés et non étiquetés pour générer une fonction ou un classificateur approprié.
- **Apprentissage par renforcement** : l'algorithme apprend une politique d'action à partir d'une observation du monde. Chaque action a un certain impact sur l'environnement, et l'environnement fournit un retour d'information qui guide l'algorithme d'apprentissage.
- **Transduction** : similaire à l'apprentissage supervisé, mais ne construit pas explicitement une fonction : au lieu de cela, il essaie de prédire les nouvelles sorties en fonction des entrées d'apprentissage, des sorties d'apprentissage et des nouvelles entrées.
- **Apprendre à apprendre** : où l'algorithme apprend son propre biais inductif basé sur l'expérience précédente.

Dans notre thèse on s'est concentré sur l'algorithme supervisé, dans les algorithmes supervisés, les classes sont prédéterminées. Ces classes sont créées à la manière d'un ensemble fini, défini par l'homme, ce qui signifie en pratique qu'un certain segment de données sera étiqueté avec ces classifications. La tâche de l'algorithme d'apprentissage automatique est de trouver des modèles et de construire des modèles mathématiques. Ces modèles sont ensuite évalués en fonction de leur capacité de prédiction par rapport aux mesures de variance des données elles-mêmes [160].

### **3. Développement de l'apprentissage automatique**

Les mots "intelligence artificielle" et "apprentissage automatique" ne sont pas nouveaux. Ils ont été étudiés, utilisés, appliqués et réinventés par des informaticiens, des ingénieurs, des chercheurs, des étudiants et des professionnels du secteur depuis plus de 60 ans. Le fondement mathématique de l'apprentissage automatique repose sur l'algèbre, les statistiques

et les probabilités. Le développement sérieux de l'apprentissage automatique et de l'intelligence artificielle a commencé dans les années 1950 et 1960 avec les contributions de chercheurs comme Alan Turing, John McCarthy, Arthur Samuels, Alan Newell et Frank Rosenblatt. Samuel a proposé le premier modèle d'apprentissage automatique fonctionnel sur le programme Optimizing Checkers. Rosenblatt a créé le Perceptron, un algorithme populaire d'apprentissage automatique basé sur des neurones biologiques, qui a jeté les bases du réseau neuronal artificiel [161, 162, 163]. Le **tableau 3.1** ci-dessous décrit le développement illustre, expansif et pratique de l'apprentissage automatique [157].

**Tableau 3.1**-Développement de la LM [157]

<b>1950</b>	Alan Turing a créé le "test Turning" pour vérifier l'intelligence d'une machine. Pour réussir le test Turning, la machine doit être capable de convaincre les humains qu'ils sont en train de parler à un humain et non à une machine.
<b>1952</b>	Samuel a créé un algorithme d'apprentissage très performant, capable de jouer au jeu de dames avec lui-même et de s'auto-entraîner.
<b>1956</b>	Martin Minsky et John McCarty, avec Claude Shannon et Nathan Rochester, ont organisé une conférence à Dartmouth en 1956 où l'intelligence artificielle est née.
<b>1958</b>	Frank Rosenblatt a créé le Perceptron, qui a posé la première pierre du développement du réseau neuronal artificiel (ANN).
<b>1967</b>	L'algorithme du plus proche voisin a été proposé pour la "reconnaissance des formes".
<b>1979</b>	Des étudiants de l'université de Stanford ont mis au point le "Stanford Cart", un robot sophistiqué capable de se déplacer dans une pièce et d'éviter les obstacles sur son chemin.
<b>1981</b>	L'apprentissage basé sur l'explication (EBL) a été proposé par Gerald Dejong, selon lequel un ordinateur peut analyser les données d'apprentissage et créer des

	règles pour écarter les données inutiles [164].
<b>1985</b>	NetTalk a été inventé par Terry Sejnowski [165], qui a appris à prononcer les mots anglais de la même manière que les enfants.
<b>1990s</b>	L'apprentissage automatique s'est déplacé de la connaissance vers les données. L'apprentissage automatique a été mis en œuvre pour analyser de grandes quantités de données et en tirer des conclusions [166].
<b>1997</b>	IBM a inventé l'ordinateur Deep Blue qui a pu battre le champion du monde d'échecs Gary Kasparov.
<b>2006</b>	Le terme "apprentissage profond" a été inventé par Geoffery Hinton pour désigner une nouvelle architecture de réseaux neuronaux qui utilise plusieurs couches de neurones pour l'apprentissage.
<b>2011</b>	Watson d'IBM, conçu pour répondre à des questions posées en langage naturel, bat un concurrent humain au jeu Jeopardy.
<b>2012</b>	Jeff Dean, de Google, a développé GoogleBrain, un réseau neuronal profond permettant de détecter des motifs dans les vidéos et les images.
<b>2014</b>	Facebook a inventé l'algorithme "DeepFace" basé sur des réseaux neuronaux profonds capables de reconnaître les visages humains dans les photos.
<b>2015</b>	Amazon a proposé sa propre plateforme d'apprentissage automatique. Microsoft a créé le "Distributed Machine Learning Toolkit" pour une distribution efficace des problèmes d'apprentissage automatique à plusieurs ordinateurs afin d'améliorer l'efficacité de l'apprentissage.  travaillent en parallèle pour trouver une solution [167,168]. Elon Musk et Sam Altman ont créé une organisation à but non lucratif, OoeaAI, dans le but d'utiliser l'intelligence artificielle au service des êtres humains.

<b>2016</b>	Google a proposé DeepMind qui est considéré comme le jeu de société le plus complexe. Le programme AlphaGo de Google devient le premier programme informatique de Go à battre un joueur humain professionnel. Il est basé sur la combinaison de techniques d'apprentissage automatique et de recherche d'arbres [169].
<b>2017</b>	Google a proposé Google Lens, Google Clicks, Google Home Mini et des téléphones basés sur Google Nexus qui utilisent des algorithmes d'apprentissage automatique et d'apprentissage profond. Nvidia a proposé les GPU NVIDIA - le moteur de l'apprentissage profond. Apple a proposé Home Pod, un dispositif interactif d'apprentissage automatique.

#### 4. Les applications de l'apprentissage automatique

Il y a divers domaines d'application de l'apprentissage automatique : La vision par ordinateur, la prédiction, l'analyse sémantique, le traitement du langage naturel et la recherche d'informations.

- Vision par ordinateur : La reconnaissance, la détection et le traitement des objets sont des sous-domaines du domaine de la vision par ordinateur.
- La prédiction : Les différents sous-domaines sont ici la classification, l'analyse et la recommandation. La classification de textes, la classification de documents, l'analyse d'images, le diagnostic médical, la prédiction de la détection des intrusions dans les réseaux et la prédiction des attaques par déni de service ont été mis en œuvre avec succès grâce à l'apprentissage automatique.
- Analyse sémantique, traitement du langage naturel et recherche d'information : L'analyse sémantique est le processus qui consiste à relier les structures syntaxiques des paragraphes, des phrases, des mots au niveau de l'écriture dans son ensemble. Le traitement du langage naturel est la façon de programmer les ordinateurs pour traiter correctement les données en langage naturel. La recherche d'informations est la science de la recherche d'informations dans un document, de la recherche de documents et de la recherche de métadonnées qui décrivent

les données et de bases de données de sons et d'images. Ce sont trois domaines dans lesquels les techniques d'apprentissage automatique ont été explorées dans le passé [156].

## **5. L'apprentissage automatique et ses applications**

Au cours de la dernière décennie, un grand nombre de méthodes d'apprentissage supervisé ont été introduites dans le domaine de l'apprentissage automatique. L'apprentissage supervisé est devenu un domaine où se déroulent de nombreuses activités de recherche en apprentissage automatique. De nombreuses techniques d'apprentissage supervisé ont trouvé une application dans le traitement et l'analyse de diverses données. L'une des principales caractéristiques de l'apprentissage supervisé est qu'il permet d'annoter les données d'apprentissage. Les étiquettes ainsi appelées sont des étiquettes de classe dans le processus de classification. Il existe une variété d'algorithmes qui sont utilisés dans les méthodes d'apprentissage supervisé [160].

### **5.1. Arbres de décision**

L'arbre de décision [170] représente un classificateur exprimé comme une partition récursive de l'espace d'instance. L'arbre de décision est constitué de noeuds qui forment ce qu'on appelle un arbre racine, ce qui signifie qu'il s'agit d'un arbre distribué avec un noeud de base appelé racine et sans arêtes entrantes. Tous les autres noeuds ont exactement un bord entrant. Le noeud qui a des bords sortants est appelé noeud interne ou noeud test. Les autres noeuds sont appelés feuilles. Dans un arbre de décision, chaque noeud de test divise l'espace d'instance en deux sous-espaces ou plus selon une certaine fonction discrète des valeurs d'entrée. Dans le cas le plus simple, chaque test considère un seul attribut, de sorte que l'espace d'instance est divisé en fonction de la valeur de l'attribut. Dans le cas d'attributs numériques, la condition fait référence à une plage [160].

Chaque feuille est affectée à une classe qui représente la valeur cible la plus appropriée. La feuille peut contenir un vecteur de probabilité qui indique la probabilité que l'attribut cible ait une certaine valeur. Les instances sont classées en les faisant naviguer de la racine de l'arbre vers les feuilles, en fonction du résultat des tests effectués le long du chemin. La **figure 3.1** décrit une utilisation simple de l'arbre de décision. Chaque noeud est étiqueté avec l'attribut qu'il teste, et ses branches sont étiquetées avec les valeurs correspondantes.

Avec ce classificateur, l'analyste peut prédire la réponse d'un client potentiel et comprendre les caractéristiques comportementales de l'ensemble de la population de clients potentiels [171].

Dans le cas d'attributs numériques, les arbres de décision peuvent être interprétés géométriquement comme une collection d'hyperplans, chacun orthogonal à l'un des axes. Les décideurs préfèrent les arbres décisionnels moins complexes, car ils peuvent être considérés comme plus complets [160].

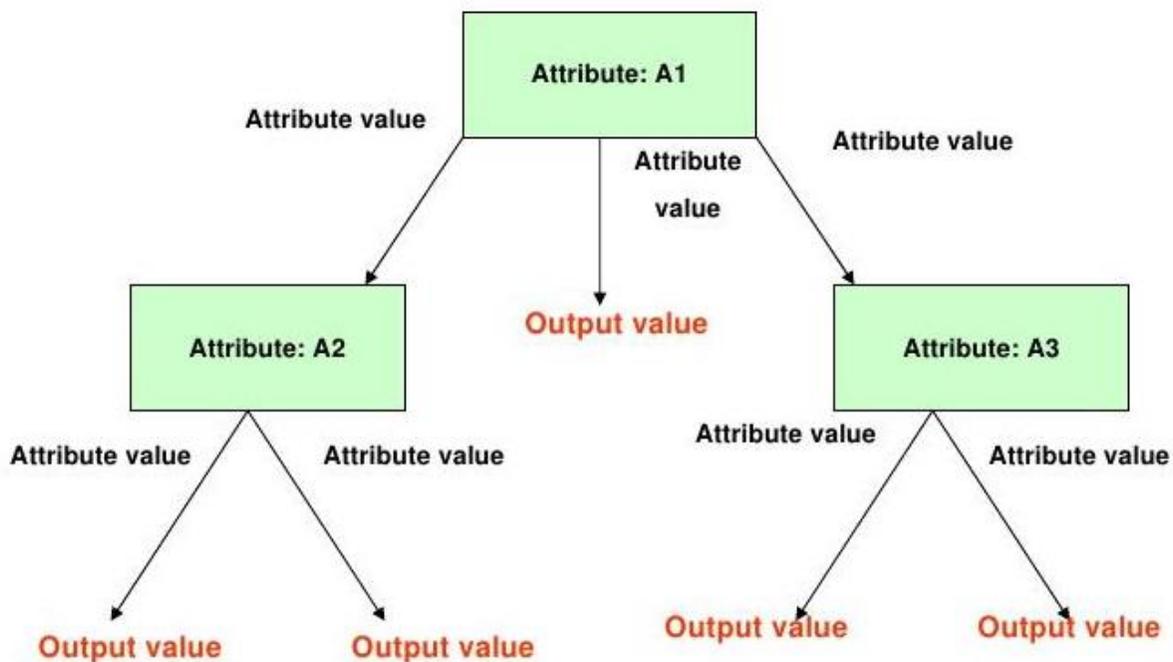


Figure 3.1-Exemple d'arbre de décision [172]

## 5.2. Régression linéaire

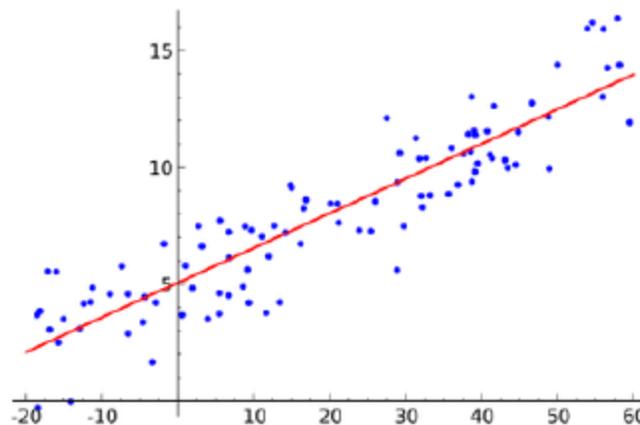
L'objectif de la régression linéaire, qui fait partie de la famille des algorithmes de régression, est de trouver des relations et des dépendances entre les variables. Elle représente une relation de modélisation entre une variable dépendante scalaire continue  $y$  (également étiquette ou cible en terminologie de l'apprentissage automatique) et une ou plusieurs (un vecteur à  $D$  dimensions) variables explicatives (également variables indépendantes, variables d'entrée, caractéristiques, données observées, observations, attributs, dimensions, point de données,

etc.) désignées par  $X$  à l'aide d'une fonction linéaire. Dans l'analyse de régression, l'objectif est de prédire une variable cible continue, alors qu'un autre domaine appelé classification consiste à prédire une étiquette à partir d'un ensemble fini. Le modèle pour une régression multiple qui implique une combinaison linéaire des variables d'entrée prend la forme :

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + e \quad (1)$$

La régression linéaire [11] appartient également à la catégorie des algorithmes d'apprentissage supervisé. Cela signifie que nous entraînons le modèle sur un ensemble de données étiquetées (données d'apprentissage) et que nous utilisons ensuite le modèle pour prédire les étiquettes sur des données non étiquetées (données de test)[160].

Comme le montre la **figure 3.2**, le modèle (ligne rouge) est calculé à partir de données d'apprentissage (points bleus) où chaque point a un label connu (axe y) pour ajuster les points le plus précisément possible en minimisant la valeur d'une fonction de perte choisie. On peut alors utiliser le modèle pour prédire des étiquettes inconnues (on ne connaît que la valeur x et on veut prédire la valeur y) [160].



**Figure 3.2-**Représentation visuelle de la régression linéaire [174].

### 5.3. Bayes naïfs

La classification bayésienne [175] est une autre méthode d'apprentissage supervisé ainsi qu'une méthode statistique de classification. Elle suppose un modèle probabiliste sous-jacent

et permet de capturer l'incertitude du modèle d'une manière raisonnée en déterminant les probabilités des résultats. L'objectif fondamental de la classification bayésienne est de pouvoir résoudre des problèmes de prédiction.

Cette classification fournit des algorithmes d'apprentissage pratiques et peut combiner les données observées. La classification bayésienne fournit une perspective utile pour comprendre et évaluer les algorithmes d'apprentissage. Elle calcule des probabilités explicites pour les hypothèses et elle résiste au bruit des données d'entrée [160].

Considérons une distribution de probabilité générale de deux valeurs  $\mathbf{P}(\mathbf{x}_1, \mathbf{x}_2)$ . En utilisant la règle de Bayes, sans perte de généralité, nous obtenons cette équation :

$$\mathbf{P}(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{P}(\mathbf{x}_1|\mathbf{x}_2)\mathbf{P}(\mathbf{x}_2) \quad (2)$$

De même, s'il existe une autre variable de classe  $c$ , nous obtenons l'équation suivante :

$$\mathbf{P}(\mathbf{x}_1, \mathbf{x}_2|c) = \mathbf{P}(\mathbf{x}_1|\mathbf{x}_2, c)\mathbf{P}(\mathbf{x}_2|c) \quad (3)$$

Si la situation est généralisée avec deux variables à une indépendance conditionnelle hypothèse pour un ensemble de variables  $y_1, \dots, x_n$  conditionnelles à une autre variable  $c$ , on obtient ce qui suit :

$$\mathbf{P}(\mathbf{x}\backslash c) = \prod_{i=1}^N \mathbf{P}(x_i\backslash c) \quad (4)$$

#### 5.4. Régression logistique

Comme la régression naïve de Bayes, la régression logistique [176] fonctionne en extrayant un ensemble de caractéristiques pondérées de l'entrée, en prenant des logs et en les combinant de manière linéaire, ce qui signifie que chaque caractéristique est multipliée par un poids et ensuite additionnée [160].

La différence la plus importante entre Bayes naïf et la régression logistique est que la régression logistique est un classificateur discriminatif alors que Bayes naïf est un classificateur génératif.

La régression logistique [177] est un type de régression qui prédit la probabilité d'occurrence d'un événement en ajustant les données à une fonction logistique. Comme de nombreuses formes d'analyse de régression, la régression logistique utilise plusieurs variables prédictives qui peuvent être numériques ou catégorielles.

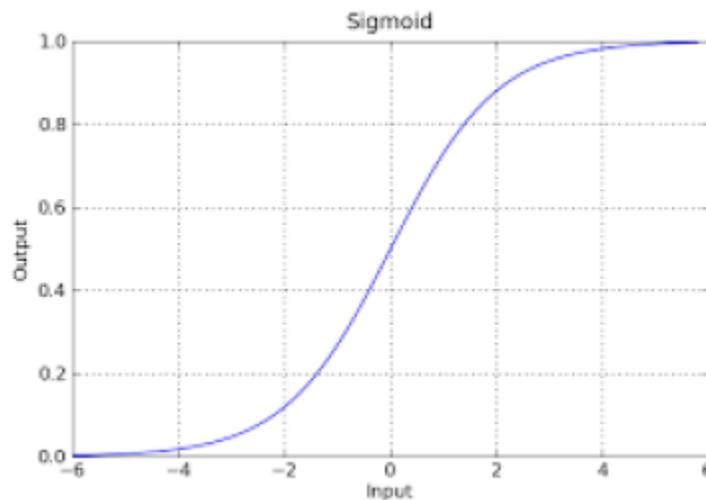
L'hypothèse de régression logistique est définie comme suit :

$$h_{\theta}(x) = g(\theta^T x) \quad (5)$$

Où la fonction  $g$  est une fonction sigmoïde définie comme suit :

$$g(z) = \frac{1}{1 + e^{-z}} \quad (6)$$

La fonction sigmoïde a des propriétés particulières qui font que les valeurs se situent dans l'intervalle  $[0,1]$ , visualisée sur la **figure 3.3**.



**Figure 3.3-** Représentation visuelle de la fonction logistique [23].

La fonction de coût pour la régression logistique est donnée comme suit :

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \left[ -y^{(i)} \log(h_{\theta}(x^{(i)})) - (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)})) \right] \quad (7)$$

### 5.5. Machine à vecteur de support

La méthode SVM est une méthode puissante pour construire un classificateur. Il vise à créer une frontière de décision entre deux classes qui permet la prédiction d'étiquettes à partir d'un ou plusieurs vecteurs de caractéristiques [179]. Cette frontière de décision, connue sous le nom de hyperplan, est orientée de manière à être aussi éloignée que possible des points de données les plus proches de chacune des classes. Ces points les plus proches sont appelés vecteurs de support [180].

Étant donné un ensemble de données d'apprentissage étiquetées :

$$(x_1, y_1), \dots, (x_n, y_n), x_i \in \mathbf{R}^d \text{ et } y_i \in (-1, +1) \quad (8)$$

où  $x_i$  est une représentation du vecteur de caractéristiques et  $y_i$  l'étiquette de classe (négative ou positive) d'un composant d'apprentissage  $i$ .

L'hyperplan optimal peut alors être défini comme suit :

$$wx^T + b = 0 \quad (9)$$

où  $w$  est le vecteur de poids,  $x$  est le vecteur de caractéristiques d'entrée, et  $b$  est le biais.

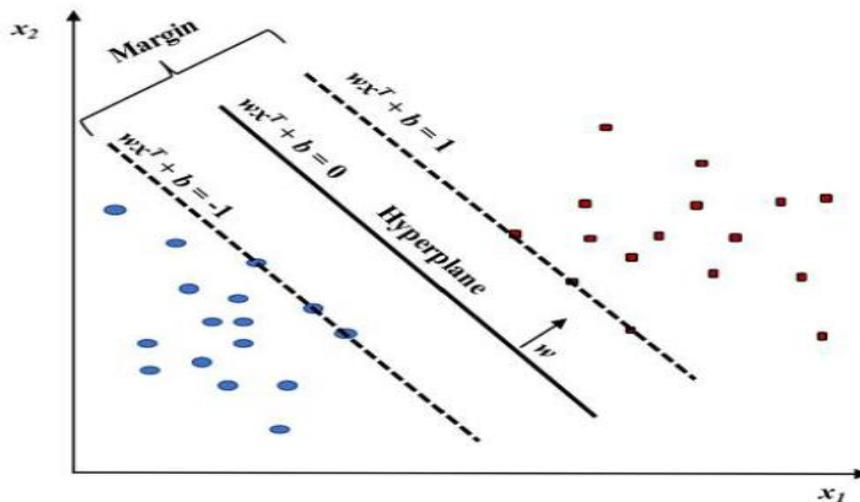
$w$  et  $b$  doivent satisfaire les inégalités suivantes pour tous les éléments de l'ensemble d'apprentissage :

$$wx_i^T + b \geq +1 \text{ if } y_i = 1 \quad (10)$$

$$wx_i^T + b \leq -1 \text{ if } y_i = -1$$

L'objectif de l'apprentissage d'un modèle SVM est de trouver les valeurs  $w$  et  $b$  de façon à ce que l'hyperplan sépare les données et maximise la marge  $1 / \|w\|^2$  [180].

Les vecteurs  $x_i$  pour lesquels  $|y_i|(wx_i^T + b)$  sont appelés vecteurs de support (**figure 3.4**).



**Figure 3.4-** Modèle SVM linéaire. Deux classes (rouge contre bleu) ont été classées [180].

## **6. L'apprentissage profond vs l'apprentissage automatique**

Aujourd'hui, les systèmes intelligents qui offrent des capacités d'intelligence artificielle reposent souvent sur l'apprentissage automatique. L'apprentissage automatique décrit la capacité des systèmes à apprendre à partir de données d'apprentissage spécifiques à un problème afin d'automatiser le processus d'élaboration de modèles analytiques et de résoudre les tâches associées. L'apprentissage profond ou Deep Learning (DL) est un concept d'apprentissage automatique basé sur les réseaux neuronaux artificiels. Les chercheurs ont déployé de nombreux efforts pour améliorer la précision de ces algorithmes d'apprentissage automatique, mais une autre dimension a été prise en compte, ce qui a conduit au concept d'apprentissage profond. L'apprentissage profond est un sous-ensemble de l'apprentissage automatique. Jusqu'à présent, peu d'applications de l'apprentissage profond ont été explorées. L'apprentissage profond va certainement permettre de résoudre des problèmes dans plusieurs nouveaux domaines et sous-domaines d'application [156]. Pour de nombreuses applications, les modèles d'apprentissage profond sont plus performants que les modèles d'apprentissage automatique peu profonds et les approches traditionnelles d'analyse des données [181].

Au cours des dernières décennies, le domaine du ML a donné lieu à une variété de progrès remarquables dans les algorithmes d'apprentissage sophistiqués et les techniques de prétraitement efficaces. L'une de ces avancées a été l'évolution des réseaux de neurones artificiels (ANN) vers des architectures de réseaux de neurones de plus en plus profondes avec des capacités d'apprentissage améliorées, résumées sous le nom d'apprentissage profond (DL) (Goodfellow et al. 2016 ; LeCun et al. 2015). Pour des applications spécifiques dans des environnements fermés, le DL montre déjà des performances surhumaines en surpassant les capacités humaines (Madani et al. 2018 ; Silver et al. 2018). Cependant, ces avantages ont aussi un prix, car il y a plusieurs challenges à surmonter pour mettre en oeuvre avec succès des modèles analytiques dans des contextes commerciaux réels. Il s'agit notamment du choix approprié parmi de nombreuses options de mise en oeuvre, du biais et de la dérive des données, de l'atténuation des propriétés de la boîte noire et de la réutilisation de modèles préconfigurés [181].

### **7. Vue d'ensemble de l'apprentissage profond**

L'apprentissage profond est un sous-ensemble de l'apprentissage automatique. Il s'agit d'un réseau neuronal comportant un grand nombre de couches et de paramètres. La plupart des méthodes d'apprentissage profond utilisent des architectures de réseaux neuronaux. C'est pourquoi on parle également de réseaux neuronaux profonds [157].

En bref, l'apprentissage profond utilise une cascade de plusieurs couches d'unités de traitement non linéaires pour l'extraction et la transformation des caractéristiques. Les couches inférieures proches de l'entrée des données apprennent des caractéristiques simples, tandis que les couches supérieures apprennent des caractéristiques plus complexes dérivées des caractéristiques des couches inférieures. Cette architecture forme une représentation hiérarchique et puissante des caractéristiques. Cela signifie que l'apprentissage profond est adapté à l'analyse et à l'extraction de connaissances utiles à partir de grandes quantités de données et de données collectées à partir de sources différentes [182].

Les principales raisons de la popularité de l'apprentissage profond aujourd'hui sont l'augmentation radicale des capacités de traitement des puces (par exemple, les unités GPU), le faible coût du matériel informatique et les progrès récents de la recherche sur l'apprentissage automatique et le traitement du signal et de l'information [15].

### **8. Les Applications de l'apprentissage profond**

Quelques applications les plus récentes de l'apprentissage profond sont décrites dans les exemples ci-dessous :

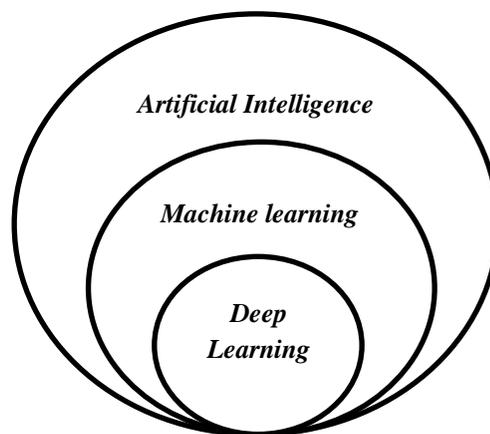
- Un exemple d'application de l'apprentissage profond dans le domaine du Big Data est la reconnaissance vocale de Microsoft (MAVIS). Utilisant l'apprentissage profond permet de rechercher des fichiers audio et vidéo à partir de voix et de discours humains [16].
- L'apprentissage profond dans un environnement Big Data est utilisé par Google pour son service de recherche d'images. Ils ont utilisé l'apprentissage profond pour comprendre les images afin qu'elles puissent être utilisées pour l'annotation et le marquage d'images qui sont

ensuite utiles dans les moteurs de recherche d'images et la récupération d'images ainsi que l'indexation d'images [16].

- En 2016, le programme AlphaGo de Google a battu Lee Sedol dans la compétition de Go, ce qui a montré que l'apprentissage profond avait une forte capacité d'apprentissage.
- Deep Dream de Google est un logiciel qui peut non seulement classer des images mais aussi générer des peintures étranges et artificielles sur la base de ses propres connaissances.
- Facebook a annoncé un nouveau système d'intelligence artificielle nommé Deep Text. Il s'agit d'un moteur de compréhension de texte basé sur l'apprentissage profond qui peut classer des quantités massives de données, fournir des services correspondants pour identifier les messages de chat des utilisateurs et nettoyer les messages de spam.
- La vision par ordinateur, la prédiction, l'analyse sémantique, le traitement du langage naturel, la recherche d'informations et la gestion de la relation client sont les domaines d'application de l'apprentissage profond.
- La vision par ordinateur comprend la reconnaissance d'objets, la détection d'objets et le traitement d'objets, ainsi que la reconnaissance automatique de la parole, la reconnaissance d'images, le traitement de la parole et du son et le traitement des arts visuels des restrictions s'appliquent dont de nouvelles applications sont explorées à l'aide de l'apprentissage profond.
- La prédiction : Les différents sous-domaines sont la classification, l'analyse et la recommandation. La découverte de médicaments et la toxicologie, la bioinformatique, la publicité mobile sont des applications plus récentes développées à l'aide de l'apprentissage profond.
- l'analyse sémantique, le traitement du langage naturel et la recherche d'informations : Il s'agit de trois domaines dans lesquels les techniques d'apprentissage automatique ont été explorées dans le passé, mais les chercheurs appliquent désormais des techniques d'apprentissage profond.
- Gestion de la relation client : Compte tenu de l'historique du client, on procède à une analyse des données qui est utilisée pour améliorer les relations commerciales. Dans ce cas, l'apprentissage profond peut être utile.

### **9. L'apprentissage profond (Deep Learning)**

Les informaticiens ont déployé des efforts depuis les années 1950 dans le domaine de l'apprentissage automatique. Cela conduit à des attentes plus élevées de la part des machines. L'apprentissage profond est une tentative dans ce sens. Il s'agit d'un sous-ensemble de l'apprentissage automatique. Comme le travail dans l'apprentissage est mis en avant dans de nombreux nouveaux domaines et l'applicabilité des nouveaux domaines est toujours une tâche en cours dans la communauté de recherche. La **figure 3.5** illustre les recherches antérieures qui ont donné naissance à l'apprentissage profond. L'apprentissage profond fait référence aux réseaux neuronaux artificiels profonds. Le terme "profond" fait référence à un certain nombre de couches dans un réseau neuronal. Le réseau profond possède plus d'une couche cachée, alors qu'un réseau peu profond n'en possède qu'une seule[157].

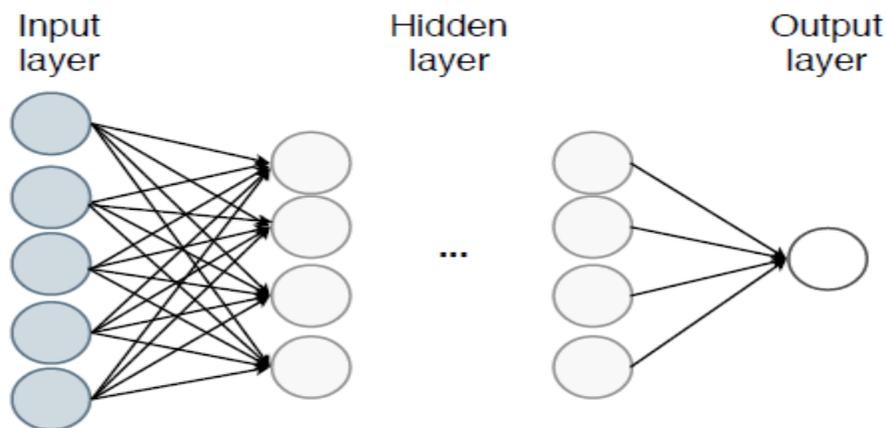


**Figure 3.5**-la naissance de l'apprentissage profond

#### **9.1. Les réseaux neuronaux artificiels ANN :**

Les réseaux neuronaux artificiels (ANN), communément appelés réseaux de neurones, sont des modèles inspirés du système neuronal biologique. Bien que les ANN s'inspirent du système nerveux de l'homme, ils sont loin d'en être une véritable réplique. Le réseau se compose le plus souvent d'une couche d'entrée, d'une ou plusieurs couches cachées et d'une

couche de sortie. La ou les couches cachées sont constituées d'une collection finie de neurones artificiels positionnés en couches qui peuvent être comparées aux neurones d'un cerveau biologique [185]. Chaque neurone d'une couche cachée ou d'une couche de sortie reçoit une entrée d'un ou de plusieurs neurones d'une couche connectée précédente. En fonction de l'entrée provenant du ou des neurones, le ou les neurones actuels peuvent prendre diverses décisions. Un algorithme appelé propagation vers l'avant est utilisé pour, à partir d'une entrée, produire une sortie. L'algorithme propage l'entrée à travers les couches cachées jusqu'à la couche de sortie qui produit une sortie [186]. Plus le nombre de couches cachées utilisées est élevé, ce qui donne un réseau neuronal plus profond, plus le réseau peut prendre des décisions complexes [185]. La **figure 3.6** ci-dessous illustre un réseau typique.



**Figure 3.6**-Modèle simple de la façon dont les neurones artificiels sont connectés entre les différentes couches d'un réseau neuronal à action direct [187]

## 9.2. Réseau neuronal convolutif (CNN :Convolutional neural network)

Dans cette section, la théorie des CNNs sera expliquée. Elle comprend les propriétés mathématiques des blocs de construction qui sont les couches convolutionnelles, les fonctions d'activation, les couches de mise en commun et les couches entièrement connectées. Les réseaux neuronaux convolutifs sont une sous-classe des réseaux neuronaux profonds (DNN), où les réseaux neuronaux profonds sont des réseaux neuronaux avec plusieurs couches cachées pour extraire des caractéristiques d'une entrée. Les réseaux convolutifs sont le plus

souvent utilisés lorsque les données présentent une structure en grille. Un exemple de structure en grille est une image dans laquelle l'image elle-même est constituée de pixels représentés sur une grille ou un tableau bidimensionnel. Les CNNs se sont avérés efficaces pour des tâches liées à l'analyse/classification/segmentation d'images [186].

### 9.2.1. Blocs de construction

L'architecture des CNN varie en fonction, par exemple, du résultat souhaité ou de la complexité de la tâche. Des blocs de construction sont utilisés pour façonner différentes architectures où les blocs utilisés sont des couches convolutionnelles, des couches de mise en commun, des fonctions d'activation ou des couches entièrement connectées. Pour mieux comprendre comment ces blocs fonctionnent et ce qu'ils font, un résumé est donné dans cette section [187].

#### 9.2.1.1. Couche convolutive

La couche convolutionnelle d'un CNN est l'endroit où une convolution entre une entrée et un noyau (filtre contenant des poids) a lieu afin de produire une sortie. Cette sortie est souvent appelée carte de caractéristiques. La couche convolutionnelle est utilisée pour extraire les caractéristiques d'une entrée et créer une carte de caractéristiques à partir de celles-ci [188]. Pour une image donnée, la couche de convolution peut extraire des caractéristiques concernant, par exemple, chacune des différentes régions de l'œil. Dans le domaine de l'apprentissage automatique, il est courant d'utiliser des tableaux multidimensionnels, généralement appelés tenseurs, plutôt que des tableaux unidimensionnels [186]. Par exemple, si l'on prend en entrée une image constituée d'un tableau bidimensionnel et que le noyau doit également être un tableau bidimensionnel, la convolution consiste donc en deux sommes au lieu d'une seule. La formule pour calculer les cartes de caractéristiques est donnée comme suit :

$$S(i, j) = (K * D)(i, j) = \sum_m \sum_n I(i + m, j + n)K(m, n) + b \quad (11)$$

où  $S$  est la carte des caractéristiques,  $I$  est l'entrée,  $K$  est le noyau et  $b$  est le biais [186]. Lorsqu'une entrée est traitée par une couche de convolution, la sortie sera constituée de nombreuses cartes de caractéristiques empilées les unes sur les autres, où chaque carte de

caractéristiques est basée sur l'équation (11) . Cela est dû au fait que la couche de convolution peut être constituée de plusieurs filtres différents et qu'elle produit donc un nombre égal de cartes de caractéristiques. La sortie consiste en un tableau 3D (largeur, hauteur, profondeur). La sortie est ensuite envoyée à travers une fonction d'activation qui sera traitée dans la section sur les fonctions d'activation. Se distinguant d'un réseau neuronal entièrement connecté, la couche convolutive d'un CNN peut aider un système d'apprentissage automatique en raison de ses interactions éparses et du partage des paramètres [186]. Typiquement, dans les réseaux de neurones entièrement connectés, toutes les sorties sont connectées à toutes les entrées dans les couches correspondantes, ce qui produit des interactions non éparses. Cependant, dans les CNN, ce n'est pas le cas car toutes les sorties ne sont pas connectées à toutes les entrées. Ceci est réalisé en choisissant un noyau qui a une taille plus petite que l'entrée [186]. Comme le noyau est convolué de manière itérative sur les axes d'entrée, chaque pixel est utilisé plusieurs fois dans la même convolution. Le facteur déterminant combien de fois un pixel peut être utilisé dans une convolution dépend de la taille du noyau ainsi que du facteur de stride. Le facteur de stride détermine le nombre de pixels que le noyau déplace à chaque fois qu'il se déplace sur l'image. Par exemple, si le facteur de stride est fixé à un, le noyau se déplacera d'un pas (pixel). En raison de la diminution du nombre de paramètres, par rapport aux couches entièrement connectées, les couches convolutionnelles entraînent moins d'opérations et réduisent la quantité de mémoire nécessaire pour stocker le modèle[187].

#### **9.2.1.2. Fonction d'activation**

Une fonction d'activation est une fonction non linéaire qui, à partir d'entrées provenant d'autres neurones ou de l'image d'entrée, peut déterminer une sortie de ce neurone [186]. L'étendue de la sortie peut varier en fonction de la fonction d'activation utilisée. La fonction d'activation est non linéaire pour permettre des mappings non linéaires entre l'entrée et la sortie. Si une fonction d'activation linéaire était utilisée, le modèle serait limité aux transformations linéaires, et ne pourrait donc pas représenter toutes les transformations possibles [186]. La fonction d'activation joue un rôle très important lors de l'apprentissage du réseau car, selon la fonction d'activation utilisée, le réseau peut donner des résultats finaux différents. Pour calculer l'activation de chaque neurone, les cartes de caractéristiques sont

passées par une fonction d'activation selon l'équation (12)

$$\mathbf{h}_k = \mathbf{g}(\mathbf{S}_k) \quad (12)$$

où  $\mathbf{S}$  désigne des cartes de caractéristiques multiples à chaque couche de convolution,  $k$  est un indice spécifiant la profondeur et  $\mathbf{g}$  est la fonction d'activation. Chaque carte de caractéristiques  $\mathbf{S}_k$  est constituée d'une entrée, d'une matrice de poids et d'un terme de biais comme spécifié dans l'équation (11) [186].

Il existe plusieurs fonctions d'activation, chacune d'entre elles pouvant être plus ou moins performante en fonction de la tâche à accomplir. Deux fonctions d'activation couramment utilisées dans les problèmes d'apprentissage automatique actuels sont la tangente hyperbolique ( $\tanh$ ) et la fonction sigmoïde. Bien qu'aucune de ces fonctions d'activation ne soit recommandée dans les réseaux neuronaux de nos jours. Au lieu de cela, l'unité linéaire rectifiée (ReLU) est actuellement la fonction d'activation la plus recommandée [186]. La fonction ReLU est définie comme suit

$$f(\mathbf{x}) = \max(\mathbf{0}, \mathbf{x}) \Rightarrow f'(\mathbf{x}) = \begin{cases} \mathbf{x}, & \text{if } \mathbf{x} > \mathbf{0}, \\ \mathbf{0}, & \text{otherwise,} \end{cases} \quad (13)$$

où  $\mathbf{x}$  est l'entrée d'un neurone spécifique. L'une des raisons pour lesquelles ReLU est populaire est qu'elle est facile à optimiser en raison de la simplicité de ses calculs. Comme la sortie est soit zéro, soit  $\mathbf{x}$  (entrée), elle est donc moins coûteuse en calcul que la fonction  $\tanh$  ou sigmoïde. Un autre avantage est que la fonction ReLU atteint la sparsité [189]. La sparsité signifie que seuls quelques neurones sont activés pour une entrée donnée, ce qui signifie que moins de puissance de calcul est nécessaire. Cela est dû au fait que la fonction ReLU peut avoir une énergie d'activation de zéro, sautant ainsi essentiellement ce neurone, ce qui génère à son tour une véritable représentation clairsemée [189].

Bien que ReLU présente de nombreux avantages, elle a également quelques inconvénients. L'un d'entre eux est le problème de ReLU mourant, où essentiellement les neurones " meurent " en raison du gradient qui ne circule pas à travers le neurone [190]. Lorsque le gradient devient suffisamment proche ou égal à zéro, le poids des neurones concernés cesse de changer, ce qui entraîne un neurone inactif ou, en d'autres termes, la mort d'un neurone [190].

Une façon de résoudre ce problème est d'utiliser le ReLU fuyant à la place. Le leaky ReLU est défini comme suit :

$$f(x) = \begin{cases} x & \text{if } x > 0, \\ 0.01x, & \text{otherwise.} \end{cases} \quad (14)$$

Cela permet d'obtenir des gradients non nuls, ce qui empêche la disparition du gradient. Cependant, en utilisant le ReLU fuyant, la même sparsité ne sera pas atteinte que dans le cas du ReLU original, car le domaine ne contient que des gradients non nuls [191].

Une autre fonction d'activation clé couramment utilisée dans la segmentation sémantique est la fonction softmax. Avec les classificateurs de réseaux neuronaux, la fonction softmax est utilisée pour acquérir une distribution de probabilité sur un ensemble de classes différentes. En tant que telle, elle peut être utilisée pour choisir une seule classe de pixels parmi plusieurs à l'intérieur du réseau [186]. La couche finale d'un classificateur de réseau neuronal utilise généralement la fonction d'activation softmax pour effectuer la sélection mentionnée précédemment. Elle est définie comme suit :

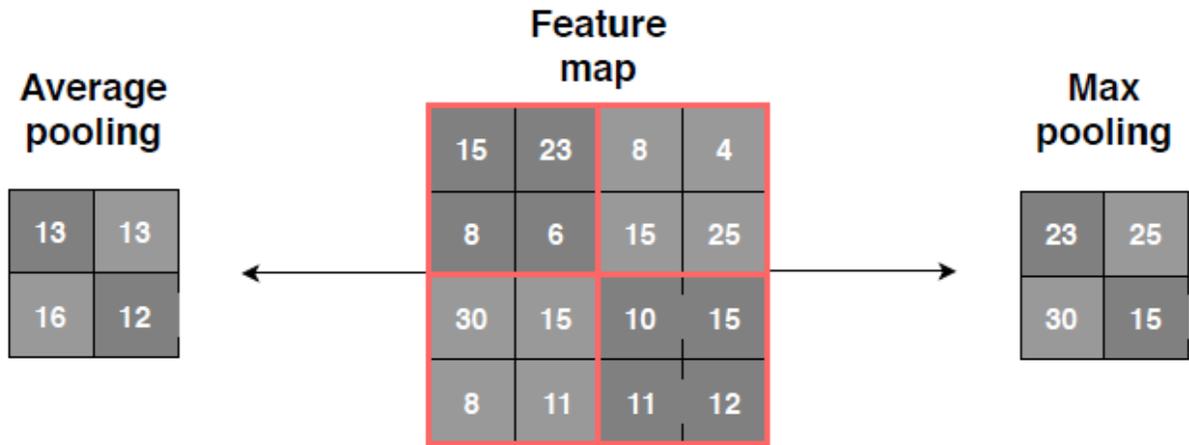
$$\text{Softmax}(x)_i = \frac{\exp(x_i)}{\sum_{j=1}^k \exp(x_j)} \quad (15)$$

où  $x$  est l'entrée dans la fonction d'activation,  $i$  est l'indice de la classe et  $K$  est le nombre total de classes [187].

### 9.2.1.3. Couche de mise en commun ( pooling layer)

Les couches de mise en commun sont utilisées pour réduire la taille de l'entrée. Ceci est réalisé en réduisant la résolution des cartes de caractéristiques après les couches convolutionnelles. Les couches de mise en commun sont donc communément appelées couches de sous-échantillonnage [192]. L'opération a principalement deux résultats importants : elle permet d'éviter une complexité de calcul élevée et agit comme une technique de régularisation pour éviter un ajustement excessif [193]. Les techniques de regroupement les plus courantes sont une moyenne ou un maximum [194]. La **figure 3.7** présente un

exemple des noyaux obtenus après les opérations de regroupement sur une carte de caractéristiques aléatoires. Ces noyaux sont ensuite transmis à la couche suivante.



**Figure 3.7** - Les cartes résultantes après le regroupement max et average sur une carte de caractéristiques [187].

Comme on peut le voir sur la **figure 3.7**, l'opération de regroupement maximal prend la valeur maximale de chaque pool-kernel (noyau de regroupement) (cases rouges sur la **figure 3.7**). L'opération de moyennage prend la valeur moyenne de chaque pool-kernel. La largeur et la hauteur du pool-kernel dépendent de la réduction souhaitée de la largeur et de la hauteur de la carte de caractéristiques. Si l'on utilise un noyau de taille 2\*2, on obtient une réduction de taille de 2. Si l'on souhaite une réduction plus importante, il faut utiliser un pool-kernel plus grand ou une longueur de pas plus grande. La longueur du pas, en d'autres termes le nombre de pas que le noyau est déplacé entre chaque opération, a le plus souvent la même longueur que la dimension du noyau [187].

#### 9.2.1.4. Couches entièrement connectées (ou couches denses)

Les couches denses, ou couches entièrement connectées, sont généralement placées en dernier dans un CNN et sont utilisées, par exemple, dans les tâches de classification d'images. Le nombre de couches denses peut varier, c'est-à-dire qu'il peut y en avoir une ou plusieurs en

fonction de la tâche. Cependant, comme une couche dense attend une taille fixe en entrée, elle ne peut pas traiter des dimensions d'image arbitraires avec le même réseau [195].

### **9.3. Apprentissage d'un réseau neuronal CNN**

L'entraînement vise à trouver les poids optimaux du réseau pour la tâche à accomplir. En ayant la "meilleure adaptation" des poids, le réseau devrait être capable de produire de bons résultats pour un problème donné. Dans cette section, les étapes importantes avant, pendant et après l'apprentissage seront abordées, ainsi que les problèmes qui peuvent survenir [187].

#### **9.3.1. Pré- apprentissage (Pre -training)**

Avant l'entraînement d'un réseau neuronal, il faut prendre en compte plusieurs éléments, dont certains sont cruciaux pour l'apprentissage. La plus importante est celle des données et de leur traitement [187].

##### **9.3.1.1. Collecte et répartition des données**

Il est vital de valider que les données collectées par des processus automatisés ou par annotation manuelle sont correctement étiquetées. Dans le cas contraire, le réseau n'a pas un bon apprentissage. Une autre partie essentielle est le partitionnement des données. En général, trois ensembles de données sont utilisés : un ensemble d'apprentissage, un ensemble de validation et un ensemble de test. Comme le réseau ne doit pas être biaisé, il est important de partager les images aléatoirement entre les trois ensembles. Si le réseau a déjà été exposé à certaines images lors de l'apprentissage, une évaluation des performances du réseau peut ne pas être précise en raison du biais mentionné précédemment, si elle contient les mêmes images. Selon la quantité de données existantes, le rapport entre les ensembles de données peut varier. Cependant, comme le suggère la littérature, une règle empirique est de garder le ratio pour l'apprentissage autour de 70-80% et le reste pour la validation et le test. Enfin, l'augmentation des données est un moyen d'accroître l'ensemble de données et de rendre le réseau plus robuste au bruit [196].

##### **9.3.2.2. Apprentissage par transfert**

L'apprentissage par transfert est un concept d'apprentissage automatique qui consiste à transférer les connaissances déjà acquises pour une tâche similaire, où des caractéristiques

similaires ont déjà été apprises, et à incorporer ces connaissances dans une nouvelle tâche. Cela signifie que les filtres déjà entraînés peuvent être réutilisés dans le nouveau réseau. Plus les tâches sont similaires les unes aux autres, mieux c'est [197].

La manière d'appliquer l'apprentissage par transfert pour les CNN consiste à utiliser les couches et leurs poids d'un modèle déjà formé et à modifier ou ajouter des couches supplémentaires si nécessaire. Les premières couches sont utilisées pour extraire des caractéristiques non abstraites plus générales telles que des lignes ou des gouttes, etc., tandis que les dernières couches extraient des caractéristiques plus abstraites spécifiques à la tâche [198]. Par conséquent, l'utilisation d'un modèle déjà entraîné, tel que celui proposé dans [199], qui a été entraîné sur une base de données d'images appelée ImageNet [200], puis de l'affiner pour l'adapter à une nouvelle tâche, devrait accélérer l'apprentissage et permettre d'obtenir une précision élevée. Comme le réseau n'a pas à repartir de zéro lors de l'apprentissage, il peut éventuellement apprendre plus rapidement les caractéristiques spécifiques à la tâche [197]. Une autre raison d'utiliser l'apprentissage par transfert est le fait de ne pas disposer d'une quantité suffisante de données pour former complètement un réseau à une nouvelle tâche [201].

### **9.3.2. L'étape d'apprentissage**

Pendant l'apprentissage, les paramètres entraîlables (poids) tentent de s'adapter aux données afin de faire de meilleures prédictions. Pour optimiser l'apprentissage d'un réseau, de multiples facteurs doivent être pris en compte et décidés avant le début de l'apprentissage proprement dite.

L'apprentissage d'un CNN est un problème d'optimisation et nécessite donc une méthode d'optimisation. Les méthodes d'optimisation courantes sont généralement la descente de gradient stochastique ou la descente de gradient en mini-batch. La descente de gradient par mini-lots (mini batch) est une combinaison de la descente de gradient classique et de la descente de gradient stochastique, où elle converge avec plus de stabilité par rapport à la méthode stochastique. Cela est dû au fait que la mise à jour ne se fait pas avec un seul exemple, mais plutôt avec un lot d'exemples, ce qui entraîne moins d'oscillations autour des

minima locaux. Elle est également plus rapide que la méthode classique, car elle ne nécessite pas le chargement de l'ensemble des données dans la mémoire [202].

Le taux d'apprentissage est un facteur d'échelle pour la descente de gradient qui détermine de combien les poids doivent changer à chaque itération par rapport à une perte de gradient. En augmentant le taux d'apprentissage, l'algorithme peut avoir une convergence plus rapide. Cependant, l'algorithme peut ne pas trouver les minima optimaux ou ne pas converger du tout. Un taux d'apprentissage plus faible entraîne une convergence plus lente, mais avec des minima plus précis. Il est donc préférable d'avoir un taux d'apprentissage adaptatif, où il peut commencer avec une valeur plus grande et ensuite diminuer à chaque étape jusqu'à une valeur minimale fixe [203].

Une autre méthode d'optimisation courante et largement utilisée est Adam. Adam utilise l'estimation du moment adaptatif des gradients pour calculer les taux d'apprentissage adaptatifs pour chaque paramètre individuel [204]. Il utilise le momentum pour espérer converger plus rapidement par rapport à la descente de gradient stochastique classique, bien qu'il existe des variantes de la descente de gradient stochastique qui incluent le momentum [202].

Un CNN est entraîné par époques, c'est-à-dire lorsque l'ensemble complet de données a été exposé et utilisé par le réseau. Le processus d'apprentissage proprement dit commence par la propagation directe d'une entrée à travers le réseau pour produire une sortie/prédiction. La sortie est ensuite comparée aux étiquettes de la vérité du terrain et une erreur peut être calculée entre les deux. L'erreur est déterminée par une fonction de perte et, pour la segmentation d'images, une fonction courante est l'entropie croisée. D'autres fonctions de perte sont le coefficient de Sørensen-Dice et l'indice de Jaccard tels que définis, mais non proposés, dans [205]. Même si les CNNs peuvent être constitués d'une quantité énorme de paramètres entraînaibles, l'algorithme de rétropropagation effectue efficacement les calculs des gradients, qui sont nécessaires à l'optimiseur [206]. Les paramètres entraînaibles sont alors mis à jour de manière appropriée et une nouvelle époque peut commencer.

L'adaptation excessive est un problème dans l'apprentissage automatique, où le réseau apprend à prédire l'ensemble de données d'apprentissage, mais lorsqu'il est exposé à de nouvelles données, non vues auparavant, c'est-à-dire des données de validation, il ne parvient pas à prédire correctement. Afin d'éviter le surajustement, des méthodes telles que l'abandon et l'augmentation des données (ce qui entraîne une augmentation de la quantité de données d'entraînement). [207] ainsi que l'arrêt anticipé [208] peuvent être utilisés.

L'ensemble de données de validation est utilisé pendant la phase d'apprentissage pour vérifier si le réseau apprend les bonnes caractéristiques ou non et pour minimiser le surajustement. Si, pendant l'apprentissage, la précision de l'apprentissage continue de s'améliorer alors que la précision de la validation a cessé d'augmenter ou diminue, le réseau est en surapprentissage. L'étape de validation est utilisée pour améliorer les évaluations pendant l'apprentissage, c'est-à-dire choisir l'ensemble de poids qui donne les meilleures performances [187].

### **9.3.3. Post-apprentissage (Post-training)**

Une fois la session d'apprentissage terminée, une session de test commence, qui vise à évaluer les performances du réseau. La session dispose d'un ensemble de données distinct qui ne fait pas partie des données d'apprentissage ni de validation afin d'éviter tout biais. Au cours de cette session, de multiples cas de test de différents degrés de difficulté seront évalués afin d'établir des statistiques concernant les performances du réseau. La mesure de performance globale, où l'ensemble des données de test est évalué, est un type de scénario de test. Un autre type de scénario consiste à tester un sous-ensemble de données de test dans lequel on n'inclure que des images correspondant à un scénario spécifique, par exemple uniquement des images comportant des lunettes. Grâce à ces mesures de performance, il est possible d'établir un consensus général sur les lacunes du réseau et de savoir ainsi ce qu'il faut améliorer. Les métriques communes dans la classification sont présentées dans la section 9.3.3.1[187].

#### **9.3.3.1. Métriques pour l'évaluation des performances du réseau**

Afin d'avoir une idée de la performance du réseau, différentes mesures de performance peuvent être utilisées. Les métriques couramment utilisées pour la segmentation sémantique

sont par exemple l'exactitude, la précision, le rappel, le score F1 (également connu sous le nom de coefficient de Dice) et l'intersection moyenne sur l'union (mIoU) (également connu sous le nom d'indice de Jaccard). Ces mesures sont basées sur les pixels correctement et faussement classés, où les expressions peuvent être configurées comme dans la matrice de confusion du **tableau 3.2** [187].

**Tableau 3.2** - Tableau présentant une matrice de confusion, utilisée comme représentation visuelle de la performance d'un algorithme.

		Classe prédite	
		Positif	Négatif
Véritable Classe	Positif	Vrai positif (TP)	Faux négatif (FN)
	Négatif	Faux positifs (FP)	Vrai négatif (TN)

Ce type de matrice de confusion peut être mis en place pour chaque classe où les "positifs" correspondent à la classe d'intérêt et les "négatifs" à tout le reste. Les vrais positifs correspondent aux pixels qui sont prédits comme étant la classe d'intérêt alors qu'ils le sont réellement. Les faux négatifs correspondent aux pixels étiquetés comme une autre classe alors qu'ils étaient en fait la classe d'intérêt. Les faux positifs correspondent aux pixels prédits comme classe d'intérêt alors qu'il s'agit en fait d'une autre classe. Enfin, les vrais négatifs correspondent aux pixels qui sont prédits comme une autre classe alors qu'ils sont en fait une autre classe. Le 'c' dans les équations (16), (17), (18) et (19) fait référence à la classe qui est évaluée [187].

La précision pour une classe spécifique, c, peut être calculée comme dans l'équation (16) ,

$$\text{Accuracy} = \frac{TP_c + TN_c}{TP_c + FN_c + TN_c + FP_c} \quad (16)$$

La précision est une mesure de la fraction de pixels correctement classés. La précision pour

une classe spécifique,  $c$ , peut être calculée comme dans l'équation (17) ,

$$\text{Précision} = \frac{TP_c}{TP_c + FP_c} \quad (17)$$

La précision est une mesure de l'exactitude de la prédiction de la classe d'intérêt. Le rappel pour une classe spécifique,  $c$ , peut être calculé comme dans l'équation (18) ,

$$\text{Recall} = \frac{TP_c}{TP_c + FN_c} \quad (18)$$

Le score F1 est une métrique utilisée pour combiner la précision et le rappel. Ceci est fait car le rappel et la précision peuvent être déséquilibrés eux-mêmes. Ce déséquilibre peut être observé si, par exemple, le réseau prédit que tous les pixels appartiennent à la même classe. Il en résulterait une précision très faible mais un rappel parfait pour cette classe. Le calcul du score F1 est expliqué dans l'équation (19) [187]. ,

$$\text{F1-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (19)$$

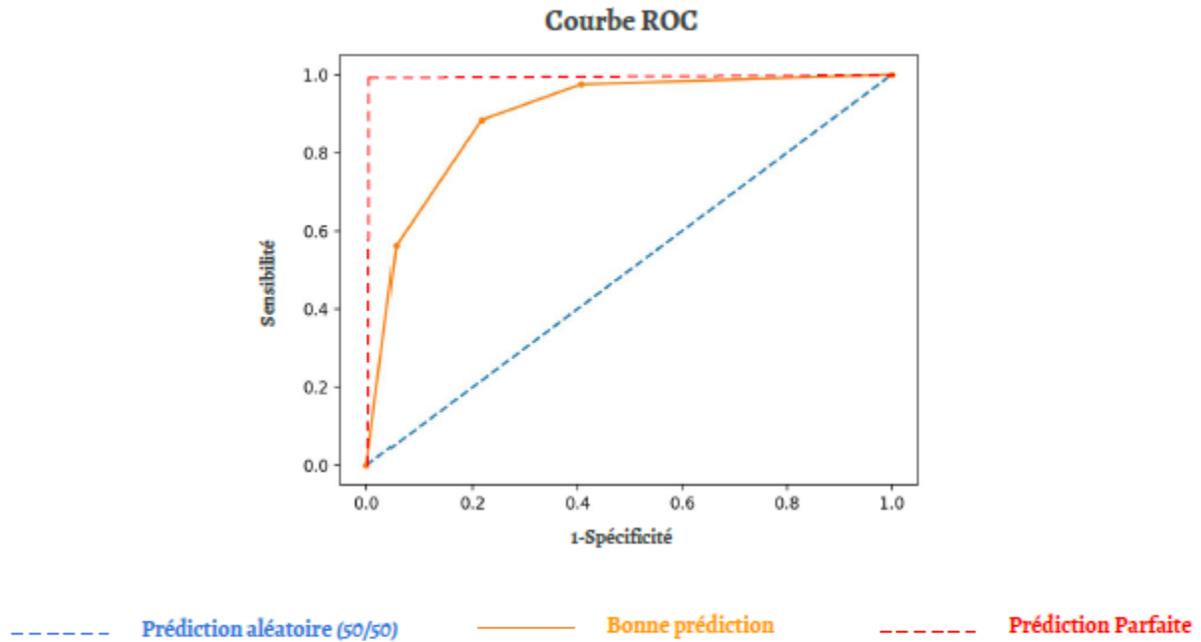
L'intersection-sur-union moyenne est une mesure du pourcentage moyen d'intersection entre les masques de vérité et les masques prédits. En d'autres termes, le nombre de pixels qui sont correctement classés par rapport au nombre total de pixels. Le mode de calcul de la métrique est illustré par l'équation (20) ,

$$mIoU = \frac{1}{|classes|} \sum_{x \in classes} \frac{TP_c}{TP_c + FP_c + FN_c} \quad (20)$$

où  $classes$  est l'ensemble des classes et  $|classes|$  correspond au nombre total de classes dans

l'ensemble [187].

Chacune de ces mesures représente un pourcentage qui décrit une sorte d'adéquation du modèle. Plus la valeur en pourcentage est élevée, meilleures sont les performances du modèle. La sensibilité et la spécificité ne prennent en compte que les échantillons de la même catégorie de test (la sensibilité est positive, la spécificité est négative). Par conséquent, les modifications du nombre d'images de test par classe n'ont aucun effet sur ces métriques. Cependant, ce n'est pas le même cas pour la précision et l'accuracy. En effet, la précision prend en compte des échantillons de test où les deux classes sont déterminées comme étant positives, et l'accuracy est une évaluation "globale" du modèle, prenant en compte toutes les prédictions (l'ensemble de la matrice de confusion). Par conséquent, si les données du test sont déséquilibrées, leur interprétation peut être erronée car elles ne reflètent pas la réalité de la situation. Par conséquent, l'interprétation des performances du classifieur ne peut se faire en exploitant les métriques proposées précédemment, qui sont complémentaires. Nous avons trouvé l'utilisation du couple de sensibilité/spécificité la **figure 3.8** dans la littérature, cette courbe permet de présentée graphiquement les performances d'un système biométrique. À partir de ce couple, une courbe ROC (Receiver Operating Characteristic) est tracée, qui peut résumer certains compromis entre le taux de faux positifs et le taux de faux négatifs en utilisant différents seuils. Ces seuils appliqués aux scores de prédiction permettent d'ajuster ces métriques en considérant les prédictions comme correctes uniquement si leurs scores sont supérieurs à ce seuil. La forme de la courbe et l'aire d'un point sous la courbe déterminent les performances du modèle biométrique. La courbe ROC est appropriée dans le cas de données de test équilibrées sur chaque classe [42].



**Figure 3.8** – Analyse des courbe ROC [42]

## 10. Etat de l'art des techniques de classification utilisées dans la reconnaissance de l'iris :

La phase de classification est la dernière phase du système de reconnaissance (RS). Le but de la classification est de mesurer le niveau de ressemblance entre les échantillons de test et les échantillons des bases de données des images d'iris. Souvent, la correspondance complète de ces échantillons n'est pas possible. Par conséquent, le taux approximatif de chaque échantillon est utilisé pour aider le système de reconnaissance à identifier les personnes. Dans cet thèse, les techniques de classification peuvent être classées en deux sections : la classification à l'aide de techniques traditionnelles et la classification à l'aide de techniques d'apprentissage profond [60].

### 10.1. Classification à l'aide de techniques traditionnelles :

L'efficacité de l'apprentissage des SVM reste un goulot d'étranglement, notamment pour les problèmes d'apprentissage à grande échelle. Pour améliorer un algorithme d'entraînement

rapide pour les SVM afin de résoudre divers problèmes, Gu et al. ont proposé une technique pour améliorer les SVM afin de classifier les caractéristiques qui ont été extraites de la région de l'iris. Pour rendre le SVM plus efficace dans les applications embarquées, ils ont présenté quelques optimisations pour augmenter la précision de performance du SVM pour IRS et garder sa haute performance de classification. La méthode suggérée transforme un problème donné de classification multi-classes à deux classes [209].

Rai et Yadav [210] ont en revanche proposé une méthode de classification en cascade. Dans cette méthode, des ondelettes de Haar (HWs) ont été appliquées pour entraîner des méthodes SVM pour chaque classe. En général, bien que cette méthode de classification en cascade ait de grandes propriétés de performance, elle prend plus de temps car c'est un classifieur séquentiel [210]. Salve et Narote ont introduit une analyse qualitative sur deux méthodes différentes, à savoir le SVM et le réseau neuronal artificiel (ANN), pour classer les caractéristiques qui ont été extraites de la texture de la région de l'iris avec la plus grande précision. Il s'avère que le SVM fonctionne mieux dans la classification [108].

Pour faire correspondre ou classer deux modèles d'iris, Raffei et al. [104, 102], Kim et al. [211], Yahiaoui et al. [84], Othman et al. [114], Gangwar et Joshi [212], Zhao et Kumar [154, 213], Zhou et al. [214], et Nazmdeh et al. [139] ont appliqué la technique de la distance de Hamming (HD) qui a été utilisée par Masek [124], A'lvarez Mariño et al. [215], et Daugman [139], [130]. Pour classer les caractéristiques de l'image de l'iris, l'échantillon enregistré a été mis en correspondance avec l'échantillon sauvegardé dans la base de données des images de l'iris.

Kumar et al. [69], Gad et al. [76], Ahuja et al. [110], Gangwar et Joshi [212], Lee et al. [155], Reddy et al. [151], et Dhage et al. [216] ont appliqué la distance euclidienne (ED) pour classer les caractéristiques qui ont été extraites de la texture de la région de l'iris dans la phase d'extraction des caractéristiques.

Une autre méthode qui a été utilisée pour classer les modèles à deux classes pour chaque sous-image est le k-nearest neighbor (k-NN). La méthode k-NN introduite par Elgamal et Al-Biqami [217] a été appliqué par Badejo et al [218] et Xia et al [219] comme classifier en raison de sa facilité de mise en œuvre, de sa faible complexité de calcul et de l'absence de

paramètres de réglage. De plus, le classifieur k-NN fonctionne bien avec des bases de données énormes. Liu et al. ont introduit une méthode coopérative collaborative pour classifier la texture de l'iris [101].

Diverses méthodes de classification et travaillant sur des modèles hybrides différents pour classer les images d'iris présentant plus de challenges tels que les sujets non coopératifs, le flou dû au mouvement, les tailles différentes de l'iris, la réflexion due aux lunettes, et la longue distance entre la caméra et les sujets [59].

### **10.2. Classification à l'aide de techniques d'apprentissage profond :**

L'approche d'apprentissage profond a gagné beaucoup d'attention parmi les chercheurs pour fournir d'excellentes solutions pour la reconnaissance de l'iris. Afin d'optimiser le paramètre de classification. Afin d'effectuer la tâche de classification, Zhang et al. [91] et Al-Waisy et al [129] ont utilisé les couches supérieures d'une technique CNN proposée qui sont deux couches FC. Ensuite, la sortie de la dernière couche FC est introduite dans le classifieur softmax [91, 129]. Afin de classifier le sexe en utilisant une image d'iris, Tapia et Aravena [277] ont proposé deux structures CNN différentes. Cette méthode a donné de bons résultats même avec un jeu de données plus petit et un modèle simple. Cependant, une amélioration est nécessaire pour augmenter l'efficacité de la classification [220]. Pour augmenter la précision de performance de l'IRS pour certains défis des images d'iris, Wang et al [221] ont introduit la méthode MiCoRe-Net qui est un réseau mixte neuronal convolutif et résiduel pour la tâche de reconnaissance de l'iris. La méthode MiCoRe-Net proposée a utilisé la couche SoftMax comme classifieur. La méthode proposée a montré diverses stratégies d'augmentation des données qui ont affecté l'efficacité des performances du système de reconnaissance et a montré que la technologie de reconnaissance de l'iris dépendait d'une stratégie d'augmentation appropriée pour fonctionner correctement [221]. Plus encore, Liu et al [120], Zhao et al [222], et Lozej et al [223] ont utilisé la couche softmax d'un CNN pour classifier le motif de l'iris [120, 222, 223]. Pour les classifications d'images d'iris ethniques asiatiques et non asiatiques, Lu et Pan [224] ont proposé un algorithme de classification d'images d'iris utilisant le réseau résiduel amélioré qui repose sur une approche d'apprentissage profond. Afin d'extraire d'autres

caractéristiques de l'image d'iris, ils ont divisé le canal de convolution ainsi que les caractéristiques extraites sous diverses dimensions de noyau de convolution.

### **11. Conclusion :**

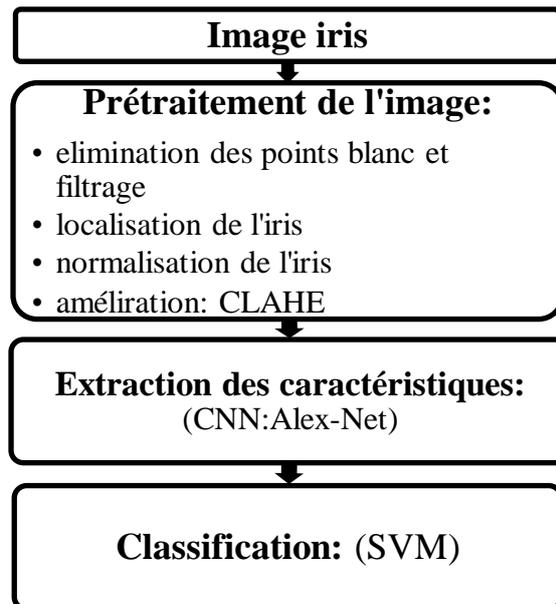
Tout au long de cette partie de la thèse, nous avons introduit les concepts de base de l'apprentissage supervisé ainsi que ceux de l'apprentissage profond. Dans ce contexte, nous avons introduit les réseaux de neurones et plus particulièrement les réseaux neuronaux convolutifs profonds (CNN) qui est la principale technique utilisée dans la thèse. Ces modèles permettent d'apprendre, grâce à l'apprentissage supervisé, les caractéristiques visuelles discriminantes à extraire sur les images afin de résoudre des problématiques de classification. Les CNN sont au centre de l'attention, en particulier en vision par ordinateur et reconnaissance de forme, vu les résultats impressionnants qu'ils produisent sur ces différentes tâches. Dans ce chapitre nous avons passé en revue les fondements essentiels des réseaux de neurones convolutifs. Nous avons défini l'architecture détaillée de ce type de réseau ainsi que les différentes couches qui le constitue. Cette étude détaillée du modèle CNN nous permet de mieux comprendre son processus et faire une conception pour une tâche de reconnaissance de l'iris qui fera l'objet du chapitre suivant.

## Chapitre 4

### Le système proposé

#### 1. Introduction :

Dans ce chapitre sont présentés les méthodologies ainsi que les résultats obtenus en implémentant le système de reconnaissance d'iris développé. La méthode proposée implique la mise en œuvre des étapes illustrées dans la **figure 4.1**, la première étape après l'acquisition de l'image est le prétraitement qui consiste à éliminer les points blancs de l'iris ainsi que d'autres types de bruit tels que les cils et les bruits de sel et de poivre dans l'image de l'œil, etc. Par la suite, nous passons à l'étape de localisation de l'iris, notre but dans cette thèse n'est pas de détecter la totalité de l'iris mais le but est de détecter une partie de l'iris qui est la collerette zigzag (plus précisément le demi-cercle inférieur), cette dernière est le sujet de notre contribution, La détection de la collerette en zigzag se fait à l'aide du détecteur de contour de Canny suivi de la transformée circulaire de Hough, après cela, nous effectuons une normalisation pour avoir un motif fixe en coordonnées polaires à l'aide du modèle de Daugman rubber sheet, puis nous passons à l'étape de l'extraction de caractéristiques à l'aide du CNN. Notre choix s'est porté sur le modèle pré-entraîné Alex-Net car il est simple et efficace, à la fin, une étape de classification par SVM est nécessaire pour calculer le taux de reconnaissance en utilisant la formule de précision [225].



**Figure 4.1** - Architecture générale de notre système de reconnaissance d'iris développé.

## 2. Acquisition de l'image

L'acquisition d'une image de l'iris est considérée comme l'un des aspects les plus difficiles de la biométrie. Tout d'abord, l'iris est sombre, il doit donc être éclairé, mais en même temps, l'iris est sensible à la lumière et un éclairage puissant peut gêner l'utilisateur. Deuxièmement, l'iris est un petit objet (environ 1 cm de diamètre), il faut donc utiliser une très forte focale ou rapprocher l'iris de l'objectif, mais ce n'est pas sans risque, car dans ce dernier cas on va rapprocher l'iris de la source d'éclairage, cela peut causer des dommages aux personnes. Enfin, l'iris est une surface qui réfléchit la lumière dans toutes les directions et se situe dans un miroir très réfléchissant derrière la cornée [226]. Ces deux dernières caractéristiques signifient que si aucune technique spécifique n'est utilisée, l'iris capturé sera recouvert par les réflexions de toutes les sources lumineuses présentes dans l'environnement d'acquisition.

La méthode proposée a été testée sur la base de données CASIA-Iris-Lamp, nombre de classe est 411. La base a été collectée à l'aide d'un capteur d'iris manuel produit par OKI. Une lampe a été allumée/éteinte à proximité du sujet pour introduire davantage de variations intra-classes lors de la collecte de CASIA-Iris-Lamp. Des conditions d'éclairage différentes entraînent l'expansion et la contraction de la pupille, ce qui provoque une déformation élastique de la texture de l'iris. C'est l'un des problèmes les plus courants et les plus difficiles à résoudre dans la reconnaissance de l'iris. Par conséquent, CASIA-Iris-Lamp est un bon outil pour étudier les problèmes de normalisation non linéaire de l'iris et de représentation robuste des caractéristiques de l'iris. Les images de la base de données CASIA sont des images JPG d'une résolution de 640\*480. Toutes

les images d'iris sont des images à niveau de gris de 8 bits et le format de fichier est JPEG (584, 2020)[227].

### 3. Prétraitement des images

Les étapes de notre système de reconnaissance sont présentées dans la **figure 4.1**. Dans la partie de prétraitement de l'image, il y a trois processus qui sont l'élimination des points blancs, la localisation de l'iris, et la normalisation de l'iris.

#### 3.1. Élimination des points blancs et filtrage

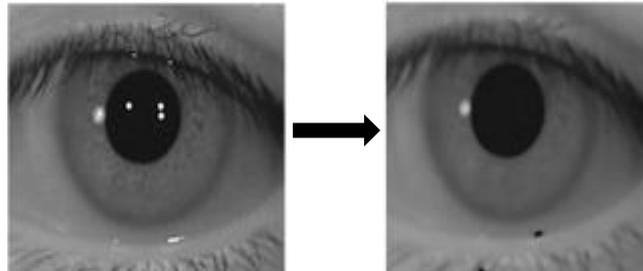
Dans la plupart des bases de données d'iris standard [228-231], des points blancs peuvent exister dans l'image de l'œil d'entrée et peuvent rester et perturber le processus de localisation de l'iris s'ils ne sont pas correctement supprimés, voir [232-236]. Cependant, dans [237], l'auteur a proposé un système efficace pour supprimer les points blancs et autres bruits étrangers tels que les fils de cheveux ou le bruit sel et poivre dans l'image de l'œil, le résultat est montré dans la **figure 4.2**. Dans ce qui suit, nous décrivons un algorithme efficace, inspiré par [237], et nous l'avons adapté sur nos images pour supprimer les points blancs et tout autre bruit- comme c'est décrit précédemment-, l'algorithme se déroule comme suit :

##### Algorithme élimination des points blancs

- Convertir l'image **Iim** (x, y) en image au niveau de gris **Igray** (x, y).
- Estimer les limites saturées -saturé de 1%- inférieure  $\alpha L$  et supérieure  $\alpha U$  de niveau de gris de l'image **Igray** (x, y).
- Convertir l'image **Igray** (x, y) en une image binaire **b**(x, y) :
 
$$b(x, y) = \begin{cases} 1 & , \text{ si } I_{\text{gray}}(x, y) < T_b \\ 0 & , \text{ sinon} \end{cases}$$

$$T_b = (\alpha U - \alpha L - 10)$$
- Dans **b**(x, y), une région noire (0s) qui est entourée d'une région blanche (1s) est appelée un trou et la région blanche est appelée un objet.
- Appliquer la méthode de 4-connectivité pour détecter et compléter l'intensité du niveau de gris des trous dans **b**(x, y) ; l'image résultante est **bhf** (x, y).
- Soustraire **b**(x, y) de **bhf** (x, y) ; l'image résultante **binf** (x, y) contient des informations concernant les trous et toute autre région blanche, éventuellement la sclérotique et la peau
- supprimer tout objet dans **binf** (x, y) dont la surface est supérieure à 80 pixels, l'image résultante est **M**(x, y).
- Utiliser les coordonnées (x, y) des pixels blancs (1s) dans **M**(x, y) pour remplacer les valeurs de niveau de gris, avec  $\alpha L$ , dans **Igray**(x, y), l'image résultante est **Ir** (x, y).
- Appliquer un filtre médian 7\*7 sur **Ir** (x, y) pour réduire les artefacts de frontière; l'image résultante est appelée image d'œil prétraitée **Ipre**(x, y).

La base de données CASIA lamp contient beaucoup de réflexions spéculaires dans la région de l'iris et de la pupille. Il est évident, d'après les résultats, que l'algorithme appliqué supprime effectivement ces points blancs comme montré dans la **figure 4.2**.



**Figure 4.2-**Élimination des points blancs et filtrage

### 3.2. Localisation de l'iris (collerette zigzag)

- **Détection des pupilles**

Pendant le processus d'acquisition, les changements de conditions d'éclairage peuvent influencer la qualité de la région de l'iris obtenue, puis affecter la localisation de l'iris et, par la suite, le résultat de la reconnaissance. Pour améliorer la précision et la fiabilité d'un système de reconnaissance de l'iris, il faut disposer d'une localisation précise de la région de l'iris, car les performances des étapes suivantes du système dépendent directement de la qualité de la région de l'iris détectée. Un système ordinaire de localisation de l'iris vise à détecter les deux limites de la région de l'iris : la limite intérieure (pupille-iris) et la limite extérieure (iris-sclérotique). Cependant, la tâche devient plus difficile lorsque les paupières et les cils recouvrent une partie de l'iris. Pour cette raison, une nouvelle idée a été proposée dans cette thèse qui se base sur la détection de la collerette - et plus précisément la détection du demi-cercle inférieur de la collerette -, après avoir détecté la pupille, et cela au lieu de détecter l'iris entier.

Le processus de segmentation de l'iris commence par la détection des limites de la pupille. Pour cela, la détection des bords de Canny est appliquée [238] pour générer une carte des bords, puis la transformation de Hough circulaire (CHT) est appliquée par la suite [239].

- **Le détecteur de bords de Canny** : Le détecteur de bords de Canny est l'un des outils de traitement d'images les plus couramment utilisés. Il détecte les arêtes de manière très robuste. L'intention de Canny était d'améliorer les nombreux détecteurs de bords déjà existants au moment où il a commencé son travail. L'algorithme s'exécute en 5 étapes distinctes :

Algorithme Canny

1)- Lissage : flou de l'image pour éliminer le bruit.

2)- Trouver les gradients : Il faut marquer les bords où les gradients de l'image sont importants. Calculez les dérivées ( $D_x(x, y)$  et  $D_y(x, y)$ ) de l'image dans les directions x et y, c'est-à-dire utiliser la différenciation centrale en utilisant les noyaux  $3 \times 3$  suivants :

Calculez ensuite la magnitude du gradient et l'angle de magnitude :

$$D = \sqrt{D_x^2(x, y) + D_y^2(x, y)}$$

$$\theta = \arctan\left(\frac{D_x(x, y)}{D_y(x, y)}\right)$$

3)-Suppression des maxima locaux : Seuls les maxima locaux doivent être marqués comme des arêtes. L'étape de "suppression non maximale" ne conserve que les pixels d'un bord ayant la magnitude de gradient la plus élevée. Ces amplitudes maximales doivent se produire juste à la limite du bord, et l'amplitude du gradient doit diminuer avec la distance du bord.

Ainsi, trois pixels dans un espace  $3 \times 3$  autour du pixel  $(x, y)$  sont examinés :

-Si  $(x, y) = 0^\circ$ , alors se sont les pixels  $(x + 1, y)$ ,  $(x, y)$  et  $(x - 1, y)$ .

-Si  $(x, y) = 90^\circ$ , alors se sont les pixels  $(x, y + 1)$ ,  $(x, y)$  et  $(x, y - 1)$ .

-Si  $(x, y) = 45^\circ$ , alors se sont les pixels  $(x + 1, y + 1)$ ,  $(x, y)$  et  $(x - 1, y - 1)$ .

-Si  $(x, y) = 135^\circ$ , alors se sont les pixels  $(x + 1, y - 1)$ ,  $(x, y)$  et  $(x - 1, y + 1)$ .

-Le pixel  $(x, y)$  présente l'amplitude de gradient la plus élevée des trois pixels examinés ; il est conservé comme un bord. Si l'un des deux autres pixels a une magnitude de gradient plus élevée, alors le pixel  $(x, y)$  n'est pas au "centre" du bord et ne doit pas être classé comme un pixel de bord.

4)-Double seuillage : Les arêtes potentielles sont déterminées par seuillage.

5)-Suivi des arêtes par hystérésis : Les bords finaux sont déterminés en supprimant tous les bords qui ne sont pas connectés à un bord très certain (fort).

La détection des contours étant une étape fondamentale de la vision par ordinateur, il est nécessaire de mettre en évidence les véritables contours pour obtenir les meilleurs résultats du processus de comparaison. C'est pourquoi il est important de choisir les détecteurs de bords les mieux adaptés à l'application. L'opérateur de Canny est optimal même pour les images bruitées,

car la méthode comble le fossé entre les bords forts et faibles de l'image en connectant les bords faibles dans la sortie uniquement s'ils sont connectés à des bords forts. Les bords sont donc plus susceptibles d'être les vrais. Par conséquent, comparé à d'autres méthodes de détection des bords, cet opérateur de Canny est moins trompé par le bruit parasite [240].

**-La transformée de Hough circulaire :** La transformée de hough est considérée comme un outil très puissant dans la liaison des bords pour l'extraction des lignes. Ses principaux avantages sont son insensibilité au bruit et sa capacité à extraire des lignes même dans des zones où il n'y a pas de pixels (trous de pixels). La transformée de hough standard proposée par Duda et Hart (Duda et Hart, 1972) est largement appliquée pour l'extraction de lignes dans les scènes naturelles, tandis que certaines de ses modifications ont été ajustées pour l'extraction de linéaments géologiques [241].

La transformée de Hough circulaire standard (CHT) détecte les formes circulaires d'un rayon donné dans l'image. La carte des contours de l'image de l'œil est générée en calculant les dérivées premières des valeurs d'intensité. Chaque point de la carte des bords donne un cercle de rayon  $r$  et de centre  $(x_c, y_c)$  à un tableau d'accumulateurs de sortie. Ensuite, une procédure de vote est utilisée pour trouver le plus grand pic dans le tableau d'accumulateurs résultant dans l'espace des paramètres, qui correspond au cercle le mieux défini par les points d'arête [20]. En considérant l'ensemble obtenu de points d'arête comme  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$ . Le CHT pour l'ajustement d'un cercle de rayon  $r$  et de coordonnées centrales  $(x_c, y_c)$  est défini comme suit :

$$H(x_c, y_c, r) = \sum_{i=1}^n h(x_i, y_i, x_c, y_c, r) \quad (1)$$

Où

$$h(x_i, y_i, x_c, y_c, r) = \begin{cases} 1, & \text{if } g(x_i, y_i, x_c, y_c, r) = 0 \\ 0 & \text{sinon} \end{cases} \quad (2)$$

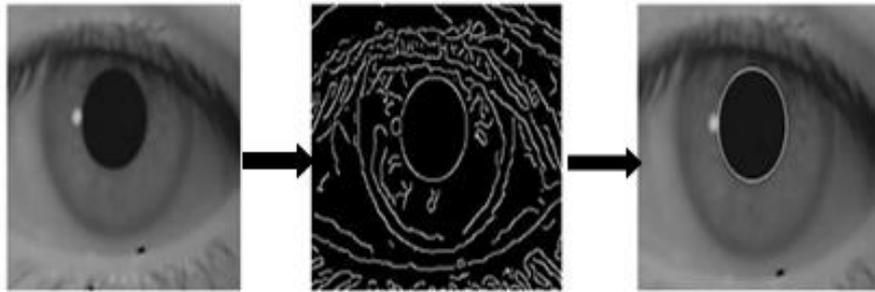
et la fonction paramétrique  $g$  est définie comme suit:

$$g(x_i, y_i, x_c, y_c, r) = (x_i - x_c)^2 + (y_i - y_c)^2 - r^2 \quad (3)$$

La valeur de la fonction  $g$  de chaque point est égale à 1, s'il est situé sur le cercle de paramètres  $(x_c, y_c, r)$  par (2). Enfin, la procédure de vote est mise en œuvre dans l'espace de Hough pour détecter le cercle correct [242].

La transformée de Hough circulaire standard est utilisée -après le détecteur de bords de Canny - pour détecter les formes circulaires d'un rayon donné dans l'image. Dans le travail de [243] l'auteur a conclu qu'il est impossible de calculer un rayon automatique - de la pupille ou de l'iris

- parce que cette mesure diffère suivant les races. Dans notre expérience, le rayon limite des pupilles ont été trouvées selon des techniques heuristiques après examen de toutes les images de la base de données CASIA V4 lamp, la **figure 4.3** montre le cercle correct détecté après la mise en œuvre de la procédure de vote dans l'espace de Hough.



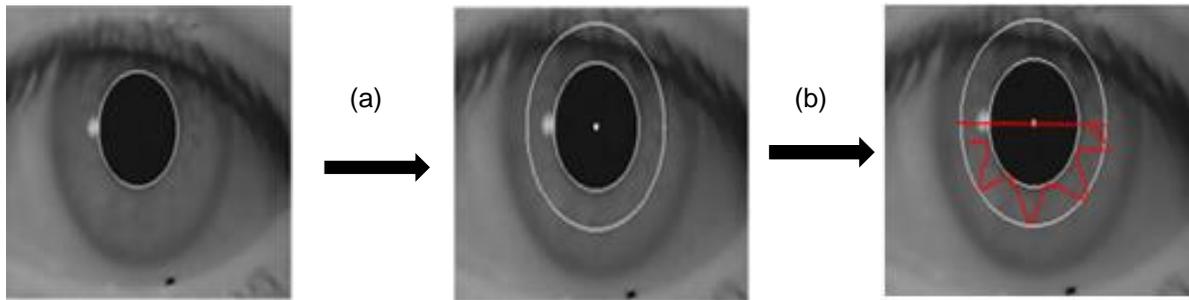
**Figure 4.3-**Détection de la pupille

- **Détection de la région en zigzag de Collarete**

Le traitement suivant consiste à isoler la zone de la collerette en zigzag (voir **figure 4.4**). Cette zone est généralement concentrique et proche de la pupille. Elle sera donc très facile à détecter à l'aide de la formule de traçage d'un cercle en utilisant les coordonnées du centre de la pupille et un rayon qui est égal à 24 selon une étude menée par Rai, ce dernier a conclu que la collerette zigzag se trouve à moins de 24 pixels de la pupille [49].

Dans la recherche de la détection de la zone de la collerette en zigzag, de nombreux problèmes sont rencontrés. Dans la plupart des cas, il y a du bruit autour de la pupille, ce qui donne une forme géométrique de la pupille qui n'est pas complètement circulaire. Il y a toujours des cils et des paupières qui couvrent la zone. Par conséquent, les chercheurs dans ce domaine ont essayé de faire un certain traitement du bruit, comme dans l'article [244].

Dans notre thèse, ce grand problème est surmonté en utilisant seulement le demi-cercle inférieur du zigzag de la collerette (voir **figure 4.4**) qui n'est pas affecté ni par les paupières ni par les cils. Le principe de notre idée est de sélectionner la région la plus importante comportant le motif le plus complexe avec le moindre bruit possible, cela pour éviter un autre traitement du bruit qui réduit la qualité de l'image dans la plupart des cas[225].



**Figure 4.4**-Isolement de la zone de collarete en zigzag (a), et le demi-cercle inférieur de la zone de collarete en zigzag (b).

### 3.3. Normalisation de l'iris (collarete zigzag)

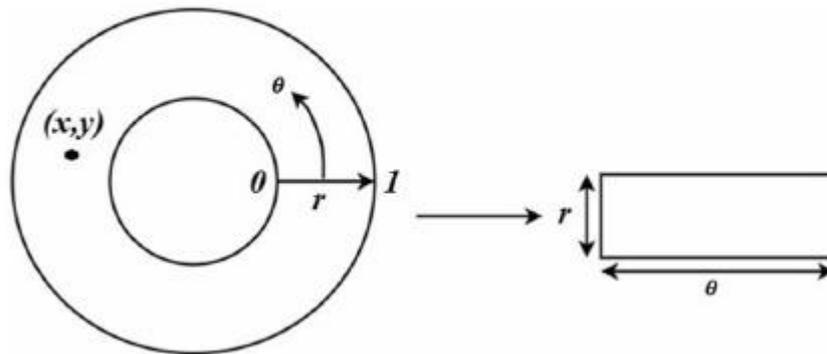
Une fois la collerette en zigzag est détectée, une normalisation est mise en œuvre pour produire un vecteur de caractéristiques de taille standard qui permet de comparer deux images d'iris différentes. Le principal avantage du processus de normalisation est de supprimer les incohérences dimensionnelles qui peuvent se produire en raison de l'étirement de la région de l'iris causé par la pupille ce phénomène peut se produire en raison de l'étirement de la région de l'iris causé par la dilatation de la pupille avec différents niveaux d'illumination. D'autres causes d'incohérences dimensionnelles sont les changements de la distance d'acquisition de l'image, la rotation de la caméra ou de l'œil et la distorsion élastique de la texture de l'iris qui peuvent affecter le résultat de la correspondance de l'iris et ainsi de suite. Pour résoudre tous ces problèmes, le processus de normalisation de l'iris est appliqué en utilisant le modèle de Daugman (Daugman's rubber sheet model) pour transformer l'image de l'iris de coordonnées cartésiennes en coordonnées polaires, comme le montre la **figure 4.5** ci-dessous. Le modèle de Daugman prend chaque point  $(x, y)$  dans la région de l'iris en une paire de coordonnées polaires non concentriques normalisées  $(r, \theta)$  où  $r$  est sur l'intervalle  $[0, 1]$  et  $\theta$  est l'angle sur l'intervalle  $[0, 2\pi]$ . Ce mappage de la région de l'iris peut être défini mathématiquement comme suit :

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta)$$

$$x(r, \theta) = (1-r)x_p(\theta) + rx_l(\theta)$$

$$y(r, \theta) = (1-r)y_p(\theta) + ry_l(\theta)$$

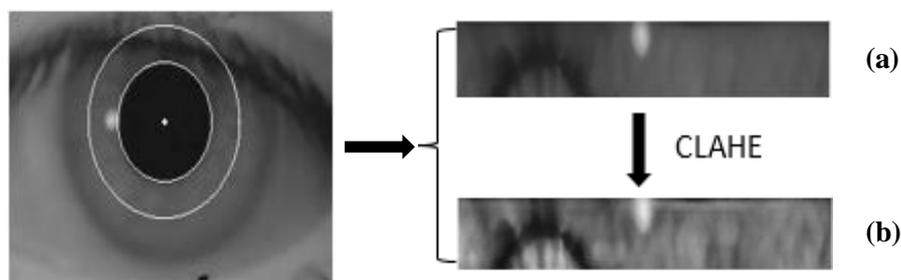
Ici,  $I(x, y)$  est la valeur d'intensité en  $(x, y)$  dans l'image de la région de l'iris, les paramètres  $x_p$ ,  $x_l$ ,  $y_p$  et  $y_l$  sont les coordonnées des limites de la pupille et de l'iris le long la direction  $\theta$ . Le résultat de la normalisation est illustré à la **figure 4.6(a)** [245].



**Figure 4.5** - le modèle de Daugman pour transférer la région de l'iris des coordonnées cartésiennes aux coordonnées polaires[245].

**3.4. Post-traitement par CLAHE :** CLAHE (L'égalisation d'histogramme adaptative limitée par le contraste) est une méthode d'ajustement du contraste pour obtenir une image avec des niveaux d'intensité uniformément distribués, c'est une méthode modifiée de l'égalisation adaptative par histogramme (AHE). Dans cette méthode, la fonction d'amélioration est appliquée à tous les pixels voisins et la fonction de transformation est dérivée. Cette méthode diffère de l'AHE en raison de sa limitation du contraste. Dans l'article [246], Zhiyuan Xu, Xiaoming Liu, Xiaonan Chen ont utilisé la méthode CLAHE pour le processus d'amélioration, dans cette méthode, la valeur maximale est utilisée pour couper l'histogramme et le redistribuer dans l'image de niveau de gris [247].

Dans cette thèse, CLAHE est appliquée comme post-traitement aux images normalisées (voir **figure 4.6(b)**).



**Figure 4.6** - (a) normalisation de l'iris (b) amélioration de l'iris avec la méthode CLAHE

Les images normalisées sont améliorées afin d'éviter la perte de caractéristiques, d'extraire les points clés avec précision, et donc d'augmenter la précision de la reconnaissance, l'algorithme CLAHE est présenté dans les étapes suivantes.

Algorithme CLAHE

**Étape 1** : L'image est chargée.

**Étape 2** : Les zones contextuelles sont définies. La taille de la zone contextuelle peut être régulée en modifiant la limite.

**Étape 3** : Les pixels égaux, c'est-à-dire les pixels ayant la même valeur dans chaque zone contextuelle, sont comptés.

**Étape 4** : Une limite de clip est fixée. Ce seuil limite déterminera quelles valeurs passent sur l'histogramme après l'égalisation, sa valeur doit être comprise entre 0 et 1. La limite de clip doit être différente selon la densité de l'image.

**Étape 5** : La valeur de chaque pixel est comparée à la valeur limite du clip. La partie de l'histogramme qui dépassait la limite de clip n'a pas été éliminée. Au lieu de cela, elle sera redistribuée.

La redistribution se fait comme suit. Si un pixel dont la valeur est supérieure à la limite du clip est détecté, alors l'opération suivante se produit

$$\text{incrément} = \text{limite de clip} / \text{valeur du pixel examiné}$$

Ainsi, l'incrément est incrémenté de la valeur de la limite du clip divisée par l'intensité du pixel. Sinon, si la valeur du pixel est inférieure à la limite du clip, l'incrément est de 1.

**Étape 7** : Un rang partiel est calculé avec la fonction

$$\text{rang partiel} = \text{rang partiel} + \text{incrément} ;$$

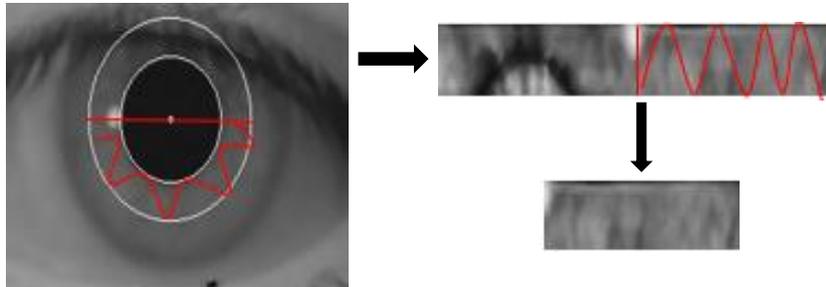
**Étape 8** : La redistribution a lieu. Les nouvelles valeurs de pixel distribuées peuvent être trouvées à partir de la valeur de redistribution et seront incrémentées par le rang partiel.

$$\text{Redistribution} = (\text{valeur totale de clip} / (\text{limite de la fenêtre} * \text{limite de la fenêtre})) * \text{données}(x,y)$$

$$\text{sortie}(x,y) = \text{rang partiel} + \text{valeur de redistribution}[248]$$

Notre contribution dans cette thèse est de prendre uniquement le demi-cercle inférieur du zigzag de la collerette (comme décrit précédemment), cela se fait en divisant l'iris normalisé en deux parties, voir **figure 4.7**. D'après la base de données d'images utilisées, on a remarqué que la

partie inférieure de l'iris est la moins affectée par les cils et les paupières par rapport à la partie supérieure, et dans la plupart des cas, cette partie représente une région sans bruit.



**Figure 4.7**-Normalisation de l'iris : demi-cercle inférieur de la région de la collerette

Dans la suite de cette thèse, nous allons effectuer deux expériences, dans la première expérience, on voudra tester les performances de nos traitements sur l'image, dans la deuxième expérience, nous allons effectuer une comparaison des taux de reconnaissance de l'iris en utilisant le cercle complet de la collerette et le demi-cercle inférieur de la collerette afin de montrer la grande amélioration du taux de reconnaissance.

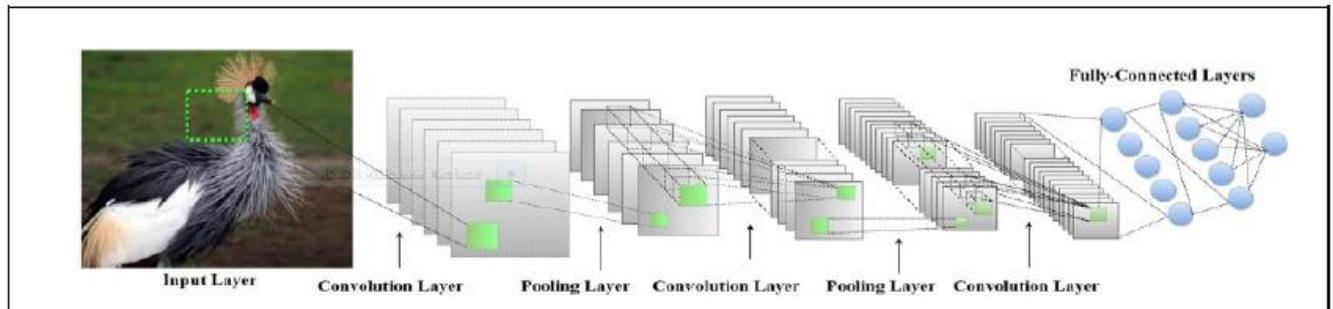
#### 4. Extraction de caractéristiques

##### - Réseau neuronal convolutif (CNN)

Les réseaux neuronaux convolutifs (CNN) appartiennent à une catégorie spécifique de méthodes de réseaux neuronaux. Les CNN ont non seulement été capables d'apprendre automatiquement des représentations de caractéristiques d'image, mais ils ont également surpassé de nombreuses techniques conventionnelles d'extraction des caractéristiques manuellement [244].

Les modèles de réseaux neuronaux ont une représentation hiérarchique des données et dépendent du calcul des couches qui ont une mise en oeuvre séquentielle, la sortie de la couche précédente sera l'entrée de la couche suivante. Chaque couche donne un niveau de représentation, et il y a un ensemble de poids qui paramètrent les couches. De même, les unités d'entrée sont reliées aux unités de sortie par les poids, en plus d'un groupe de biais [249]. Les poids dans les réseaux neuronaux convolutifs (CNN) sont partagés localement, ce qui signifie que chaque emplacement de l'entrée a des poids similaires. Le filtre formé par les poids est lié à la sortie similaire [249].

Un réseau neuronal convolutif (CNN) comprend des couches alternées de couches convolutives connectées localement où chaque couche possède le même nombre de filtres, des couches de sous-échantillonnage et des couches entièrement connectées qui fonctionnent comme un classificateur [250]. **La figure 4.8** montre l'architecture globale d'un CNN.



**Figure 4.8-** Une illustration de l'architecture des réseaux neuronaux convolutifs. Les carrés gris mentionnent les cartes de caractéristiques et les carrés verts mentionnent le filtre de convolution. Les lignes croisées entre les deux dernières couches mentionnent les neurones entièrement connectés [245].

L'architecture des réseaux neuronaux convolutifs repose sur trois concepts qui la rendent efficace : les champs réceptifs locaux, le partage des poids et les opérations de sous-échantillonnage [245]. Le champ réceptif local signifie que chaque neurone accepte l'entrée d'une petite partie de la couche précédente. De plus, il a la même taille que le filtre de convolution. Les champs réceptifs locaux sont utilisés dans les systèmes convolutifs et les systèmes de couches de sous-échantillonnage. Le partage des poids est appliqué à la couche convolutionnelle pour contrôler la capacité et diminuer la complexité du modèle. Enfin, le sous-échantillonnage non linéaire est utilisé dans les couches de sous-échantillonnage pour diminuer la taille spatiale de l'image ainsi que le nombre de paramètres libres du modèle. Ces concepts aident le CNN à être fort et efficace dans les tâches de reconnaissance [245]. Voir le chapitre 3 la section 9.2 pour plus de détail sur les CNN[233].

Dans la suite de cette thèse nous avons défini les réseaux de convolution utilisés, le réseau AlexNet a été utilisé pour l'extraction des caractéristiques dans le processus de reconnaissance de l'iris, le réseau GoogleNet a été utilisé d'un côté pour tester les performances des traitements appliqués dans notre chaîne de reconnaissance et de l'autre côté pour faire une comparaison avec AlexNet.

- **Google-Net** : GoogLeNet est un réseau neuronal convolutif profond à 22 couches qui est une variante de l'Inception Network, un réseau neuronal convolutif profond développé par les chercheurs de Google. L'architecture GoogLeNet présentée lors de l'ImageNet Large-Scale Visual Recognition Challenge 2014(ILSVRC14) a résolu des tâches de vision par ordinateur telles que la classification d'images et la détection d'objets. Aujourd'hui, GoogLeNet est utilisé pour d'autres tâches de vision par ordinateur, telles que la détection et la reconnaissance des visages, l'entraînement adversarial, etc. L'architecture GoogLeNet est constituée de 22 couches, et une partie de ces couches est constituée d'un total de 9 modules d'inception [251], la figure **figure 4.9**, montre l'architecture de GoogLeNet et le **tableau 4.1** représente l'architecture Inception.

**Tableau 4.1** - GoogLeNet : l'architecture Inception [252].

Type	patch size/ stride	Output Size	Depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj	params	Ops
Convolution	7×7/2	112×112×64	1							2.7K	34M
max pool	3×3/2	56×56×64	0								
Convolution	3×3/1	56×56×192	2		64	192				112K	360M
max pool	3×3/2	28×28×192	0								
inception (3a)		28×28×256	2	64	96	128	16	32	32	159K	128M
inception (3b)		28×28×480	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	14×14×480	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		14×14×512	2	160	112	224	24	64	64	437K	88M
inception (4c)		14×14×512	2	128	128	256	24	64	64	463K	100M
inception (4d)		14×14×528	2	112	144	288	32	64	64	580K	119M
inception (4e)		14×14×832	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	7×7×832	0								
inception (5a)		7×7×832	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	1×1×1024	0								
dropout (40%)		1×1×1024	0								
Linear		1×1×1000	1							1000K	1M
softmax		1×1×1000	0								

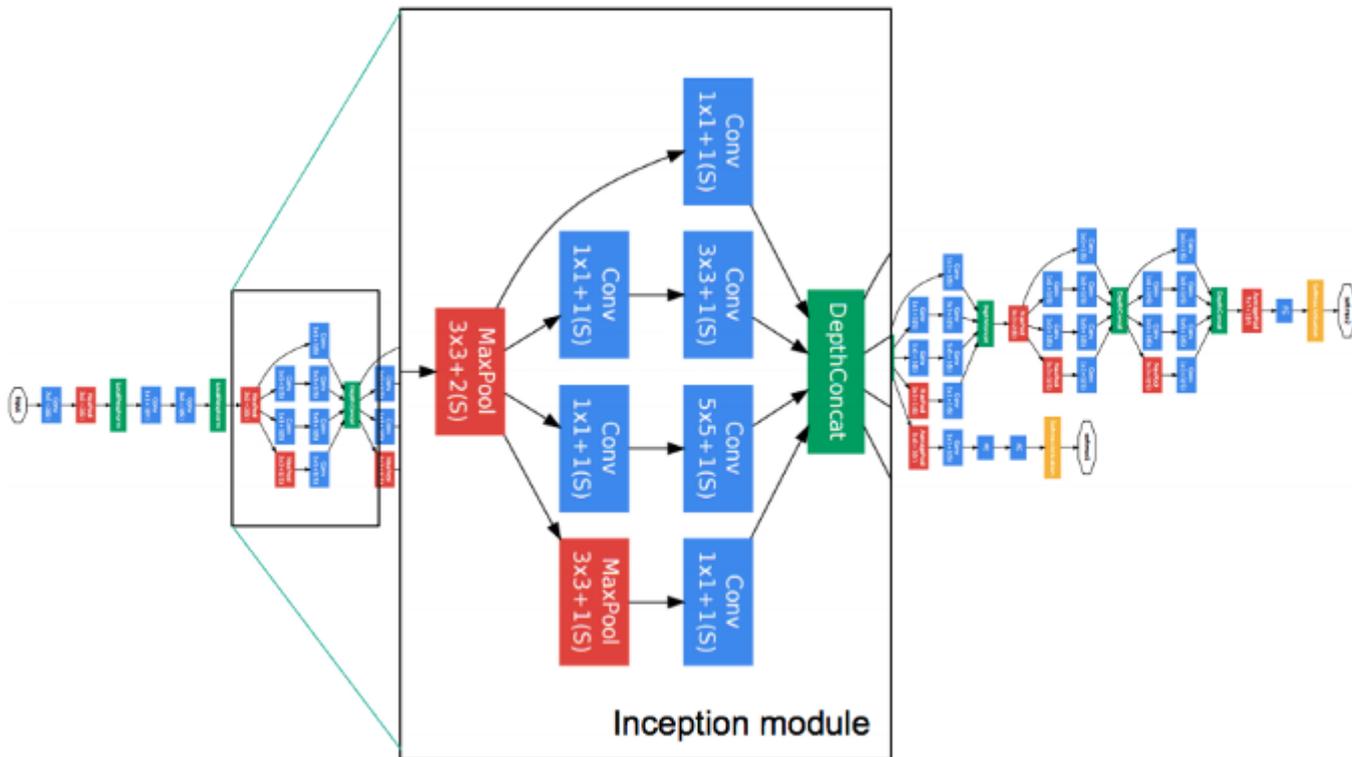


Figure 4.9- GoogLeNet module d'inception [252]

- **Alex-Net** : Le modèle de réseau neuronal convolutionnel pré-entraîné (Alex-Net) est utilisé pour le processus d'extraction des caractéristiques. Ce modèle a été conçu par le groupe Super Vision [233]. L'Alex-Net est une version à échelle réduite du Le-Net conventionnel [244]. Alex-Net est entraîné sur le défi de reconnaissance visuelle à grande échelle Image-Net (ILSVRC). Alex-Net est entraîné à classer les 1,2 million d'images de la base de données Image-Net en 1000 classes différentes. La **figure 4.10** montre l'architecture globale du modèle Alex-Net [47].

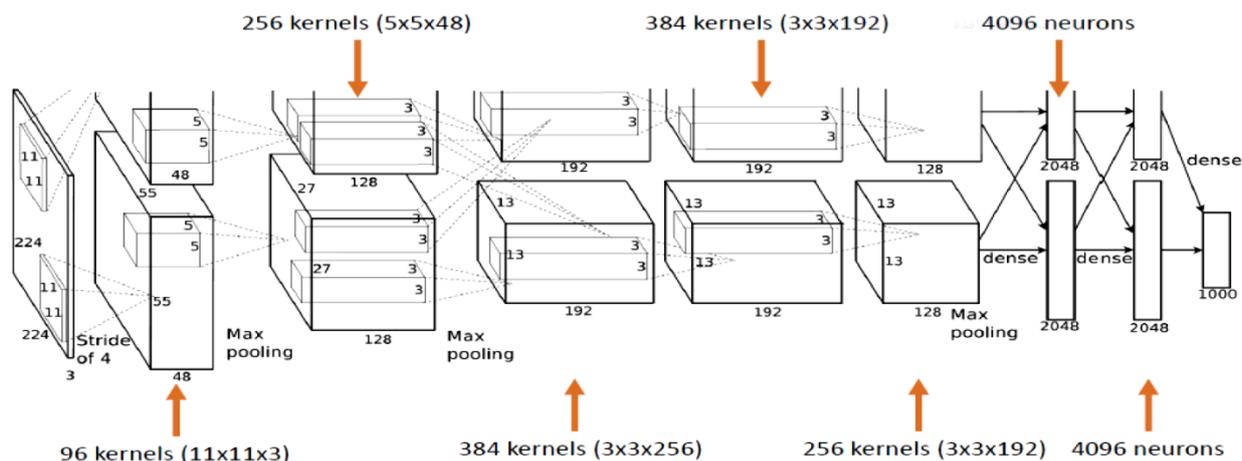


Figure 4.10-Architecture détaillée d'Alex-Net [233].

Le modèle Alex-Net contient un ensemble de couches ; la couche d'entrée est la première couche qui définit les dimensions d'entrée. Le modèle Alex-Net a besoin d'une image d'entrée de 227 x 227 x 3. Les couches intermédiaires constituent l'essentiel du modèle AlexNet. Ces couches sont constituées d'une série de cinq couches convolutionnelles, suivies d'unités linéaires rectifiées (ReLU) et des couches de max-pooling. À côté de ces couches, trois couches entièrement connectées. La couche de classification est la dernière couche [233]. Les premières couches convolutionnelles effectuent des convolutions 11x11 avec un stride de 4 et sans padding, convolutions 5x5 avec stride 1 et pad 2, les autres couches convolutionnelles couches convolutionnelles effectuent des convolutions 3x3 avec stride 1 et pad 1, et un pooling 2x2 (sans padding). L'explication détaillée des couches d'Alex-Net est présentée dans le **tableau 4.2**

Alex-Net est le réseau pré-entraîné utilisé dans cette thèse pour le processus d'extraction de caractéristiques, le choix du modèle Alex-Net par rapport à un autre modèle pré-entraîné n'est pas arbitraire, Alex-Net est un modèle simple et il offre la possibilité de tester les performances sans compromettre la mémoire et le temps.

**Tableau 4.2-** La couche d'Alex-Net [233]

Type of Layer	No. of Filter	Feature Map Size (height x width x channel)	Kernel Size	No. Of Stride	No. of Padding
Image input layer		227x227x3			
(1st convolutional layer) Relu-1 Cross-channel normalization Max pooling1	96 1	55x55x96 27x27x96	11x11 3x3	4x4 2x2	0x00x0
(2nd convolutional layer) Relu-2 Cross-channel normalization Max pooling2	256 1	27x27x256 13x13x256	5x5 3x3	1x1 2x2	2x20x0
(3rd convolutional layer) Relu-3	384	13x13x384	3x3	1x1	1x1
(4th convolutional layer) Relu-4	384	13x13x384	3x3	1x1	1x1
(5th convolutional layer) Relu-5 Max pooling5	256 1	13x13x256 6x6x256	3x3 3x3	1x1 2x2	1x10x0
Fully connected layer-6(fc6) Relu-6		4096x1 4096x1			
Fully connected layer-7 (fc7) Relu-7		4096 x1 4096 x1			
Fully connected layer-8 (fc8) Softmax layer Output layer		1000x1 1000 class			

Les réseaux neuronaux profonds apprennent des caractéristiques de haut niveau dans les couches cachées. C'est l'une des plus grandes forces des CNN et cela permet de réduire les besoins en ingénierie des caractéristiques. En fait, l'image subit plusieurs transformations. Tout d'abord, l'image passe par de nombreuses couches convolutionnelles où le réseau apprend de nouvelles caractéristiques de plus en plus complexes. Ensuite, l'information de l'image transformée passe par les couches entièrement connectées et est transformée en une classification. Les caractéristiques de haut niveau peuvent être récupérées à partir des dernières couches de convolution FC7.

Pour éviter un temps d'apprentissage très long en utilisant le CNN, et puisque toutes les couches sont responsables de l'apprentissage de certaines caractéristiques des images, les caractéristiques peuvent être extraites du réseau à tout moment pendant le processus d'apprentissage. Nous utilisons ces caractéristiques extraites comme données d'entrée pour un modèle de classification avec des Machines à Vecteur de Support (SVM).

Dans notre travail, le vecteur caractéristique est récupéré à partir de la couche entièrement connectée (FC7) car nous obtenons des caractéristiques de haut niveau à partir de ces couches convolutives [225].

### **5. La classification**

Le classifieur est appliqué après l'extraction des caractéristiques pour trouver l'étiquette correspondante pour chaque image de test. De nombreux classificateurs peuvent être utilisés pour la classification, avec différents types. Par exemple, nous citons le réseau neuronal, la régression Softmax et la machine à vecteur de support. Dans notre travail, un classificateur multiclasse Support Vector Machine est utilisé (voir la section 5.5 du chapitre 3 pour plus de détails). Le SVM est un algorithme d'apprentissage supervisé ; il construit un hyperplan optimal comme surface de décision pour maximiser la marge de séparation entre les deux classes de données. Les vecteurs de support font référence à un petit sous-ensemble d'observations d'apprentissage utilisées comme support pour l'emplacement optimal de la surface de décision

Après le chargement de la base de données, celle-ci est divisée en deux parties : 70% pour l'apprentissage et 30% pour le test. Ensuite, nous extrayons les étiquettes de classe des données d'entraînement et de test. Le principe de l'algorithme de classification SVM est de placer chaque point de données dans un espace à  $n$  dimensions, où  $n$  est le nombre de caractéristiques. La valeur de chaque caractère est la valeur d'une coordonnée particulière. Ensuite, on effectue la classification en cherchant l'hyperplan qui différencie très bien les deux classes. L'algorithme

SVM va donc classer les images et finalement calculer l'accuracy ou le taux d'apprentissage, c'est-à-dire l'efficacité de la classification. Cette valeur exprime la fraction des étiquettes que le réseau prédit correctement, comme défini par la formule suivante :

$$Accuracy = \frac{\text{correctly classified Iris Images}}{\text{Total Number}} \times 100 \quad (21)$$

## 6. Résultats expérimentaux et analyse

Nos expériences sont réalisées dans MATLAB R2018a sur un PC Intel core i5, une RAM de 6 Go, et un système d'exploitation Windows 64 bits. Les images utilisées dans nos expériences proviennent de la base de données d'images CASIA V4-Iris-Lamp. Cette version la plus récente de CASIA a été utilisée avec CNN pour étudier l'efficacité de l'utilisation de la région de la collerette en zigzag dans la reconnaissance de l'iris (deuxième expérimentation) et d'autre part pour tester les performances de nos méthodes et traitements utilisés sur l'image et justifier leur choix (première expérimentation).

### 6.1. Première expérimentation (détection de la région de l'iris)

Dans la première expérience, on a voulu tester les performances de nos traitements sur l'image, nous avons commencé par le post-traitement, pour cela on a utilisé 100 images de la région de l'iris (10 classes et 10 échantillons), pour comparer le taux de reconnaissance des images d'iris sans prétraitement et sans post-traitement, en utilisant le réseau AlexNet (comme c'est expliqué dans la section 4) et le réseau GoogleNet pour l'extraction des caractéristiques et le SVM pour la classification, les résultats sont montrés dans le **tableau 4.3**

Pour le réseau GoogleNet, nous l'avons traité comme nous l'avons fait avec le réseau AlexNet, le réseau GoogleNet a été utilisé pour l'extraction des caractéristiques en suite le résultat a été récupéré de la couche FC 142 et réinjecté dans le SVM, le réseau GoogleNet n'est pas utilisé dans notre processus de reconnaissance mais il a été utilisé pour tester les performances et justifier le choix des méthodes utilisé dans la chaîne de reconnaissance.

Le résultat dans le **tableau 4.3** montre clairement l'efficacité de notre post-traitement parce que le taux de reconnaissance a amélioré avec le réseau AlexNet et le réseau GoogleNet, cela justifie le choix de notre méthode de post-traitement.

**Tableau 4.3-** comparaison de taux de reconnaissance (accuracy), avec et sans post-traitement en utilisant les deux modèles AlexNet et GoogleNet.

Détection de la région de l'iris	AlexNet	GoogleNet
- Sans élimination des points blancs et sans filtrage - Sans post-traitement : CLAHE	42.00 %	40.00 %
- Sans élimination des points blancs et sans filtrage -Avec post-traitement : CLAHE	<b>76.67 %</b>	66.67 %

D'autre part, on a voulu savoir l'efficacité de prétraitement (élimination des points blancs et filtrage) proposé, pour cela on a utilisé 100 images de la région de l'iris (10 classes et 10 échantillons), pour comparer le taux de reconnaissance des images d'iris avec prétraitement et sans prétraitement, en utilisant le réseau AlexNet et GoogleNet pour l'extraction des caractéristiques et le SVM pour la classification, les résultats sont montrés dans le **tableau 4.4**

Le résultat dans le **tableau 4.4** montre l'amélioration très perceptible de taux de reconnaissance avec le réseau AlexNet et le réseau GoogleNet, cela justifie le choix de notre méthode de prétraitement.

**Tableau 4.4-** comparaison de taux de reconnaissance (accuracy), avec et sans prétraitement en utilisant les deux modèles AlexNet et GoogleNet

Détection de la région de l'iris	AlexNet	GoogleNet
- Sans élimination des points blancs et sans filtrage -Avec post-traitement	76.67 %	66.67 %
- avec élimination des points blancs et filtrage -Avec post-traitement	<b>83.33 %</b>	76.67 %

D'après le **tableau 4.3** et le **tableau 4.4** on remarque que le taux de reconnaissance en utilisant le modèle AlexNet est toujours supérieur aux taux de reconnaissance en utilisant le modèle GoogleNet cela justifie notre choix.

## 6.2. Deuxième expérimentation (détection de la région de la collerette zigzag)

Pour notre deuxième expérience, 400 images ont été utilisées à partir de la base de données CASIA (40 classes et 10 échantillons par classe), plusieurs expériences sont réalisées avec 100 images (10 classes, 10 échantillons), 200 images (20 classes, 10 échantillons), 300 images (10 classes, 10 échantillons), 400 images (40 classes, 10 échantillons) comme mentionné dans le **tableau 4.5**. Alex-Net est utilisé pour l'extraction des caractéristiques. Les données sont divisées en deux parties, 70% pour l'apprentissage et 30% pour le test. Toutes les images de la base de données sont redimensionnées à 227 par 227, ce qui est la taille d'entrée de l'Alex-Net. Toutes les images en niveaux de gris sont converties en RVB. A la fin, les caractéristiques apprises en utilisant le CNN (Alex-Net) et extraites de la couche FC7 seront injectées dans le classificateur SVM multiclasse pour la classification des images[225].

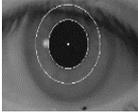
L'idée proposée dans la section précédente a été mise en œuvre et l'accuracy a été calculée en utilisant le cercle complet du collerette zigzag et comparée à l'accuracy calculée en utilisant le demi-cercle inférieur du collerette zigzag. Le résultat expérimental est présenté dans le **tableau 4.5**.

Le modèle CNN pré-entraîné utilisé a déjà été entraîné sur plus d'un million. Alex-Net avec le classificateur SVM a obtenu un taux très intéressant avec un temps d'apprentissage assez court, de l'ordre de quelques secondes à quelques minutes seulement. L'avantage du modèle CNN pré-entraîné est l'élimination de la tâche laborieuse d'ingénierie des caractéristiques, ce qui facilite l'apprentissage de la nouvelle tâche assignée. Le **tableau 4.5** illustre l'accuracy avec les différents cas étudiés : 100 images, 200 images, 300 images et 400 images comme expliqué ci-dessous ; pour chaque ensemble d'images de collerette zigzag, on a comparé la collerette entière par rapport à son demi-cercle inférieur. Le résultat montre que le meilleur accuracy est obtenue avec les échantillons en utilisant le demi-cercle inférieur de la collerette zigzag.

Sur la base des résultats, notre contribution qui utilise uniquement le demi-cercle inférieur de la collerette a obtenu le meilleur taux de reconnaissance par rapport aux anciennes méthodes qui utilisent le cercle complet de la collerette, [45-51,54-56].

Dans notre étude, nous avons effectué plusieurs apprentissages en augmentant le nombre de classes : (10,20,30,40) et en gardant le nombre d'échantillons toujours fixe, nous remarquons que le taux de reconnaissance diminue, d'autre part, si nous augmentons le nombre d'échantillons par classe (quelque chose d'impossible avec la base de données CASIA) l'apprentissage par CNN s'améliorera grandement et l'accuracy aussi.

**Tableau 4.5** - Accuracy : collarete cercle complet et collarete demi-cercle inférieur.

Échantillons (images)	Collarete : Le cercle complet 	Collarete : demi-cercle inférieur (notre proposition) 
100	93. 33%	100%
200	91. 67%	100%
300	88. 89%	100%
400	88. 33%	<b>94. 17%</b>

En plus de l'accuracy, l'analyse de la performance des deux expériences (Collarete : cercle complet / Collarete : demi-cercle inférieur) a été basée sur les mesures d'évaluation les plus utilisées pour les tests statistiques, tels que (sensibilité, spécificité, précision, rappel, fscore ), voir la section 9.3.3.1 du chapitre 3 pour plus de détail

Afin de vérifier davantage la performance de notre classifieur CNN/SVM, pour ceux-ci nous avons effectué le premier test avec 100 images (10 classes) -comme exemple- et calculer la matrice de confusion pour afficher les résultats de classification de nos tests. En calculant les paramètres statistiques, la performance de notre système est évaluée et présentée dans le **tableau 4.6**. Après l'analyse, nous pouvons facilement voir la supériorité de tous les paramètres dans le cas du demi-cercle inférieur de Collarete – par apport au cercle complet- et aussi dans la matrice de confusion, voir **figure 4.11** et **figure 4.12**.

**Tableau 4.6**-Paramètres de performance

Collarete : Le cercle complet				
Sensibilité	Spécificité	Précision	Recall	Fscore
0.91	0.997	0.88	0.91	0.87
Collarete : demi-cercle inférieur (notre proposition)				
Sensibilité	Spécificité	Précision	Recall	Fscore
0.94	0.998	0.95	0.94	0.93

**Confusion Matrix**

001	3 10.0%	0 0.0%	100% 0.0%								
002	0 0.0%	3 10.0%	0 0.0%	100% 0.0%							
003	0 0.0%	0 0.0%	3 10.0%	0 0.0%	100% 0.0%						
004	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
005	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
006	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
007	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
008	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	0 0.0%	100% 0.0%
009	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	100% 0.0%
010	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	100% 0.0%
	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%
	001	002	003	004	005	006	007	008	009	010	
	<b>Target Class</b>										

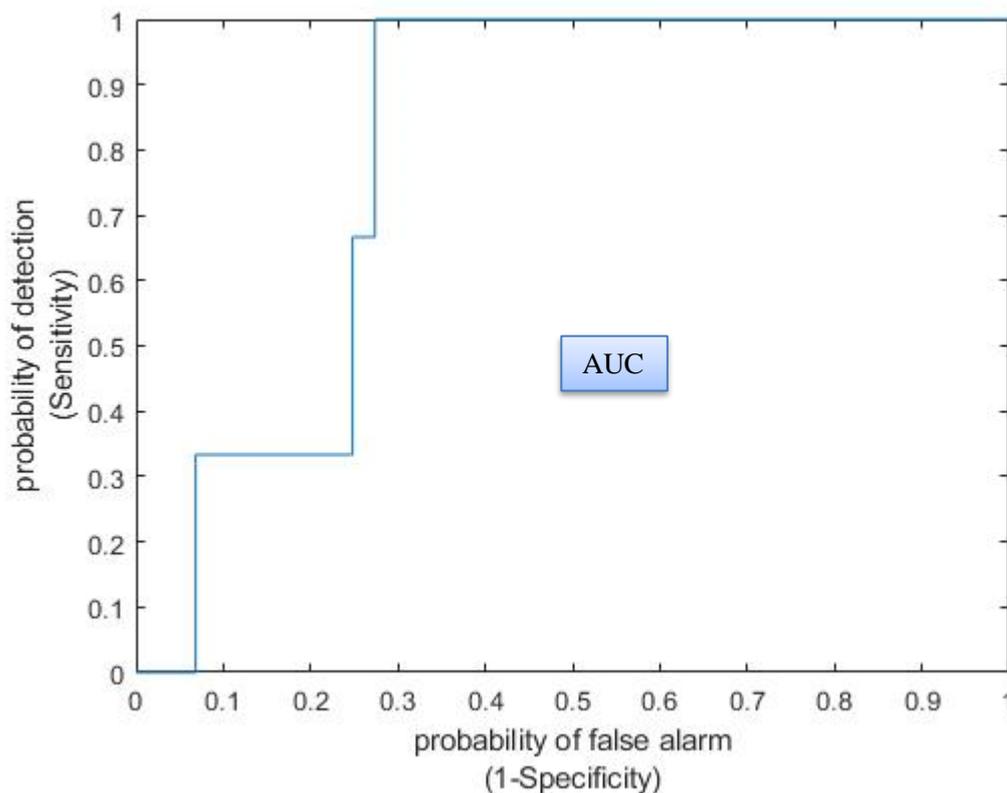
**Figure 4.11.** Matrice de confusion pour le premier test  
(Collarette : demi-cercle inférieur)

**Confusion Matrix**

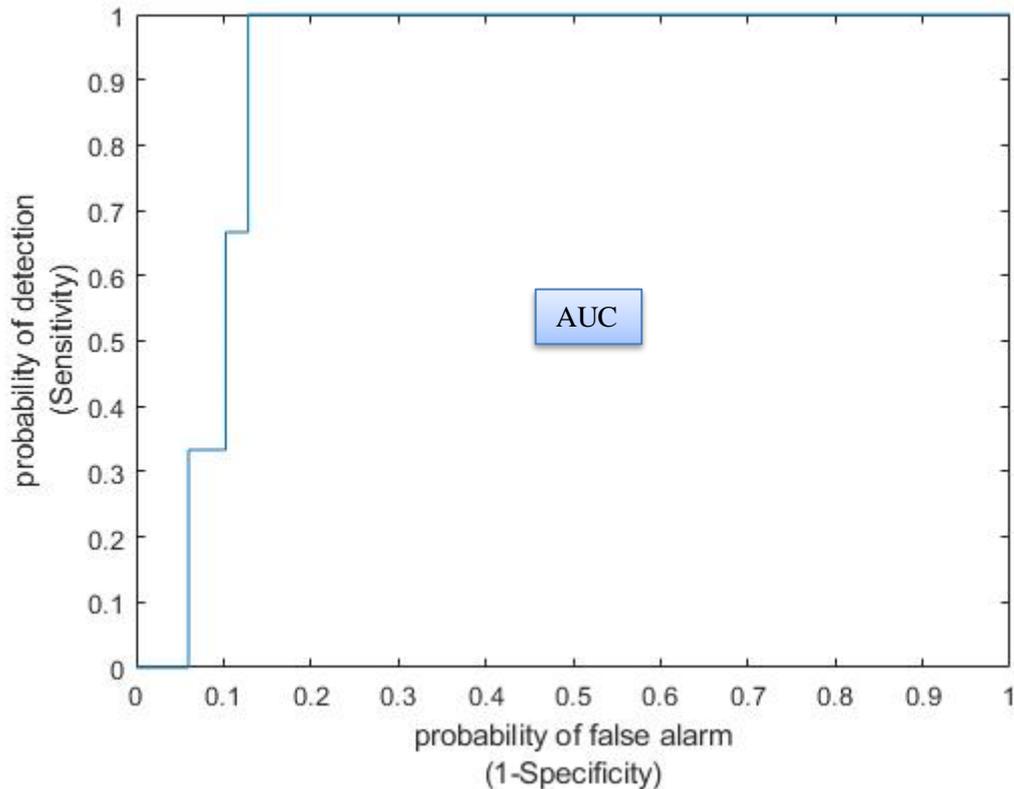
001	3 10.0%	0 0.0%	1 3.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 3.3%	0 0.0%	0 0.0%	60.0% 40.0%
002	0 0.0%	3 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
003	0 0.0%	0 0.0%	2 6.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
004	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
005	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
006	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
007	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
008	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 6.7%	0 0.0%	0 0.0%	100% 0.0%
009	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	0 0.0%	100% 0.0%
010	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 10.0%	100% 0.0%
	100% 0.0%	100% 0.0%	66.7% 33.3%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	66.7% 33.3%	100% 0.0%	100% 0.0%	93.3% 6.7%
	001	002	003	004	005	006	007	008	009	010	
	<b>Target Class</b>										

**Figure 4.12-**Matrice de confusion pour le deuxième test  
(Collarette : cercle complet)

En ce qui concerne la courbe ROC (Receiver Operating Characteristics), cette courbe représente en ordonnée la proportion de tests positifs parmi les utilisateurs authentique (la sensibilité) en fonction de la proportion de tests positifs parmi les imposteurs (complément de la spécificité ou  $1 - \text{spécificité}$ , en abscisse). Pour pouvoir déterminer la validité d'un test à partir cette courbe, il est nécessaire de calculer la surface située sous la courbe ROC appelée AUC (Area Under the Curve). Plusieurs méthodes ont été proposées dans [253] pour estimer l'AUC. Ainsi, quand le test est parfaitement discriminant, la surface sous la courbe (AUC) vaut 1 mais cela ne peut jamais être atteint. En fait, plus l'AUC est grande, meilleures sont les performances de l'algorithme, les deux **figure 4.13** et **figure 4.14** montrent que la courbe ROC de notre proposition (collerette : demi-cercle inférieur) surpasse largement celle de (collerette : le cercle complet) parce que la surface située sous la courbe ROC : (AUC) de la première est supérieur au dernière, cela traduit aussi l'efficacité de notre proposition.



**Figure 4.13-** courbe ROC avec (collerette : cercle complet)



**Figure 4.14-** courbe ROC avec (collerette : demi-cercle inférieur)

## 7. Conclusion

Dans ce chapitre, nous avons présenté une méthode de reconnaissance d'iris pour l'identification de personnes basée sur le CNN et la région de la collerette en zigzag. L'impact du choix de la caractérisation CNN sur le demi-cercle inférieur de la région de la collerette nous a permis en même temps de cibler la zone la moins bruyante de la collerette avec un vecteur de caractéristiques optimal. Nous avons utilisé un modèle Alex-Net pré-entraîné sur plus d'un million d'images comme extracteur de caractéristiques, et un SVM multi-classes pour la classification. Alex-Net avec le classificateur SVM a atteint une bonne précision avec un temps d'apprentissage assez court, de l'ordre de quelques secondes à quelques minutes. Le taux de détection de l'iris en utilisant le cercle complet de la collerette en zigzag a été comparé au taux de détection en utilisant uniquement son demi-cercle inférieur. Les images de la base de données CASIA-Iris-Lamp V4 ont été utilisées pour évaluer notre approche. La version 4 de la base de données CASIA a été appliquée pour la première fois - afin d'étudier la contribution de la zone de la collerette zigzag - avec le CNN pour la reconnaissance de l'iris.

Le choix de demi-cercle inférieur de la collerette en zigzag a eu une contribution très efficace sur l'accuracy (taux de reconnaissance) comme le montre le **tableau 4.5**, nous remarquons que l'accuracy de demi-cercle inférieur de la collerette zigzag est toujours supérieure à celle de cercle complet de collerette zigzag. La classification de l'ensemble de test est représentée par une matrice de confusion, elle montre la performance d'un classificateur sur un ensemble de données de test, si une classe est mal étiquetée comme l'autre classe parmi plusieurs classes, nous pouvons facilement l'identifier à partir d'une matrice de confusion, d'autres paramètres statistiques ont été prospectés : (Sensibilité, Spécificité, Précision, Recall, Fscore, la courbe ROC, l'espace AUC) ; les résultats confirment notre choix (collerette basse/CNN).

Nous avons choisi l'algorithme SVM car il peut être adapté aux problèmes de classification impliquant plus de 2 classes. Contrairement à un réseau de neurones qui demande beaucoup de travail pour déterminer la bonne structure et les paramètres à utiliser, les SVM sont performants même sans aucune préparation.

Le CNN facilite la tâche d'extraction des caractéristiques et rend plus facile l'apprentissage de la nouvelle tâche assignée et l'extraction de plus de caractéristiques de l'image, permettant à un classificateur SVM d'être mieux informé et d'atteindre une bonne précision. L'avantage crucial de notre approche combinée (CNN/SVM) est que nous pouvons extraire suffisamment de caractéristiques (4096 caractéristiques de la couche FC7) de chaque image en représentant le détail de chaque image à partir d'un modèle Alex-Net pré-entraîné, et profiter du SVM pour classer les caractéristiques, ce qui permet de gagner du temps.

# Conclusion générale

Au cours de cette thèse, nous nous sommes intéressés d'une manière générale à la biométrie et d'une manière plus particulière à la reconnaissance des personnes à partir des images d'iris. Cette dernière est considérée comme étant parmi les meilleures technologies biométriques. Cela est lié essentiellement aux caractéristiques et aux propriétés de l'iris. Pour cela, nous avons proposé de traiter la zone de la collerette de l'iris de la même personne afin d'avoir une meilleure performance.

Tout d'abord, nous avons exposé la méthode de segmentation à savoir la transformée de Hough en citant notre contribution à savoir localiser la partie basse de la collerette de l'iris, puisqu'elle est la région qui contient le plus d'informations pertinentes et qui est la moins affectée par le bruit (paupières, cils). Elle a été testée sur un sous-ensemble de la base de données CASIA-IrisV4-lamp. Ensuite, la collerette segmentée est normalisée par la transformation pseudo-polaire pour remédier à la non concentricité des deux bordures et à la variation de la taille de l'iris due à la dilatation/contraction de la pupille.

Les résultats principaux de cette thèse se résument ci-dessous :

- La thèse décrit et présente un système biométrique de reconnaissance de l'iris basée sur un apprentissage profond, pour y parvenir, nous avons proposé une nouvelle approche basée sur la caractérisation de la collerette basse. Les solutions proposées se basent sur l'utilisation des réseaux de neurones convolutifs profonds comme modèle de caractéristiques. Ce choix est motivé par leur capacité à résoudre des problèmes de vision et reconnaissance de forme très complexes grâce à l'extraction de caractéristiques discriminantes de haut niveau. Cependant, ces réseaux ont besoin d'un grand nombre de données d'apprentissage pour être performants. - Nous avons également étudié l'impact de l'augmentation de la taille de l'ensemble d'apprentissage et l'intégration d'une architecture de réseau par pré-apprentissage (transfert learning) pour améliorer les performances de reconnaissance.

Nos travaux ont démontré que l'efficacité des caractéristiques extraites par les CNNs dépendent de la complexité du problème cible. Plus particulièrement, les descripteurs des réseaux CNNs se sont avérés beaucoup plus efficaces que les descripteurs conçus par des méthodes classiques. Néanmoins, les CNNs sont souvent entraînés sur de grandes quantités de données pour bien extraire des caractéristiques discriminantes.

La majeure partie de ce travail a porté sur le développement de méthodes performantes d'extraction de caractéristiques d'une part, et la mise en œuvre des réseaux de neurones convolutifs, et les SVMs pour la classification d'autre part. L'association de ces techniques a permis d'obtenir des taux de reconnaissance très prometteurs, pour la base de données casia v4.

Par la suite, nous présentons les perspectives de notre travail :

- Utiliser une méthode de réduction de caractéristiques afin de remédier au problème de la dimensionnalité rencontré.
- Utiliser des bases de données plus grandes pour les expérimentations.
- Tester d'autres modèles CNN prédéfinis.
- Proposer un nouveau modèle en jouant sur la profondeur du modèle

## *Bibliographie*

- [1] Clarke, R. (1994). Human identification in information systems: Management challenges and public policy issues. *Information Technology & People*.
- [2] Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1), 4-20.
- [3] Yahiaoui, M. (2017). Modèles statistiques avancés pour la segmentation non supervisée des images dégradées de l'iris (Doctoral dissertation, Université Paris-Saclay).
- [4] Aloui, K. (2012). *Caractérisation du cerveau humain: application à la biométrie* (Doctoral dissertation, Université Paris-Est).
- [5] Morizet, N. (2009). *Reconnaissance biométrique par fusion multimodale du visage et de l'iris* (Doctoral dissertation, Télécom ParisTech).
- [6] Krichen, E. (2007). *Reconnaissance des personnes par l'iris en mode dégradé* (Doctoral dissertation, Evry, Institut national des télécommunications).
- [7] Hilal, A. (2013). *Système d'identification à partir de l'image d'iris et détermination de la localisation des informations* (Doctoral dissertation, Université de Technologie de Troyes; Université libanaise).
- [8] Masek, L. (2003). *Recognition of human iris patterns for biometric identification* (Doctoral dissertation, Master's thesis, University of Western Australia).
- [9] Bowyer, K. W., Hollingsworth, K., & Flynn, P. J. (2008). Image understanding for iris biometrics: A survey. *Computer vision and image understanding*, 110(2), 281-307.
- [10] Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1), 4-20.
- [11] Sanderson, S., & Erbetta, J. H. (2000). Authentication for secure environments based on iris scanning technology, IEE Colloquium on Visual Biometrics, Vol., No., p.p.8/1-8/7, London
- [12] Daugman, J. G. (1993). High confidence visual recognition of persons by a test of statistical independence. *IEEE transactions on pattern analysis and machine intelligence*, 15(11), 1148-1161.
- [13] Zhang, D., Jing, X. Y., & Yang, J. (2006). An Introduction to Biometrics Image Discrimination (BID). In *Biometric Image Discrimination Technologies: Computational Intelligence and its Applications Series* (pp. 1-20). IGI Global.
- [14] Arvacheh, E. M. (2006). *A study of segmentation and normalization for iris recognition systems*. University of Waterloo.
- [15] Bertillon, A. (1886). *La couleur de l'iris*. Masson.
- [16] Flom, L., & Safir, A. (1987). Iris recognition system, US" patent 4641349, patent and trademark office. *Washington, DC*.
- [17] Burghardt, B. M. I. T. (2002). Inside iris recognition. *Report on Identity, Assignment Information Security, COMS40213*.
- [18] Daugman, J. G. (1994). Biometric personal identification system based on iris analysis. *United States Patent, No. 5291560*.

## Bibliographie

- [19] Oyeniyi, J., Oyeniran, O., Omotosho, L., & Adebayo, O. (2020). Iris Recognition System: Literature Survey and Technical Overview. *International Journal of Engineering and Artificial Intelligence*, 1(3), 34-43.
- [20] Wildes, R. P. (1997). Iris recognition: an emerging biometric technology. *Proceedings of the IEEE*, 85(9), 1348-1363.
- [21] Boles, W. W., & Boashash, B. (1998). A human identification technique using images of the iris and wavelet transform. *IEEE transactions on signal processing*, 46(4), 1185-1188.
- [22] Kong, W. K., & Zhang, D. (2001, May). Accurate iris segmentation based on novel reflection and eyelash detection model. In *Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing. ISIMP 2001 (IEEE Cat. No. 01EX489)* (pp. 263-266). IEEE.
- [23] Daouk, C. H., El-Esber, L. A., Kammoun, F. D., & Al Alaoui, M. A. (2002). Iris recognition. In *IEEE ISSPIT* (No. 4, p. 558).
- [24] Seung-In, N., Bae, K., Park, Y., & Kim, J. (2003, June). A novel method to extract features for iris recognition system. In *International Conference on Audio-and Video-Based Biometric Person Authentication* (pp. 862-868). Springer, Berlin, Heidelberg.
- [25] Jan, J. D. (2004). How Iris Recognition works, *IEEE Transactions on Circuits and systems for video Technology*.
- [26] Proença, H., & Alexandre, L. A. (2006). Iris segmentation methodology for non-cooperative recognition. *IEE Proceedings-Vision, Image and Signal Processing*, 153(2), 199-205.
- [27] Daugman, J. (2007). New methods in iris recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(5), 1167-1175.
- [28] Abiyev, R. H., & Altunkaya, K. (2008). Personal iris recognition using neural network. *International Journal of Security and its Applications*, 2(2), 41-50.
- [29] Azizi, A., & Pourreza, H. R. (2009). Efficient iris recognition through improvement of feature extraction and subset selection. *arXiv preprint arXiv:0906.4789*.
- [30] Narote, S. P., Narote, A. S., & Waghmare, L. M. (2009). Iris based recognition system using wavelet transform. *IJCSNS International Journal of Computer Science and Network Security*, 9(11), 101-104.
- [31] Hussain, M. A. (2010, December). Eigenspace based accurate iris recognition system. In *2010 Annual IEEE India Conference (INDICON)* (pp. 1-3). IEEE.
- [32] Alheeti, K. M. A. (2011). Biometric iris recognition based on hybrid technique. *International Journal on Soft Computing*, 2(4), 1.
- [33] Shams, M. Y., Rashad, M. Z., Nomir, O., & El-Awady, R. M. (2011). Iris recognition based on LBP and combined LVQ classifier. *arXiv preprint arXiv:1111.1562*.
- [34] Sathish, G., Narmadha, S., Saravanan, S. V., & Maheswari, S. U. (2012). Multi-algorithmic iris recognition. *International Journal of Computer Applications*, 38(11), 13-21.
- [35] Jan, F., & Usman, I. (2014). Iris segmentation for visible wavelength and near infrared eye images. *Optik*, 125(16), 4274-4282.

## *Bibliographie*

- [36] Abikoye Oluwakemi, C., Sadiku, J. S., Adewole Kayode, S., & Jimoh Rasheed, G. (2014). Iris feature extraction for personal identification using fast wavelet transform (FWT). *structure*, 6(9).
- [37] Oyeniran, O. A., Adegoke, B. O., & Oyeniyi, J. O. (2019). MULTIPLE CLASSIFIERS: AN APPROACH IN IRIS RECOGNITION SYSTEM. *Annals. Computer Science Series*, 17(2).
- [38] Khanam, R., Haseen, Z., Rahman, N., & Singh, J. (2019). Performance analysis of iris recognition system. In *Data and Communication Networks* (pp. 159-171). Springer, Singapore.
- [39] Oyeniyi, J., Oyeniran, O., Omotosho, L., & Adebayo, O. (2020). An Enhanced Iris Feature Extraction Technique Using Continuous Wavelet Transform. *International Journal of Engineering and Artificial Intelligence*, 1(2), 16-21.
- [40] Quinn, G. W., Matey, J., Tabassi, E., & Grother, P. J. (2014). IREX V: Guidance for iris image collection.
- [41] Proença, H., & Alexandre, L. A. (2010). Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage. *Image and vision computing*, 28(1), 202-206.
- [42] Bonazza, P. (2019). *Système de sécurité biométrique multimodal par imagerie, dédié au contrôle d'accès* (Doctoral dissertation, Bourgogne Franche-Comté).
- [43] Ducasse, A. (2002). Anatomie et physiologie de l'iris. *Encycl Med Chir*.
- [44] Albadarneh, A., Albadarneh, I., & Alqatawna, J. F. (2015, November). Iris recognition system for secure authentication based on texture and shape features. In *2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)* (pp. 1-6). IEEE.
- [45] Sung, H., Lim, J., Park, J. H., & Lee, Y. (2004, August). Iris recognition using collarette boundary localization. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.* (Vol. 4, pp. 857-860). IEEE.
- [46] He, X., & Shi, P. (2005, August). An efficient iris segmentation method for recognition. In *International Conference on Pattern Recognition and Image Analysis* (pp. 120-126). Springer, Berlin, Heidelberg.
- [47] Alaslani, M. G. (2018). Convolutional neural network based feature extraction for iris recognition. *International Journal of Computer Science & Information Technology (IJCSIT)* Vol, 10.
- [48] Jeng, R. H., Chen, W. S., & Hsieh, L. (2012, June). Does Collarette of Iris Work For Recognizing Persons. In *Proceedings of the International Multi-Conference on Computing in the Global Information Technology* (pp. 24-29).
- [49] Rai, H., & Yadav, A. (2014). Iris recognition using combined support vector machine and Hamming distance approach. *Expert systems with applications*, 41(2), 588-593.
- [50] Punyani, P., Kumar, A., & Gupta, R. (2016). An optimized Iris Recognition System using MOGA followed by Combined Classifiers. *International Journal of Research in Advent Technology*, 4(3), 221-226.
- [51] Joshi, A., Gangwar, A., & Saquib, Z. (2012). Collarette Region Recognition based on Wavelets and Direct Linear Discriminant Analysis. *International Journal of Computer Applications*, 40(9), 0975-8887.

## *Bibliographie*

- [52] Frazier, M. W. (2006). *An introduction to wavelets through linear algebra*. Springer Science & Business Media.
- [53] Yu, H., & Yang, J. (2001). A direct LDA algorithm for high-dimensional data—with application to face recognition. *Pattern recognition*, 34(10), 2067-2070.
- [54] Faundra, M. R., & Sulistyaningrum, D. R. (2017). Iris Segmentation and Normalization Algorithm Based on Zigzag Collarete. In *Journal of Physics: Conference Series* (Vol. 795, No. 1, p. 012049). IOP Publishing.
- [55] Roy, K., & Bhattacharya, P. (2006, October). An iris recognition method based on zigzag collarete area and asymmetrical support vector machines. In *2006 IEEE International Conference on Systems, Man and Cybernetics* (Vol. 1, pp. 861-865). IEEE.
- [56] Jauro, S. S., & Yadav, R. (2020). An efficient interior and exterior boundary detection-based two level iris segmentation. *International Journal of Information and Computer Security*, 13(1), 97-117.
- [57] Hollingsworth, K. P. (2010). *Increased use of available image data decreases errors in iris biometrics* (Doctoral dissertation).
- [58] Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1), 4-20.
- [59] Fang, B., & Tang, Y. Y. (2006). Elastic registration for retinal images based on reconstructed vascular trees. *IEEE Transactions on Biomedical Engineering*, 53(6), 1183-1187.
- [60] Malgheet, J. R., Manshor, N. B., Affendey, L. S., & Abdul Halin, A. B. (2021). Iris recognition development techniques: a comprehensive review. *Complexity*, 2021.
- [61] Nixon, M. and Aguado, (A. S. (2008). *Feature Extraction and Image Processing*, Academic Press, Cambridge, MA, USA.
- [62] Huang, X., Ren, L., & Yang, R. (2009, June). Image deblurring for less intrusive iris capture. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1558-1565). IEEE.
- [63] Lee, S. T., Yoon, T. H., Kim, K. S., Kim, K. D., & Park, W. (2010, July). Removal of specular reflections in tooth color image by perceptron neural nets. In *2010 2nd International Conference on Signal Processing Systems* (Vol. 1, pp. V1-285). IEEE.
- [64] Santos, G., & Hoyle, E. (2012). A fusion approach to unconstrained iris recognition. *Pattern Recognition Letters*, 33(8), 984-990.
- [65] Dehkordi, A. B., & Abu-Bakar, S. A. (2013, October). Noise reduction in iris recognition using multiple thresholding. In *2013 IEEE International Conference on Signal and Image Processing Applications* (pp. 140-144). IEEE.
- [66] Liu, J., Sun, Z., & Tan, T. (2014). Distance metric learning for recognizing low-resolution iris images. *Neurocomputing*, 144, 484-492.
- [67] Raffei, A. F. M., Asmuni, H., Hassan, R., & Othman, R. M. (2015). A low lighting or contrast ratio visible iris recognition using iso-contrast limited adaptive histogram equalization. *Knowledge-Based Systems*, 74, 40-48.
- [68] Bakshi, K. A., Prasad, B. G., & Sneha, K. (2015, December). An efficient iris code storing and searching technique for Iris Recognition using non-homogeneous Kd tree. In *2015*

## ***Bibliographie***

*International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)* (pp. 34-38). IEEE.

- [69] Kumar, D., Sastry, M., & Manikantan, K. (2016, February). Iris recognition using contrast enhancement and spectrum-based feature extraction. In *2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)* (pp. 1-7). IEEE.
- [70] Baqar, M., Ghani, A., Aftab, A., Arbab, S., & Yasin, S. (2016, December). Deep belief networks for iris recognition based on contour detection. In *2016 International Conference on Open Source Systems & Technologies (ICOSST)* (pp. 72-77). IEEE.
- [71] Djoumessi, M. (2016). Iris Segmentation Using Daugman's Integro-Differential Operator. *NSF REU at Utah State University*.
- [72] Gangwar, A., Joshi, A., Singh, A., Alonso-Fernandez, F., & Bigun, J. (2016, June). IrisSeg: A fast and robust iris segmentation framework for non-ideal iris images. In *2016 international conference on biometrics (ICB)* (pp. 1-8). IEEE.
- [73] Radman, A., Zainal, N., & Suandi, S. A. (2017). Automated segmentation of iris images acquired in an unconstrained environment using HOG-SVM and GrowCut. *Digital Signal Processing*, *64*, 60-70.
- [74] Ribeiro, E., Uhl, A., Alonso-Fernandez, F., & Farrugia, R. A. (2017, August). Exploring deep learning image super-resolution for iris recognition. In *2017 25th European Signal Processing Conference (EUSIPCO)* (pp. 2176-2180). IEEE.
- [75] Arsenovic, M., Sladojevic, S., Stefanovic, D., & Anderla, A. (2018, March). Deep neural network ensemble architecture for eye movements classification. In *2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH)* (pp. 1-4). IEEE.
- [76] Gad, R., Talha, M., Abd El-Latif, A. A., Zorkany, M., Ayman, E. S., Nawal, E. F., & Muhammad, G. (2018). Iris recognition using multi-algorithmic approaches for cognitive internet of things (CIoT) framework. *Future Generation Computer Systems*, *89*, 178-191.
- [77] Susitha, N., & Subban, R. (2019). Reliable pupil detection and iris segmentation algorithm based on SPS. *Cognitive Systems Research*, *57*, 78-84.
- [78] Das, R., & Derakhshani, R. (2019, November). Using transfer learning and BPDFHE to improve ocular image recognition accuracy. In *2019 IEEE International Symposium on Technologies for Homeland Security (HST)* (pp. 1-6). IEEE.
- [79] Raffei, A. F. M., Asmuni, H., Hassan, R., & Othman, R. M. (2013). Feature extraction for different distances of visible reflection iris using multiscale sparse representation of local Radon transform. *Pattern Recognition*, *46*(10), 2622-2633.
- [80] Chai, T. Y., Goi, B. M., Tay, Y. H., Chin, W. K., & Lai, Y. L. (2015, November). Local chinese segmentation for non-ideal visible wavelength iris images. In *2015 Conference on Technologies and Applications of Artificial Intelligence (TAAI)* (pp. 506-511). IEEE.
- [81] Chen, Y., Adjouadi, M., Han, C., Wang, J., Barreto, A., Risse, N., & Andrian, J. (2010). A highly accurate and computationally efficient approach for unconstrained iris segmentation. *Image and Vision Computing*, *28*(2), 261-269.
- [82] Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001* (Vol. 1, pp. I-I). IEEE.

- [83] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1), 62-66.
- [84] Yahiaoui, M., Monfrini, E., & Dorizzi, B. (2016). Markov Chains for unsupervised segmentation of degraded NIR iris images for person recognition. *Pattern Recognition Letters*, 82, 116-123.
- [85] Winston, J. J., & Hemanth, D. J. (2019). A comprehensive review on iris image-based biometric system. *Soft Computing*, 23(19), 9361-9384.
- [86] Jan, F. (2017). Segmentation and localization schemes for non-ideal iris biometric systems. *Signal Processing*, 133, 192-212.
- [87] Alvarez-Betancourt, Y., & Garcia-Silvente, M. (2016). A keypoints-based feature extraction method for iris recognition under variable image quality conditions. *Knowledge-Based Systems*, 92, 169-182.
- [88] Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. *Graphics gems*, 474-485.
- [89] Saeedi, J., Safabakhsh, R., & Mozaffari, S. (2009, October). Document image segmentation using fuzzy classifier and the dual-tree DWT. In *2009 14th International CSI Computer Conference* (pp. 385-391). IEEE.
- [90] Bai, X. (2013). Morphological image fusion using the extracted image regions and details based on multi-scale top-hat transform and toggle contrast operator. *Digital Signal Processing*, 23(2), 542-554.
- [91] Zhang, W., Wang, C., & Xue, P. (2017, November). Application of convolution neural network in Iris recognition technology. In *2017 4th International Conference on Systems and Informatics (ICSAI)* (pp. 1169-1174). IEEE.
- [92] Zhang, X., & Shen, L. (2001). Wavelet-based homomorphic filters for image contrast enhancement. *Journal of Electronics*, 29(4), 531-533.
- [93] Chang, Y. T., Ou, C. W., Shih, T. K., & Li, Y. H. (2017, July). Performing Iris segmentation by using geodesic active contour (GAC). In *International Conference on Frontier Computing* (pp. 24-35). Springer, Singapore.
- [94] Chang, Y. T., Shih, T. K., Li, Y. H., & Kumara, W. G. C. W. (2020). Effectiveness evaluation of iris segmentation by using geodesic active contour (GAC). *The Journal of Supercomputing*, 76(3), 1628-1641.
- [95] Hofbauer, H., Jalilian, E., & Uhl, A. (2019). Exploiting superior CNN-based iris segmentation for better recognition accuracy. *Pattern Recognition Letters*, 120, 17-23.
- [96] Khan, T. M., Bailey, D. G., Khan, M. A., & Kong, Y. (2020). Real-time iris segmentation and its implementation on FPGA. *Journal of Real-Time Image Processing*, 17(5), 1089-1102.
- [97] Umer, S., Dhara, B. C., & Chanda, B. (2015). Iris recognition using multiscale morphologic features. *Pattern Recognition Letters*, 65, 67-74.
- [98] Gad, R., EL-Sayed, A., Zorkany, M., & EL-Fishawy, N. (2016). An efficient approach for simple iris localization and normalization technique. *Menoufia Journal of Electronic Engineering Research*, 25(2), 213-226.
- [99] Chawla, S., & Oberoi, A. (2011). A robust segmentation method for iris recognition. *International Journal of Advanced Research in Computer Science*, 2(5).

## Bibliographie

- [100] Gupta, R., & Kumar, A. (2013). An effective segmentation technique for noisy iris images. *International Journal of Application or Innovation in Engineering & Management (JAIEM)*, 2(12), 118-125.
- [101] Verma, P., Dubey, M., Verma, P., & Basu, S. (2012). Daughman's algorithm method for iris recognition—a biometric approach. *International journal of emerging technology and advanced engineering*, 2(6), 177-185.
- [102] Raffei, A. F. M., Asmuni, H., Hassan, R., & Othman, R. M. (2014). Fusing the line intensity profile and support vector machine for removing reflections in frontal RGB color eye images. *Information Sciences*, 276, 104-122.
- [103] Aligholizadeh, M. J., Javadi, S., Sabbaghi-Nadooshan, R., & Kangarloo, K. (2011, October). Eyelid and eyelash segmentation based on wavelet transform for iris recognition. In *2011 4th International Congress on Image and Signal Processing* (Vol. 3, pp. 1231-1235). IEEE.
- [104] Raffei, A. F. M., Asmuni, H., Hassan, R., & Othman, R. M. (2015). Frame detection using gradients fuzzy logic and morphological processing for distant color eye images in an intelligent iris recognition system. *Applied Soft Computing*, 37, 363-381.
- [105] Davies, E. R. (2012). *Computer and machine vision: theory, algorithms, practicalities*. Academic Press.
- [106] Zhao, Z., & Ajay, K. (2015). An accurate iris segmentation framework under relaxed imaging constraints using total variation model. In *Proceedings of the IEEE international conference on computer vision* (pp. 3828-3836).
- [107] Malgheet, J. R., Manshor, N. B., Affendey, L. S., & Abdul Halin, A. B. (2021). Iris recognition development techniques: a comprehensive review. *Complexity*, 2021.
- [108] Salve, S. S., & Narote, S. P. (2016, March). Iris recognition using SVM and ANN. In *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 474-478). IEEE.
- [109] Ma, L., Tan, T., Wang, Y., & Zhang, D. (2004). Efficient iris recognition by characterizing key local variations. *IEEE Transactions on Image processing*, 13(6), 739-750.
- [110] Ahuja, K., Islam, R., Barbhuiya, F. A., & Dey, K. (2016, December). A preliminary study of CNNs for iris and periocular verification in the visible spectrum. In *2016 23rd International conference on pattern recognition (ICPR)* (pp. 181-186). IEEE.
- [111] Ballard, D. H. (1987). Generalizing the Hough transform to detect arbitrary shapes. *Readings in computer vision*.
- [112] Daugman, J. G. (1993). High confidence visual recognition of persons by a test of statistical independence. *IEEE transactions on pattern analysis and machine intelligence*, 15(11), 1148-1161.
- [113] Haindl, M., & Krupička, M. (2015). Unsupervised detection of non-iris occlusions. *Pattern Recognition Letters*, 57, 60-65.
- [114] Othman, N., Dorizzi, B., & Garcia-Salicetti, S. (2016). OSIRIS: An open source iris recognition software. *Pattern Recognition Letters*, 82, 124-131.
- [115] Xu, C., & Prince, J. L. (1998). Snakes, shapes, and gradient vector flow. *IEEE Transactions on image processing*, 7(3), 359-369.

- [116] Sutra, G., Garcia-Salicetti, S., & Dorizzi, B. (2012, March). The Viterbi algorithm at different resolutions for enhanced iris segmentation. In *2012 5th IAPR International Conference on Biometrics (ICB)* (pp. 310-316). IEEE.
- [117] Benmiloud, B., & Pieczynski, W. (1995). Estimation des paramètres dans les chaînes de Markov cachées et segmentation d'images. *TS. Traitement du signal*, 12(5), 433-454.
- [118] Baum, L. E., & Eagon, J. A. (1967). An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of the American Mathematical Society*, 73(3), 360-363.
- [119] He, F., Han, Y., Wang, H., Ji, J., Liu, Y., & Ma, Z. (2017). Deep learning architecture for iris recognition based on optimal Gabor filters and deep belief network. *Journal of Electronic Imaging*, 26(2), 023005.
- [120] Liu, M., Zhou, Z., Shang, P., & Xu, D. (2019). Fuzzified image enhancement for deep learning in iris recognition. *IEEE Transactions on Fuzzy Systems*, 28(1), 92-99.
- [121] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6), 679-698.
- [122] Osorio-Roig, D., Morales-González, A., & Garea-Llano, E. (2017, November). Semantic segmentation of color eye images for improving iris segmentation. In *Iberoamerican Congress on Pattern Recognition* (pp. 466-474). Springer, Cham.
- [123] Morales-González, A., García-Reyes, E., & Sucar, L. E. (2013, November). Improving image segmentation for boosting image annotation with irregular pyramids. In *Iberoamerican Congress on Pattern Recognition* (pp. 399-406). Springer, Berlin, Heidelberg.
- [124] Masek, L. (2003). *Recognition of human iris patterns for biometric identification* (Doctoral dissertation, Master's thesis, University of Western Australia).
- [125] Nguyen, K., Fookes, C., Ross, A., & Sridharan, S. (2017). Iris recognition with off-the-shelf CNN features: A deep learning perspective. *IEEE Access*, 6, 18848-18855.
- [126] Reddy, N., Rattani, A., & Derakhshani, R. (2016, May). A robust scheme for iris segmentation in mobile environment. In *2016 IEEE Symposium on Technologies for Homeland Security (HST)* (pp. 1-6). IEEE.
- [127] Campilho, A., & Kamel, M. (Eds.). (2010). *Image Analysis and Recognition: 7th International Conference, ICIAR 2010, Póvoa de Varzin, Portugal, June 21-23, 2010, Proceedings, Part I* (Vol. 6111). Springer.
- [128] Labati, R. D., Muñoz, E., Piuri, V., Ross, A., & Scotti, F. (2019). Non-ideal iris segmentation using Polar Spline RANSAC and illumination compensation. *Computer Vision and Image Understanding*, 188, 102787.
- [129] Al-Waisy, A. S., Qahwaji, R., Ipson, S., Al-Fahdawi, S., & Nagem, T. A. (2018). A multi-biometric iris recognition system based on a deep learning approach. *Pattern Analysis and Applications*, 21(3), 783-802.
- [130] Daugman, J. (2009). How iris recognition works. In *The essential guide to image processing* (pp. 715-739). Academic Press.
- [131] Kaur, B., Singh, S., & Kumar, J. (2019). Cross-sensor iris spoofing detection using orthogonal features. *Computers & Electrical Engineering*, 73, 279-288.

- [132] Shamsi, M., & Rasouli, A. (2011). An innovative trapezium normalization for iris recognition systems. In *Int Conf Comput Softw Model* (Vol. 14, pp. 130-134).
- [133] Hilal, A., Beausery, P., & Daya, B. (2014). Elastic strips normalisation model for higher iris recognition performance. *Int Biometrics*, 3(4), 190-197.
- [134] Gong, Y., Zhang, D., Shi, P., & Yan, J. (2012). High-speed multispectral iris capture system design. *IEEE Transactions on instrumentation and measurement*, 61(7), 1966-1978.
- [135] Jan, F., Usman, I., & Khan, S. A. (2013). Robust iris biometric system for visible wavelength data. *Chinese Optics Letters*, 11(8), 081501.
- [136] Jan, F., Usman, I., Khan, S. A., & Malik, S. A. (2014). A dynamic non-circular iris localization technique for non-ideal data. *Computers & Electrical Engineering*, 40(8), 215-226.
- [137] Jan, F., & Usman, I. (2014). Iris segmentation for visible wavelength and near infrared eye images. *Optik*, 125(16), 4274-4282.
- [138] Raja, K. B., Raghavendra, R., Vemuri, V. K., & Busch, C. (2015). Smartphone based visible iris recognition using deep sparse filtering. *Pattern Recognition Letters*, 57, 33-42.
- [139] Nazmdeh, V., Mortazavi, S., Tajeddin, D., Nazmdeh, H., & Asem, M. M. (2019, January). Iris recognition; from classic to modern approaches. In *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0981-0988). IEEE.
- [140] Jang, J., Park, K. R., Kim, J., & Lee, Y. (2008). New focus assessment method for iris recognition systems. *Pattern recognition letters*, 29(13), 1759-1767.
- [141] Barpanda, S. S., Sa, P. K., Marques, O., Majhi, B., & Bakshi, S. (2018). Iris recognition with tunable filter bank based feature. *Multimedia Tools and Applications*, 77(6), 7637-7674.
- [142] Barpanda, S. S., Majhi, B., Sa, P. K., Sangaiah, A. K., & Bakshi, S. (2019). Iris feature extraction through wavelet mel-frequency cepstrum coefficients. *Optics & Laser Technology*, 110, 13-23.
- [143] Minaee, S., Abdolrashidiy, A., & Wang, Y. (2016, December). An experimental study of deep convolutional features for iris recognition. In *2016 IEEE signal processing in medicine and biology symposium (SPMB)* (pp. 1-6). IEEE.
- [144] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- [145] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [146] Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4), 433-459.
- [147] Carvalho, T., De Rezende, E. R., Alves, M. T., Balieiro, F. K., & Sovat, R. B. (2017, December). Exposing computer generated images by eye's region classification via transfer learning of VGG19 CNN. In *2017 16th IEEE international conference on machine learning and applications (ICMLA)* (pp. 866-870). IEEE.
- [148] Nguyen, K., Fookes, C., & Sridharan, S. (2020). Constrained design of deep iris networks. *IEEE Transactions on Image Processing*, 29, 7166-7175.

## *Bibliographie*

- [149] Arandjelović, R., & Zisserman, A. (2012, June). Three things everyone should know to improve object retrieval. In *2012 IEEE conference on computer vision and pattern recognition* (pp. 2911-2918). IEEE.
- [150] Nguyen, H. V., & Bai, L. (2010, November). Cosine similarity metric learning for face verification. In *Asian conference on computer vision* (pp. 709-720). Springer, Berlin, Heidelberg.
- [151] Reddy, N., Rattani, A., & Derakhshani, R. (2018, October). Ocularnet: deep patch-based ocular biometric recognition. In *2018 IEEE international symposium on technologies for homeland security (HST)* (pp. 1-6). IEEE.
- [152] Chen, Y., Wu, C., & Wang, Y. (2020). T-center: a novel feature extraction approach towards large-scale iris recognition. *IEEE Access*, 8, 32365-32375.
- [153] Marra, F., Poggi, G., Sansone, C., & Verdoliva, L. (2018). A deep learning approach for iris sensor model identification. *Pattern Recognition Letters*, 113, 46-53.
- [154] Zhao, Z., & Kumar, A. (2019). A deep learning based unified framework to detect, segment and recognize irises using spatially corresponding features. *Pattern Recognition*, 93, 546-557.
- [155] Lee, Y. W., Kim, K. W., Hoang, T. M., Arsalan, M., & Park, K. R. (2019). Deep residual CNN-based ocular recognition based on rough pupil detection in the images by NIR camera sensor. *Sensors*, 19(4), 842.
- [156] Shinde, P. P., & Shah, S. (2018, August). A review of machine learning and deep learning applications. In *2018 Fourth international conference on computing communication control and automation (ICCUBEA)* (pp. 1-6). IEEE.
- [157] Alzubi, J., Nayyar, A., & Kumar, A. (2018, November). Machine learning from theory to algorithms: an overview. In *Journal of physics: conference series* (Vol. 1142, No. 1, p. 012012). IOP Publishing.
- [158] Samuel, A. L. (1959, July). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3):210-29.
- [159] Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], 9, 381-386.
- [160] Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons*. b, 4, 51-62.
- [161] Sandhya, N., & Charanjeet, K. R. (2016). A review on machine learning techniques. *International Journal on Recent and Innovation Trends in Computing and Communication*, 4(3), 451-458.
- [162] Ayon Dey, (2016) *International Journal of Computer Science and Information Technologies (IJCSIT)*, Vol. 7 (3), 1174-1179 [www.ijcsit.com](http://www.ijcsit.com) 117410
- [163] A report by Royal Society, (April 2017), "Machine learning: the power and promise of computers that learn by example ", ISBN: 978-1-78252-259-1. and Information Technologies, Vol. 7 (3) , 2016, 1174- 1179.
- [164] Minton, S., & Zweben, M. (1993). Learning, planning, and scheduling: An overview. *Machine Learning Methods for Planning*, 1-29.
- [165] Sejnowski, T. J., & Rosenberg, C. R. (1988). NETtalk: A parallel network that learns to read aloud. In *Neurocomputing: foundations of research* (pp. 661-672).

## *Bibliographie*

- [166] Han, J., Cai, Y., & Cercone, N. (1993). Data-driven discovery of quantitative rules in relational databases. *IEEE transactions on Knowledge and Data Engineering*, 5(1), 29-40.
- [167] Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017, May). Membership inference attacks against machine learning models. In *2017 IEEE symposium on security and privacy (SP)* (pp. 3-18). IEEE.
- [168] Li, M., Andersen, D. G., Smola, A. J., & Yu, K. (2014). Communication efficient distributed machine learning with the parameter server. *Advances in Neural Information Processing Systems*, 27.
- [169] Chen, J. X. (2016). The evolution of computing: AlphaGo. *Computing in Science & Engineering*, 18(4), 4-7.
- [170] Maimon, O. Z., & Rokach, L. (2014). *Data mining with decision trees: theory and applications* (Vol. 81). World scientific.
- [171] Tan, P. N., Steinbach, M., & Kumar, V. (2013). *Introduction to Data Mining*. 6. m.
- [172] <http://www.slideshare.net/cnu/machine-learning-lecture-3>.
- [173] Lin, C. J., Weng, R. C., & Keerthi, S. S. (2008). Trust region Newton method for large-scale logistic regression. *Journal of Machine Learning Research*, 9(4).
- [174] [http://gerardnico.com/wiki/data\\_mining/simple\\_regression](http://gerardnico.com/wiki/data_mining/simple_regression)
- [175] Mitchell, T. M., & Mitchell, T. M. (1997). *Machine learning* (Vol. 1, No. 9). New York: McGraw-hill.
- [176] Jurafsky, D., & Martin, J. H. (2017). *Speech and Language Processing, 2016. Preprint draft of.*
- [177] Ng, A. (2000). CS229 Lecture notes. *CS229 Lecture notes*, 1(1), 1-3.
- [178] <http://aimotion.blogspot.mk/2011/11/machine-learning-with-python-logistic.html>
- [179] Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology*, 24(12), 1565-1567.
- [180] Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer genomics & proteomics*, 15(1), 41-51.
- [181] Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685-695.
- [182] Khan, A., Baharudin, B., Lee, L. H., & Khan, K. (2010). A review of machine learning algorithms for text-documents classification. *Journal of advances in information technology*, 1(1), 4-20.
- [183] Li, S. Z., & Jain, A. (2015). *Encyclopedia of biometrics*. Springer Publishing Company, Incorporated.
- [184] Tom, M. (1997). Mitchell: Machine Learning. *1997 Burr Ridge*, 45(37), 870-877.
- [185] Haykin, S. (2009). *Neural networks and learning machines*, 3/E. Pearson Education India.

## Bibliographie

- [186] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- [187] Hallberg, D., & Nilsagård, O. (2019). *Eye region segmentation using deep learning for Smart Eye tracking systems* (Master's thesis).
- [188] Kim, Y.(2014). Convolutional neural networks for sentence classification,” arXiv preprint arXiv:1408.5882, 2014.
- [189] Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (pp. 315-323). JMLR Workshop and Conference Proceedings.
- [190] Agarap, A. F. (2018). Deep learning using rectified linear units (ReLU). CoRR abs/1803.08375 (2018). *arXiv preprint arXiv:1803.08375*.
- [191] Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013, June). Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml* (Vol. 30, No. 1, p. 3).
- [192] Scherer, D., Müller, A., & Behnke, S. (2010, September). Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks* (pp. 92-101). Springer, Berlin, Heidelberg.
- [193] Shanmugamani, R. (2018). *Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras*. Packt Publishing Ltd.
- [194] Boureau, Y. L., Ponce, J., & LeCun, Y. (2010). A theoretical analysis of feature pooling in visual recognition. In *Proceedings of the 27th international conference on machine learning (ICML-10)* (pp. 111-118).
- [195] Alvé, J. (2017). *Improving multi-atlas segmentation methods for medical images*. Chalmers Tekniska Hogskola (Sweden).
- [196] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. CoRR abs/1505.04597 (2015). *arXiv preprint arXiv:1505.04597*.
- [197] Torrey, L., & Shavlik, J. (2010). Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (pp. 242-264). IGI global.
- [198] George, D., Shen, H., & Huerta, E. A. (2017). Deep Transfer Learning: A new deep learning glitch classification method for advanced LIGO. *arXiv preprint arXiv:1706.07446*.
- [199] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- [200] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.
- [201] Martin, S. C. O. T. T. (2019). What Is Transfer Learning?: NVIDIA Blog.
- [202] Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.
- [203] Zeiler, M. D. (2012). Adadelata: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*.

- [204] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [205] Thada, V., & Jaglan, V. (2013). Comparison of jaccard, dice, cosine similarity coefficient to find best fitness value for web retrieved documents using genetic algorithm. *International Journal of Innovations in Engineering and Technology*, 2(4), 202-205.
- [206] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- [207] Cogswell, M., Ahmed, F., Girshick, R., Zitnick, L., & Batra, D. (2015). Reducing overfitting in deep networks by decorrelating representations. *arXiv preprint arXiv:1511.06068*.
- [208] Piotrowski, A. P., & Napiorkowski, J. J. (2013). A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *Journal of Hydrology*, 476, 97-111.
- [209] Gu, H., Zhuang, Y., Pan, Y., & Chen, B. (2004, December). A new iris recognition approach for embedded system. In *International Conference on Embedded Software and Systems* (pp. 103-109). Springer, Berlin, Heidelberg.
- [210] Rai, H., & Yadav, A. (2014). Iris recognition using combined support vector machine and Hamming distance approach. *Expert systems with applications*, 41(2), 588-593.
- [211] Kim, D., Jung, Y., Toh, K. A., Son, B., & Kim, J. (2016). An empirical study on iris recognition in a mobile phone. *Expert systems with Applications*, 54, 328-339.
- [212] Gangwar, A., & Joshi, A. (2016, September). DeeplrisNet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition. In *2016 IEEE international conference on image processing (ICIP)* (pp. 2301-2305). IEEE.
- [213] Zhao, Z., & Kumar, A. (2017). Towards more accurate iris recognition using deeply learned spatially corresponding features. In *Proceedings of the IEEE international conference on computer vision* (pp. 3809-3818).
- [214] Zhou, Z., Du, E. Y., Thomas, N. L., & Delp, E. J. (2013). A comprehensive multimodal eye recognition. *Signal, Image and Video Processing*, 7(4), 619-631.
- [215] Mariño, R. Á., Álvarez, F. H., & Encinas, L. H. (2012). A crypto-biometric scheme based on iris-templates with fuzzy extractors. *Information Sciences*, 195, 91-102.
- [216] Dhage, S. S., Hegde, S. S., Manikantan, K., & Ramachandran, S. (2015). DWT-based feature extraction and radon transform based contrast enhancement for improved iris recognition. *Procedia Computer Science*, 45, 256-265.
- [217] Elgamal, M., & Al-Biqami, N. (2013). An efficient feature extraction method for iris recognition based on wavelet transformation. *Int. J. Comput. Inf. Technol*, 2(03), 521-527.
- [218] Badejo, J. A., Atayero, A. A., & Ibiyemi, T. S. (2016, December). A robust preprocessing algorithm for iris segmentation from low contrast eye images. In *2016 Future Technologies Conference (FTC)* (pp. 567-576). IEEE.
- [219] Xia, S., Chen, P., Zhang, J., Li, X., & Wang, B. (2017). Utilization of rotation-invariant uniform LBP histogram distribution and statistics of connected regions in automatic image annotation based on multi-label learning. *Neurocomputing*, 228, 11-18.

- [220] Tapia, J., & Aravena, C. (2017). Gender classification from NIR iris images using deep learning. In *Deep learning for biometrics* (pp. 219-239). Springer, Cham.
- [221] Wang, Z., Li, C., Shao, H., & Sun, J. (2018). Eye recognition with mixed convolutional and residual network (MiCoRe-Net). *IEEE Access*, 6, 17905-17912.
- [222] Zhao, T., Liu, Y., Huo, G., & Zhu, X. (2019). A deep learning iris recognition method based on capsule network architecture. *IEEE Access*, 7, 49691-49701.
- [223] Lozej, J., Štepec, D., Štruc, V., & Peer, P. (2019, May). Influence of segmentation on deep iris recognition performance. In *2019 7th International Workshop on Biometrics and Forensics (IWBF)* (pp. 1-6). IEEE.
- [224] Lu, Y., & Pan, H. (2019). Application of Iris images in racial classifications based on dilate convolution and residual network. *IEEE Access*, 7, 182395-182405.
- [225] [225] Tounsi, S., Boukari, K., & Souahi, A. (2022). The impact of collarette region-based convolutional neural network for iris recognition. *International journal of electrical and computer engineering systems*, 13(1), 37-47.
- [226] "Individual Biometrics: Iris Scan" 5 July 05, National Center for State Courts 6 July 2006, <<https://ctt.nsc.dni.us/biomet%20web/BMIris.htm>>.
- [227] <http://www.cbsr.ia.ac.cn/china/Iris%20Databases%20CH.asp>, (dernier accès 2020).
- [228] CASIA iris database, <http://www.idealtest.org/findTotalDbByMode.do?mode=Iris> (dernier accès 2012).
- [229] UBIRIS iris database, <http://iris.di.ubi.pt/>, (dernier accès 2012).
- [230] IITD iris database, <http://www.iitd.ac.in/>, (dernier accès 2012).
- [231] MMU iris databases, <http://pesona.mmu.edu.my/~ccteo/>, ((dernier accès 2012).
- [232] Daugman, J. G. (1993). High confidence visual recognition of persons by a test of statistical independence. *IEEE transactions on pattern analysis and machine intelligence*, 15(11), 1148-1161.
- [233] Alaslani, M. G. (2018). Convolutional neural network based feature extraction for iris recognition. *International Journal of Computer Science & Information Technology (IJCSIT) Vol, 10*.
- [234] Sankowski, W., Grabowski, K., Napieralska, M., Zubert, M., & Napieralski, A. (2010). Reliable algorithm for iris segmentation in eye image. *Image and vision computing*, 28(2), 231-237.
- [235] De Almeida, P. (2010). A knowledge-based approach to the iris segmentation problem. *Image and Vision Computing*, 28(2), 238-245.
- [236] He, Z., Tan, T., Sun, Z., & Qiu, X. (2008). Toward accurate and fast iris segmentation for iris biometrics. *IEEE transactions on pattern analysis and machine intelligence*, 31(9), 1670-1684.
- [237] Jan, F., Usman, I., & Agha, S. (2012). Iris localization in frontal eye images for less constrained iris recognition systems. *Digital Signal Processing*, 22(6), 971-986.
- [238] Singh, N., Gandhi, D., & Singh, K. P. (2011). Iris recognition system using a canny edge detection and a circular hough transform. *International Journal of Advances in Engineering & Technology*, 1(2), 221.

- [239] Sangwan, S., & Rani, R. (2015). A review on: Iris Recognition. *IJCSIT) International Journal of Computer Science and Information Technologies*, 6, 3871-3873.
- [240] Chouhan, B., & Shukla, S. (2011). Iris Recognition System using canny edge detection for Biometric Identification. *International Journal of Engineering Science and Technology (IJEST)*, 3(1), 31-35.
- [241] Argialas, D. P., & Mavrantza, O. D. (2004). Comparison of edge detection and Hough transform techniques for the extraction of geologic features. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34(Part XXX).
- [242] Al-Waisy, A. S., Qahwaji, R., Ipson, S., & Al-Fahdawi, S. (2015, October). A fast and accurate iris localization technique for healthcare security system. In *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing* (pp. 1028-1034). IEEE.
- [243] Lu, Y., & Pan, H. (2019). Application of Iris images in racial classifications based on dilate convolution and residual network. *IEEE Access*, 7, 182395-182405.
- [244] Nguyen, K., Fookes, C., Ross, A., & Sridharan, S. (2017). Iris recognition with off-the-shelf CNN features: A deep learning perspective. *IEEE Access*, 6, 18848-18855.
- [245] Al-Waisy, A. S., Qahwaji, R., Ipson, S., Al-Fahdawi, S., & Nagem, T. A. (2018). A multi-biometric iris recognition system based on a deep learning approach. *Pattern Analysis and Applications*, 21(3), 783-802.
- [246] Park, G. H., Cho, H. H., & Choi, M. R. (2008). A contrast enhancement method using dynamic range separate histogram equalization. *IEEE Transactions on Consumer Electronics*, 54(4), 1981-1987.
- [247] Yadav, G., Maheshwari, S., & Agarwal, A. (2014, September). Contrast limited adaptive histogram equalization based enhancement for real time video system. In *2014 international conference on advances in computing, communications and informatics (ICACCI)* (pp. 2392-2397). IEEE.
- [248] Kyriakopoulou, M. (2020). Histogram Equalization on Medical Images: CLAHE implementation on CT images.
- [249] Romero, A., Gatta, C., & Camps-Valls, G. (2015). Unsupervised deep feature extraction for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 54(3), 1349-1362.
- [250] Oyedotun, O., & Khashman, A. (2017). Iris nevus diagnosis: convolutional neural network and deep belief network. *Turkish Journal of Electrical Engineering and Computer Sciences*, 25(2), 1106-1115.
- [251] <https://towardsdatascience.com/deep-learning-googlenet-explained-de8861c82765> (dernier accès 2022 )
- [252] <https://medium.com/@RaghavPrabhu/cnn-architectures-lexnet-alexnet-vgg-googlenet-and-resnet-7c81c017b848>( dernier accès 2022)
- [253] Bhatnagar, J., & Kumar, A. (2009). On estimating performance indices for biometric identification. *Pattern recognition*, 42(9), 1803-1815.