



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

**MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE**

UNIVERSITE MOULOUD MAMMERIDE TIZI-OUZOU

FACULTE DE GENIE ELECTRIQUE ET INFORMATIQUE

DEPARTEMENT D'INFORMATIQUE

Mémoire de Fin d'Etudes

En vue de l'obtention du diplôme de master en informatique

Spécialité : *Systemes Informatiques*

Thème

**La Classification d'images d'insectes ravageurs en utilisant le
*Deep Learning***

Réalisé par :

LOUNIS Katia

MOUSSI Dahbia

Encadré par :

Promotrice : M^{me} AOUDJIT. R

Co-promotrice : M^{me} AIT ISSAD. H

Les membres du jury

Présidente: M^{me} BELATTAF

Examinatrice: M^{me} CHEMOUN

Année universitaire : 2019-2020

Remerciement

Nous tenons à remercier en premier lieu le bon Dieu, de nous avoir donné la force, la patience, la volonté afin d'accomplir ce modeste travail qui présente le fruit de plusieurs années de sacrifices.

Nous avons l'honneur et le plaisir de présenter notre profonde gratitude et nos sincères remerciements à notre promotrice : **Mme Aoudjit** et notre Copromotrice **Mme Ait Issad** pour leur précieuse aide, leurs orientations et le temps qu'elles ont accordé à notre encadrement.

Nous tenons à remercier également l'ensemble de nos professeurs chacun avec son nom pour la qualité de l'enseignement qu'ils nous ont prodigué au cours des cinq années passées à l'université de **Mouloud Mammeri**.

Et aussi l'ensemble de tous les fonctionnaires du département d'informatique.

Nous remercions par ailleurs vivement les membres du jury de nous avoir fait l'honneur d'examiner notre travail.

Arrivés au terme de la rédaction de ce mémoire, nous exprimons notre gratitude pour toutes les personnes qui ont participé de près ou de loin à notre réussite.

Dédicaces

Je tiens à dédier ce travail à

Mes très chers **parents**, qui m'ont comblé avec leur amour, sacrifices et précieux conseils. Qui m'ont soutenu moralement et financièrement jusqu' à ce jour, qui m'ont encouragé tout au long de mon parcours d'études. Que ce travail soit pour eux un modeste témoignage de ma Profonde gratitude. Qu'ALLAH les préserve, et leurs accorde santé, bonheur et longue vie .

Mon cher unique frère **Moumouh** qui est mon exemple et mes adorables sœurs **Fe tta** et **Le ticia** qui m'ont épaulé durant tout au long de mes études et qui ont été toujours là pour moi .

Une dédicace particulière au deux personnes les plus chers à mon coeur, Tante **Aldj ia Nadji** , Tonton **Said** pour leurs encouragement permanent , leurs conseils que j'appliquerai toujours, pour leurs amour sincère, pour leurs gentillesse , et je remercie surtout tonton Said qui m'a vraiment soutenu et qui essaye toujours de trouver un moyen pour me déstresser , merci pour tout les fous rires , que dieu vous garde pour moi les deux.

A ma sœur de cœur, Ma meilleure amie **Lilia** qui a été à mes cotés dans mes mauvais moments avant les bons, à tout nos beaux souvenirs ensemble.

A la mémoire de mon très cher Tonton **Achour**, que dieu l'accueille dans ce vaste paradis, qui m'a toujours aimé, et son seul vœux était de me voir réussir mes études. Je dédie ce travail pour toi.

A mon professeur du lycée, Madame **S.Khris** qui m'a toujours encouragé tout au long de mes études.

A mon cher binôme et ma chère amie **Katia**, avec laquelle J'ai vécu des souvenirs inoubliables durant les cinq ans d'études à l'université. Avec elle j'ai partagé le stress et la pression des études, la saveur de la réussite et beaucoup de fou rire. Que de la réussite dans le reste de sa vie et dans tous les domaines.

A mes chers grands parents, mes tantes et oncles.

A mes chers voisins, que j'adore beaucoup, surtout ma chère Lynda.

Tous mes ami (es) proches qui mon soutenu et qui ont été des personnes inoubliables et formidables.

Dahbia

Dédicaces

Je tiens à dédier ce travail à

Mes très chers **parents**, qui m'ont comblé avec leur amour, sacrifices et précieux conseils. Qui m'ont soutenu moralement et financièrement jusqu' à ce jour, qui m'ont encouragé tout au long de mon parcours d'études. Que ce travail soit pour eux un modeste témoignage de ma Profonde gratitude. Qu'*ALLAH* les préserve, et leurs accorde santé, bonheur et longue vie.

Mes adorables frères et sœurs qui m'ont épaulé durant tout au long de mes études qui ont été toujours là pour moi ; sans eux je ne serais jamais arrivée là où je suis aujourd'hui.

Mes chères belles sœurs que je considère comme étant des vraies sœurs, qui m'ont toujours étaient d'un énorme soutien.

Mes adorables petits neveux et nièces qui me prennent toujours comme exemple dans leurs vies. Je les adore énormément et je leur souhaite toute la réussite.

Aux deux personnes les plus chères à mon cœur, mon cher **Malik** et ma sœur de cœur **Kenza** qui ont été à mes côtés dans mes mauvais moments avant les bons, et qui m'ont supporté dans mes crises de larmes durant les préparations de ce mémoire. Que Dieu les garde pour moi à jamais.

Une dédicace particulière, à mon cher binôme et ma chère amie **Dahbia**, avec laquelle J'ai vécu des souvenirs inoubliables durant les cinq ans d'études à l'université. Avec elle j'ai partagé le stress et la pression des études, la saveur de la réussite et beaucoup de faux rires. Que de la réussite dans le reste de sa vie et dans tous les domaines.

Tous mes amis(es) proches qui m'ont soutenu et qui ont été des personnes inoubliables et formidables. En particulier **Tinhinane** et **Mohand** qui m'ont beaucoup aidé.

Katia



Liste des figures, tableaux et d'abréviations

Table des figures

• Chapitre 1 : Principes et concepts de base de l'intelligence artificielle

-Figure 1.1 : Illustration d'un panier d'achat intelligent.....07

• Chapitre 2: Outils et algorithmes mis en œuvre pour l'IA

-Figure 2.1 : Le rapport entre L'IA, Machine Learning et le *Deep Learning*.....17

-Figure 2.2 : Architecture de base d'un réseau de neurones artificiel.....18

- Figure 2.3 : Principe de fonctionnement d'un réseau de neurones artificiel.....19

- Figure 2.4 : Le processus d'auto-apprentissage pour l'identification d'un chat.....22

- Figure 2.5 : Le processus d'auto-apprentissage pour l'identification d'une personne.....23

• Chapitre 3: La classification d'images avec le Deep Learning

-Figures 3.1 : Les Domaines d'application de la classification d'images.....27

- Figures 3.2 : Voisinage à 4.....29

- Figures 3.3: Voisinage à 8.....29

- Figures 3.4 : Exemple de Matrice de confusion.....32

-Figures 3.5: Exemple d'illustration de la courbe ROC.....33

- Figures 3.6 : Architecture de base d'un réseau neuronal.....36

- Figures 3.7 : La partie convolution d'un CNN.....37

- Figures 3.8 : Parcours de la fenêtre de filtre sur l'image.....38

- Figures 3.9 : Le principe de convolution.....38

-Figure 3.10 : Le fonctionnement de la fonction *RELU*.....39

- Figure 3.11: Le principe de Pooling avec un filtre 2*2 et un pas de 2.....40

-Figure 3.12: La couche *Fullyconnected*.....41

- Figure 3.13: Le fonctionnement d'un réseau de neurones à convolution.....42

- Figure 3.14: Visualisation de l'Architecture du réseau LeNet-5.....43

- Figure 3.15: Visualisation de l'architecture du réseau AlexNet.....	44
- Figure 3.16: Visualisation de l'architecture du réseau VGG16.....	44
- Figure 3.17: Visualisation de l'architecture du réseau GoogleNet (1).....	45
- Figure 3.18: Visualisation de l'architecture du réseau GoogleNet (2).....	45
- Figure 3.19: Visualisation de l'architecture du réseau ResNet.....	46
- Figure 3.20 : Classification des gagnants du Défi de reconnaissance visuelle à grande échelle ImageNet (ILSVRC).....	46

• **Chapitre 4: Implémentation**

- Figure 4.1: Classe 1: Papillon (Lépidoptères)	51
- Figure 4.2 : Classe 2: Mouche (Diptères).....	51
- Figure 4.3 : Classe 3: Abeille (Hyménoptères).....	51
- Figure 4.4 : Classe 4: Sauterelle (Orthoptères).....	52
- Figure 4.5 : Classe 5: Fourmies blanche(Isoptères).....	52
- Figure 4.6 : Classe 6: Cafard (Blattodea)	52
- Figure 4.7 : Le fonctionnement de la méthode k-fold CV	54
- Figure 4.8 : Visualisation de l'architecture du réseau AlexNet après modifications.....	56
- Figure 4.9 : Visualisation de l'architecture du réseau GoogleNet après modifications.....	57
- Figure 4.10 : Progression de l'apprentissage du réseau AlexNet sans augmentation des données.....	60
- Figure 4.11 : Matrice de confusion d'AlexNet pour l'ensemble de test.....	61
- Figure 4.12 : Progression de l'apprentissage du réseau GoogleNet sans augmentation des données	62
- Figure 4.13 : Matrice de confusion de GoogleNet pour l'ensemble de test.....	62
- Figure 4.14 : Progression de l'apprentissage du réseau AlexNet avec augmentation des données.....	64
- Figure 4.15 : Matrice de confusion d'AlexNet pour l'ensemble de test.....	64
- Figure 4.16 : Progression de l'apprentissage du réseau GoogleNet avec augmentation des données.....	65
- Figure 4.17 : Matrice de confusion de GoogleNet pour les données de test.....	66

- Figure 4.18 :** Courbe ROC d’AlexNet sans augmentation de l’ensemble de données.....**69**
- Figure 4.19 :** Courbe ROC de GoogleNet sans augmentation de l’ensemble de données...**69**
- Figure 4.20 :** Courbe ROC d’AlexNet avec augmentation de l’ensemble de données.....**71**
- Figure 4.21 :** Courbe ROC de GoogleNet avec augmentation de l’ensemble de données...**71**

Liste des Tableaux

- Tableau 4.1 :** Tableau représentant les paramètres des modèles AlexNet et GoogleNet.....**57**
- Tableau 4.2:** Les valeurs des taux de précision et d’erreur des modèles avec et sans augmentation.....**67**
- Tableau 4.3:** Tableau comparatif des valeurs des indicateurs de performance de GoogleNet et AlexNet.....**68**

Les Abréviations

- IA** : Intelligence Artificielle.
- IAS**: Integrated Automation System.
- GPS**: Global Positioning System. -
- SQL** : Structured Query Language.
- ANN** : Artificial Neuronal Network.
- RVB** : Rouge Vert Bleu.
- PPP** : Point Par Puce.
- **ROC**: Received Operating Characterist.
- **CNN**: Convolutional neural network.
- **MLP**: Multi Layer Perceptron.
- ILSVRC** : ImageNet Large Scale Visual Recognition Challenge.
- CV** : Cross-Validation.



Table des matières

Table des matières

Introduction générale

Chapitre 1 : Principes et concepts de base de l'intelligence artificielle

1.1 Introduction.....	01
1.2 Principe de l'intelligence artificielle.....	02
1.3 Qu'est-ce que l'intelligence artificielle ?.....	02
1.4 L'utilisation de l'IA.....	04
1.5 Domaines d'application.....	05
1.5.1 La santé.....	05
1.5.2 Le secteur bancaire et financier.....	06
1.5.3 Le transport.....	06
1.5.4 Le commerce.....	06
1.5.5 L'industrie.....	07
1.5.6 L'agriculture intelligente.....	07
1.5.7 Les médias.....	08
1.5.8 Le droit.....	08
1.6 Les avantages, inconvénients et limites de l'intelligence artificielle.....	09
1.6.1 Les avantages.....	09
1.6.2 Les inconvénients.....	10
1.6.3 Les limites.....	10
1.7 Conclusion.....	11

Chapitre 2 : Outils et algorithmes mis en œuvre pour l'IA

2.1 Introduction.....	12
2.2 Techniques de l'apprentissage.....	13
2.2.1 L'apprentissage automatique (<u>machine learning</u>).....	13
2.2.2 Types d'apprentissage.....	13
2.2.2.1) Apprentissage supervisé.....	14
2.2.2.2) Apprentissage non supervisé.....	15
2.2.2.3) Apprentissage par renforcement.....	15
2.2.2.4) Apprentissage profond (<i>Deep Learning</i>).....	16
2.3 Deep Learning (Apprentissage Profond).....	16
2.3.1 Définition et origines.....	16
2.3.2 Réseaux de neurones artificiels.....	17
2.3.2.1 Définition.....	17
2.3.2.2 Architecture d'un réseau de neurones artificiel.....	17
2.3.2.3 Principe de fonctionnement d'un réseau de neurones artificiel.....	18
2.3.2.4 Types des réseaux de neurones artificiels.....	19
2.3.3 Domaines d'applications du Deep Learning.....	20
2.3.4 Exemples du Deep Learning.....	21
2.4 Conclusion.....	24

Chapitre 3 : La classification d'images

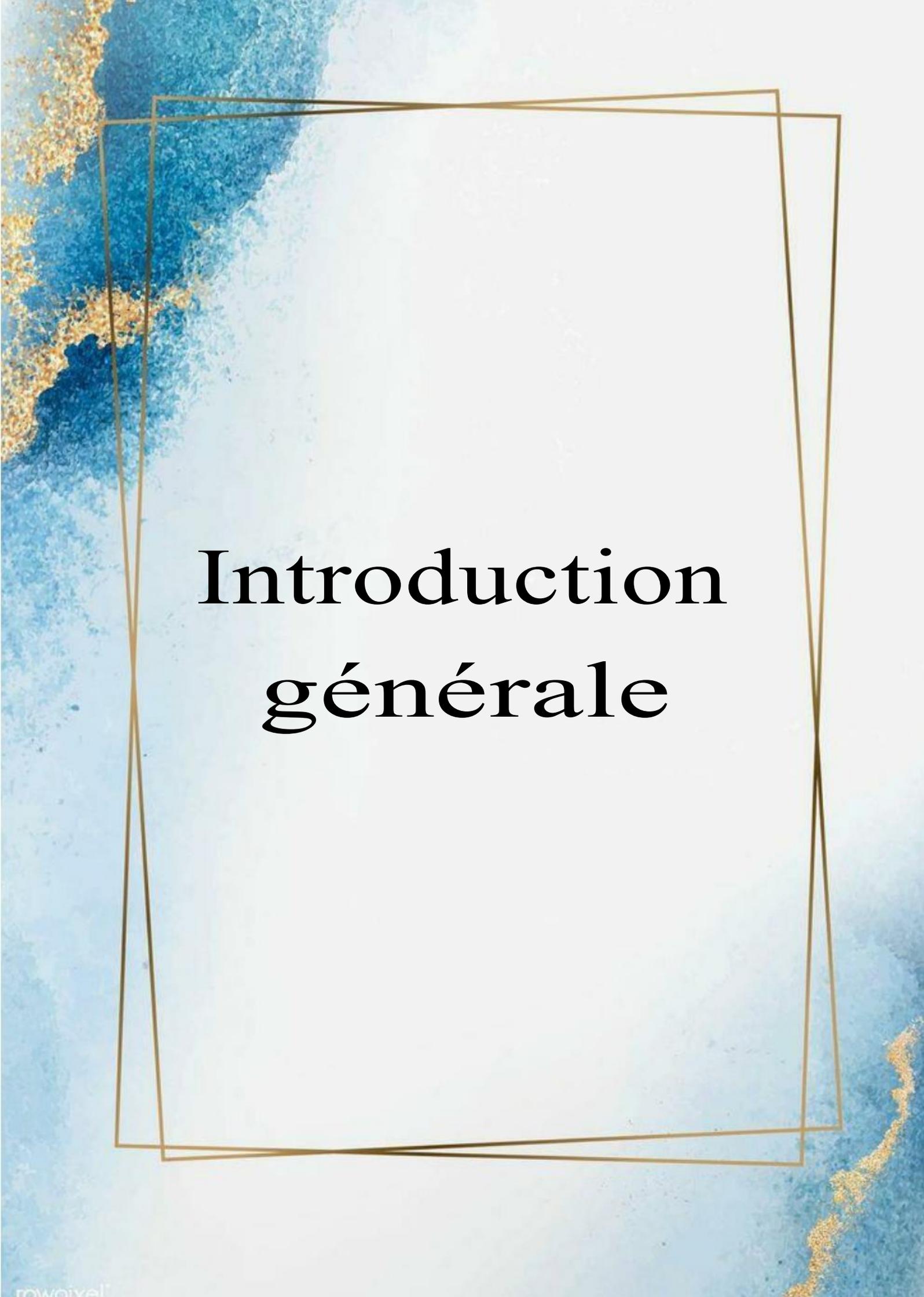
3.1 Introduction.....	25
3.2 Les motivations de la classification des images.....	26
3.3 Les domaines d'application de la classification d'images.....	26
3.4 Notions de bases.....	27
3.4.1 Définition d'une image.....	27
3.4.2 Les différents types du format d'image.....	28
3.4.3 Caractéristiques de l'image.....	28
3.4.3.1 Pixel.....	28
3.4.3.2 Dimension Résolution.....	28
3.4.3.3 Voisinage.....	29
3.4.3.4 Niveau de gris.....	29
3.4.3.5 Luminance.....	29
3.4.3.6 Bruit.....	30
3.4.3.7 Contour.....	30
3.5 Méthodes de classification.....	30
3.5.1 Méthodes supervisées.....	30
3.5.2 Méthodes non supervisées.....	31
3.6 Indicateurs de performance en classification.....	31
3.6.1 Matrice de confusion.....	31
3.6.2 Courbe ROC.....	32
3.6.3 La précision.....	33
3.6.4 Le rappel.....	33
3.6.5 La justesse.....	33
3.6.6 La spécificité.....	33
3.6.7 Le F-score.....	34
3.7 Classification des images et l'apprentissage machine.....	34
3.8 Classification des images et les réseaux de neurones.....	35
3.9 Les réseaux de neurones convolutifs(CNN).....	35
3.9.1 Définition d'un réseau neuronal convolutif.....	36
3.9.2 Architecture et fonctionnement d'un CNN.....	36
3.9.2.1 Partie convolution.....	36
3.9.2.2 Partie classification.....	41

3.9.3 Les modèles prédéfinis.....	43
3.10 Conclusion.....	47

Chapitre 4 : Implémentation

4.1 Introduction.....	48
4.2 Les Outils logiciels utilisés.....	48
4.2.1 Définition de Matlab.....	48
4.2.2 Environnement de développement.....	49
4.2.3 Boîte à outils Matlab (Toolbox).....	49
4.2.4 Matlab et le Deep Learning.....	50
4.3 Configuration matérielle utilisée.....	51
4.4 Description de l'ensemble de données.....	51
4.5 Architectures et paramètres des modèles CNNs utilisés dans l'implémentation.....	55
A) Architecture d'AlexNet et de GoogleNet.....	55
B) Les paramètres d'AlexNet et de GoogleNet.....	57
4.6 Indicateurs de performances.....	58
4.7 Résultats et Discussions.....	60
4.8 Conclusion.....	72

Conclusion générale

The background features a soft, painterly texture in shades of light blue and white, with irregular splatters of a darker blue and shimmering gold particles. A thin, gold-colored rectangular frame is centered on the page, with its lines slightly offset from each other to create a subtle 3D effect.

Introduction générale

Introduction générale

Actuellement, nous sommes devant un début d'une nouvelle notion technologique qui est l'intelligence artificielle qui n'épargnera aucun secteur. La présence d'un grand nombre d'algorithmes, la croissance de la puissance de calcul, la disponibilité d'une énorme quantité de données stockée dans des bases reliées entre elles grâce à des réseaux de plus en plus rapide, disponibles sur des machines connectées en permanence va avoir un impact immense sur la vie humaine.

L'intelligence artificielle est devenue donc un sujet d'une importance majeure dans notre époque, car cette évolution technologique a pu résoudre plusieurs problèmes complexes, chose que les autres technologies n'ont pas pu réussir à faire. Cette notion n'est pas rupture technologique mais plutôt elle est considérée comme une continuité de l'informatique.

L'intelligence artificielle, considérée comme étant une intelligence fournie par une machine, est utilisée pour développer et faire progresser de nombreux domaines, y compris la santé, l'industrie, le commerce et le domaine d'agriculture. Dans ce dernier, l'identification des insectes ravageurs est l'une des techniques clés de la gestion des insectes. En déterminant le type et le nombre de ces derniers dans une ferme, les épidémies d'insectes et les maladies des cultures peuvent être évitées. Pour notre travail, on souhaite pouvoir classifier les images d'insectes nuisibles, selon des classes déjà connues afin de faciliter leurs identifications. Pour cela il est nécessaire de développer une méthode de classification automatique rapide et fiable. Cela est possible grâce à un ensemble d'algorithmes du Deep Learning qui est une dérivée de l'apprentissage automatique qui constitue une sous discipline de l'intelligence artificielle. Ce travail sera le sujet de notre mémoire qui porte le thème " la classification d'images d'insectes ravageurs par le Deep Learning".

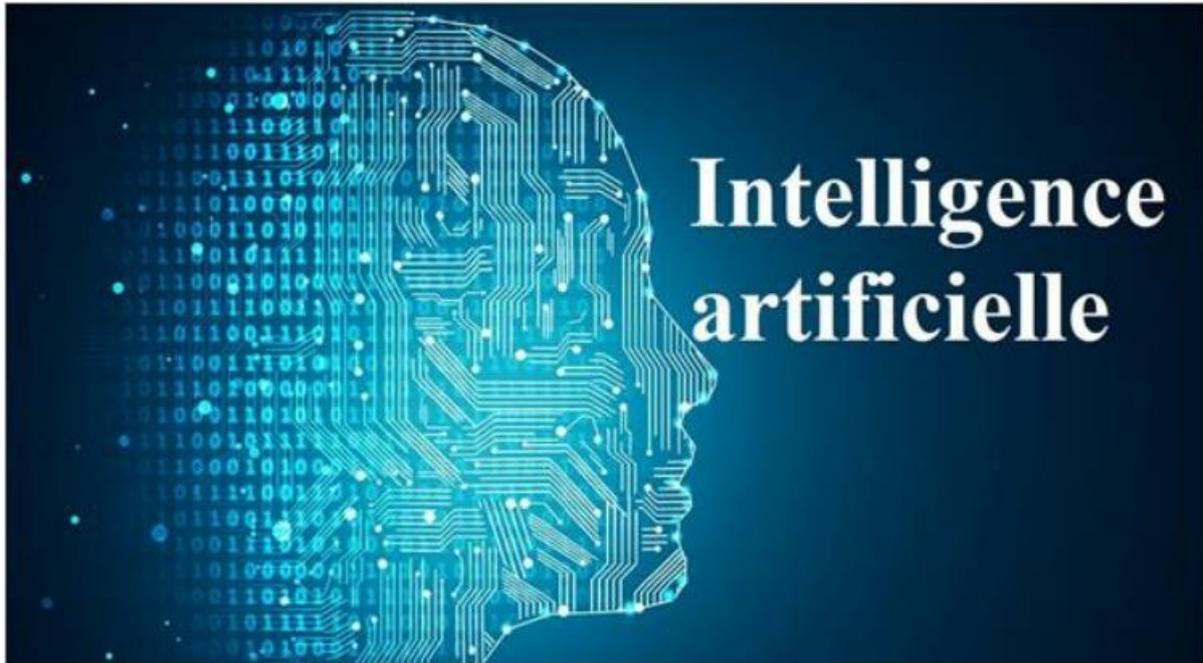
Pour cela nous avons structuré notre travail en quatre grands chapitres qui sont les suivants :

- Nous commençons par les principes et les concepts de base de l'intelligence artificielle comme premier chapitre, afin de se familiariser avec le concept de l'intelligence artificielle.

- Suivi par les outils et algorithmes mis en œuvre pour l'IA dans le deuxième chapitre, chose qui va permettre d'approfondir les premières notions.
- Pour le troisième chapitre nous allons aborder la notion de la classification d'images, où nous allons pouvoir appliquer les techniques de l'apprentissage profond.
- Enfin un quatrième chapitre qui va comporter l'implémentation d'un exemple d'illustration de la notion de classification d'images avec *Deep Learning*.

Chapitre I

Principes et concepts de base de l'intelligence artificielle



1.1) Introduction:

Avec l'arrivée d'Internet, le monde est devenu hyper connecté dans lequel chaque objet manipulé (les voitures, les réfrigérateurs, les vêtements, les réseaux sociaux...) génèrent des millions de données supplémentaires chaque jour, qui s'ajoutent à un océan de données sans fin.

Toutes ces données peuvent être utilisées pour garantir des services personnalisés et immédiats attendus. Mais comment peut-on transformer un tel océan de données sans fond en flux régulier d'informations pertinentes pour répondre à ces attentes ? La réponse est l'intelligence artificielle (IA).

Dans ce premier chapitre nous aborderons les principes et les concepts de base de l'intelligence artificielle, sa définition, son principe, son d'utilisation et les domaines d'applications de cette dernière et encore ses avantages, inconvénients et limites.

1.2) Principe de l'intelligence artificielle:

L'intelligence artificielle est présente partout, nous nous en servons déjà dans notre vie de tous les jours mais sans s'en rendre compte, que ce soit en utilisant une application de commande vocale, des jeux sur un ordinateur ou encore des outils de domotique. Son principe est de viser à reproduire le fonctionnement du cerveau humain sur les machines, et cela en utilisant un ensemble d'algorithmes d'apprentissage, qui permettent d'apporter à ces machines une sorte d'intelligence.

1.3) Qu'est-ce que l'intelligence artificielle ?

Les progrès des neurosciences et l'exploration du cerveau humain soulèvent une question remplie à la fois d'espoir et de crainte: serait-il possible un jour de reproduire l'intégralité du cerveau humain ? Aujourd'hui déjà, les ordinateurs sont supérieurs par rapport au cerveau humain en ce qui concerne la puissance de calcul, bien que le degré de complexité présenté par le cerveau humain soit bien plus élevé. Ces doutes soulèvent la question de l'intelligence artificielle.



Plusieurs définitions ont été attribuées à cette notion parmi elles nous pouvons citer ces deux:

❖ Définition 1

L'intelligence artificielle ou IA est une discipline scientifique et technologique qui vise à exécuter par machines (ordinateurs et programmes informatiques) des processus cognitifs réservés aux capacités du cerveau humain, dans les domaines de :

- la compréhension.

- la communication (dialogue entre machines et avec l'humain).
- l'adaptation.
- l'apprentissage en autonomie (*Deep Learning*).

Dans la langue courante, l'IA désigne également les dispositifs informatiques ou robotiques mettant en œuvre ces fonctions.

❖ Définition 2

On peut aussi définir l'intelligence artificielle comme étant un ensemble d'algorithmes qui traite un ensemble d'informations ou données, relatives à des tâches, d'une manière similaire ou identique à celle qu'un être humain utilisera pour prendre une décision ou résoudre un problème.

De la, on distingue que définir l'intelligence artificielle n'est pas une chose facile. Le champ est si vaste qu'il est impossible de la restreindre à un domaine de recherche spécifique.

Ainsi on peut définir l'intelligence artificielle comme :

- Un domaine de l'informatique dont le but est de recréer un équivalent technologique à l'intelligence humaine.
- Un ensemble de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle.
- L'ensemble des théories et d'algorithmes mise en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine.

Le point commun entre toutes ces définitions se résume dans l'objectif majeur de l'IA qui présente son ambition d'imiter le processus cognitifs de l'être humain. Ces processus comprennent l'apprentissage (acquisition d'informations et de règles liées à leur utilisation), le raisonnement (application des règles pour parvenir à des conclusions approximatives ou précises) et l'autocorrection. [1.1] L'IA vise aussi un peu loin et cela afin de mettre au point des systèmes qui résolvent certains problèmes bien mieux que les humains dans le futur.

1.4) L'utilisation de l'intelligence artificielle:

L'intelligence artificielle s'invite dans notre quotidien depuis quelques années et fait de grands pas. On peut citer quelques usages quotidiens de l'intelligence artificielle:

Les e-mails

L'utilisation de l'intelligence artificielle dans la messagerie électronique améliore l'expérience des utilisateurs, elle aide à filtrer les courriers indésirables, et à traiter de grandes quantités de messages, ainsi la détection des spam qui encombrant les boîtes de réception.

Les réseaux sociaux

L'intelligence artificielle influence largement sur la façon de représenter les informations, notamment sur les réseaux sociaux, par exemple sur Facebook, les publications qui sont importantes sont publiées en premier et celles jugées moins importantes sont cachées, et cela à l'aide du filtrage qui s'appuie sur l'IA. D'autres réseaux sociaux utilisent l'IA pour proposer aux utilisateurs des produits intéressants, tels que faire un achat ou une vente de produit en ligne, découvrir de nouvelles offres dans plusieurs domaines...etc.

Les moteurs de recherches:

Pour pouvoir fonctionner, les moteurs de recherches ont besoin de l'intelligence artificielle, elle est déployée sur le web sous forme de robots qui analysent les différents sites pour les indexer, les classer et déterminer l'ordre de présentation selon les termes de recherche.

Google offre également sa fonction Google Suggest, qui s'affiche pendant la saisie de la requête, qui essaye de deviner la recherche basée sur son indexation, les recherches fréquentes et l'historique de l'utilisateur.

La traduction:

Grâce aux techniques du *Deep Learning* qui s'appuient sur l'IA, le concept de la traduction automatique a beaucoup progressé ces dernières années, au point de produire des textes parfaitement compréhensibles.

Les sites marchands

L'IA occupe une grande place dans la gestion des sites de vente en ligne, elle suggère aux visiteurs de nouveaux articles à consulter et même la possibilité de poster leurs propres produits en ligne afin de les vendre. Ainsi L'IA aide à mieux gérer les marchandises en utilisant des robots qui, acheminent les produits d'un poste à l'autre.

Les applications de navigation

L'intelligence artificielle a changé nos habitudes de circulation, et ça grâce à l'utilisation des applications de navigation comme Waze ou Google Maps qui facilitent les déplacements en déterminant le chemin le plus court, et même pour désigner s'il y'a un trafic en temps réel pour modifier automatiquement la trajectoire et éviter tout encombrement.

1.5) Les domaines d'applications de L'IA

L'intelligence artificielle, définie comme intelligence présentée par les machines, à de nombreuses applications dans la société d'aujourd'hui. Elle a été utilisée pour développer et faire progresser de nombreux domaines et industries, y compris la santé, le secteur financier et bancaire, le transport, le commerce, l'industrie, l'agriculture intelligente, les médias, le droit...etc.

1.5.1 La santé

C'est l'un des secteurs les plus dynamiques en intelligence artificielle. Aujourd'hui un grand nombre de systèmes se basant sur l'intelligence artificielle commence à diagnostiquer des maladies, notamment les cancers, d'une façon aussi bien voire mieux que les spécialistes. L'intelligence artificielle peut favoriser la détection précoce d'une maladie et aussi pourrait accélérer la mise au point de nouveaux médicaments et réduire le temps écoulé entre la découverte d'une nouvelle molécule et sa mise sur le marché [1.2].

1.5.2 Le secteur bancaire et financier

Un des principaux usages de l'IA est:

- la mise en place des agents conversationnels pour répondre aux questions des clients, sur la base de milliers de conversations analysées et enregistrées.
- Aider les conseillers dans leur décision d'accorder ou non un prêt à une personne et les aider à traiter en priorité les mails urgents adressés par les clients (sinistre, perte ou vol de carte).
- La détection des opérations frauduleuses, de plus en plus nombreuses et complexes à débusquer, est facilitée par le recours à l'IA.

1.5.3 Le transport

Les voitures d'aujourd'hui ont des fonctions d'aide à la conduite basées sur l'IA telles que le stationnement automatique. L'intelligence artificielle a été utilisée dans ce domaine afin de pouvoir optimiser les applications de gestion du trafic, ce qui permet de réduire les temps d'attente, la consommation d'énergie et les émissions. Cette utilisation devrait fournir aussi un transport sûr, efficace et fiable tout en essayant de minimiser l'impact sur l'environnement. À l'avenir, des voitures entièrement autonomes sont envisagées à être développées.

1.5.4 Le commerce

L'intelligence artificielle a un impact énorme sur le commerce. Aujourd'hui, les robots effectuent plusieurs tâches d'aide, par exemple la vérification de l'inventaire, la surveillance des planchers qui sont sales et les nettoyer, faire le suivi des étiquettes de prix, le travail à la caisse et beaucoup plus. Toutes ces tâches sont réalisées par des machines afin de faire face aux défis de la main-d'œuvre que les supermarchés rencontrent comme difficultés.

❖ Exemple d'application de l'intelligence artificielle dans le Commerce

Caper, un fournisseur de technologie de vente au détail basé à Brooklyn, New York, a développé un panier d'achat intelligent et autocontrôlé qui utilise la vision par ordinateur, la fusion de capteurs et trois caméras pour mettre automatiquement les articles placés dans le panier.

Lorsqu'un produit est placé dans un panier, le client doit le scanner pour que le panier puisse en apprendre sur le produit. Après cette analyse, les fonctions de vision par ordinateur prennent le relais et l'article peut simplement être placé dans le panier. Caper calcule automatiquement le prix. Lorsqu'un client termine ses achats, il peut vérifier en utilisant les lecteurs de carte de crédit sur le panier d'achat. Il peut utiliser le paiement mobile ou une carte de crédit. Ainsi, le client peut retirer son sac et repartir.



Figure 1.1 : Illustration d'un panier d'achat intelligent.

1.5.5 L'industrie

La robotique moderne s'appuie largement en intelligence artificielle. C'est le cas des robots qui sont capables d'apprendre un geste à partir d'une manipulation réalisée par un opérateur. Les applications industrielles de l'intelligence artificielle sont nombreuses. L'IA permet d'optimiser les performances industrielles tel que le coût de production, de réduire les défaillances et de simplifier les simulations, et enfin elle aide à automatiser le contrôle de qualité pour assurer une cadence de production élevée et une meilleure planification (par exemple la production d'un nombre de pièces supérieur en un temps réduit).

1.5.6 L'agriculture intelligente

Dans le domaine agricole, le déploiement de l'IA se fait de plusieurs manières, D'abord, elle aide les agriculteurs à surveiller leurs champs et ça grâce à l'utilisation des capteurs, ensuite à l'aide des techniques d'apprentissage automatique et le Deep Learning, une analyse très précise est faite sur ces informations captées. Par exemple, dans le cas de la

reconnaissance d'images de plantes, les techniques d'IA aident à détecter certains défauts comme la présence de ravageurs sur les feuilles.

On peut aussi trouver d'autres applications de l'IA dans l'agriculture comme l'augmentation des rendements. Ceux-ci s'appuient sur des paramètres comme le climat, l'état des semences et des sols, les niveaux d'irrigation etc., pour aider les agriculteurs à savoir quoi planter, où et quand, quels plants surveiller et quand récolter.

1.5.7 Les Médias

L'analyse de contenu multimédia audiovisuel tel que des films, des programmes télévisés, des vidéos publicitaires ou les contenus générés par les utilisateurs présente un domaine vers lequel certaines applications d'IA se sont orientées. Les solutions impliquent souvent l'utilisation de la vision par ordinateur, qui représente un domaine d'application majeur de l'IA.

Les scénarios de tels cas d'utilisation comprennent l'analyse d'images à l'aide des techniques de reconnaissance d'objets ou de reconnaissance faciale, ou l'analyse de vidéos pour la reconnaissance de scènes. La motivation pour l'utilisation de l'analyse des médias basée sur l'intelligence artificielle peut être, entre autres:

- la facilité de la recherche multimédia.
- la création d'un ensemble de mots-clés descriptifs pour un élément multimédia.
- la reconnaissance automatique de la parole à des fins d'archivage ou à d'autres fins.

1.5.8 Le droit

L'intelligence artificielle intervient aussi dans le domaine de droit. Elle fournit un ensemble d'algorithmes qui sont susceptibles d'analyser une jurisprudence afin de déterminer les arguments clés, fournir un conseil juridique sur la qualité d'un contrat ou la probabilité de gagner un contentieux et même d'estimer le montant des indemnités à espérer [1.3]

1.6) Les avantages, inconvénients et les limites de l'intelligence artificielle

L'intelligence artificielle présente de nombreux avantages, concernant essentiellement le monde du travail, ses hautes performances et ses perspectives économiques intéressantes. Mais cela n'empêche pas l'existence de certains inconvénients et limites.

1.6.1. Les avantages

L'intelligence artificielle présente plusieurs avantages qui touchent à plusieurs domaines desquels on peut citer:

- L'IA pourrait remplacer l'homme dans son travail quotidien et pourrait réaliser des tâches pénibles ou dangereuses (comme ménage, courses, cuisine, jardinage...) sans présenter aucune contrainte physique (besoin de repos, nourriture ...etc) .
- Établir des calculs plus rapides et efficaces sur un ordinateur, grâce aux algorithmes de l'IA, et limiter les erreurs par rapport aux calculs réalisés par l'être humain.
- L'IA facilite les déplacements, il existe des véhicules pouvant se déplacer seuls à l'aide de caméras et de capteurs répartis sur ceux-ci ce qu'on appelle des véhicules autonomes.
- L'IA présente des avantages dans la médecine, tels que le suivi des patients à distance, les prothèses intelligentes, ainsi que les traitements personnalisés.
- Les machines, ont pas besoins de pauses et de rafraîchissements fréquents par rapport aux humains. Elles sont programmées pour de longues durées et peuvent fonctionner sans arrêt sans être fatiguées, tout ce dont elles ont besoin est une maintenance de temps en temps.
- L'IA présente aussi des avantages dans les jeux, elle permet aux joueurs de combattre des adversaires plus forts et plus expérimentés et donc de ne plus se lasser d'un jeu trop rapidement.

1.6.2. Les inconvénients

Comme toute discipline scientifique, l'intelligence artificielle présente un ensemble d'inconvénients

- L'inconvénient le plus envisageable est la présence d'une erreur dans la programmation d'un robot, ce qui pourrait nuire à son bon fonctionnement. Les machines (ordinateurs, robots, véhicules intelligents...), ne savent pas déceler les erreurs de programmation. Les conséquences d'une telle erreur pourraient être catastrophiques à grande échelle, mais le risque que cela se produise reste très faible.
- L'augmentation de taux de chômage, car les entreprises préfèrent remplacer l'homme par des robots mécanisés dotés de l'IA qui ne se fatiguent pas et qui n'ont besoin que d'une maintenance de temps en temps, ce qui entraîne la suppression de postes.
- Le prix des recherches pour l'IA est très élevé, si l'on veut fabriquer des robots capables d'être autonomes dans la vie quotidienne, cela coûterait extrêmement cher.

En résumé, l'IA ne présente pas beaucoup d'inconvénients face aux nombreux avantages mais il faut savoir rester toujours vigilant.

1.6.3. Les limites de l'IA

Le domaine de l'intelligence artificielle n'est pas infini, il est soumis à des limites. On compte trois limites principales:

- **Les limites matérielles:** Elle fonctionne à l'aide d'un réseau de transistors qui communiquent entre eux (presque aussi vite que la lumière, soit 3.00×10^8 m/s) tandis que le cerveau humain communique moins rapidement, mais contient beaucoup plus de neurones par rapport aux nombres de transistors que possède l'IA.
- **Les limites émotionnelles:** Deuxièmement, l'Homme est capable d'avoir des sentiments, de ressentir des choses, cela est très compliqué à reproduire. Les sentiments sont dus à la nature et non à une programmation.

- **Les limites cognitives:** Les machines sont de plus en plus efficaces pour traiter des données, mais ce traitement comporte des points faibles. En effet, les ordinateurs fonctionnent sous système binaire, c'est-à-dire avec des 1 et des 0, appelés bytes, c'est en quelque sorte leur langage. A la différence du cerveau humain qui réfléchit en sollicitant son réseau, d'une complexité sans pareille et très étendue de neurones [1.4].



1.7) Conclusion:

Dans ce chapitre nous avons parlé de manière générale sur les principes et concepts de base de l'intelligence artificielle.

Nous avons montré le principe de l'intelligence artificielle et cité quelques définitions afin de pouvoir tirer les grandes lignes de cette nouvelle discipline. Nous avons aussi parlé de quelques utilisations de cette dernière et de ses domaines d'application. A la fin nous avons récolté les principaux avantages de l'intelligence artificielle, ainsi que les inconvénients et les limites qu'elle présente.

Pour le chapitre à suivre nous allons nous intéresser aux outils ou algorithmes indispensables à la conception d'un système intelligent.

Chapitre II

**Outils et algorithmes mis en
œuvre pour l' intelligence
artificielle**

2.1) Introduction

L'utilisation des méthodes manuelles pour résoudre certaines tâches comme la reconnaissance d'objet dans les images ou la reconnaissance vocale, qui ont une apparence simple est avérée très difficile à appliquer. Cela est dû à la complexité et la variabilité des données venant du monde réel, comme les échantillons d'un son ou les pixels d'une image.

La machine voit une image comme étant un tableau de nombres qui représente la luminosité de chaque pixel, et un signal sonore comme étant une succession de nombres qui indique la pression de l'air à chaque instant. Grâce à l'intelligence artificielle et plus précisément l'apprentissage machine, que l'on appelle aussi apprentissage automatique, la machine est devenue capable de transcrire la suite de nombres d'un signal sonore en série de mots et d'identifier un chat dans un tableau de nombre d'une image lorsque son apparence et l'apparence des objets qui l'entourent sont variés.

Pour ce deuxième chapitre nous allons approfondir la notion d'IA en citant les outils et les algorithmes mis en œuvre pour l'IA. Premièrement nous aborderons les techniques de l'apprentissage, bien précisément l'apprentissage machine (l'apprentissage automatique ou machine Learning), nous allons citer ses différents types utilisés lors de la conception d'un système doté d'une intelligence artificielle et parler plus en détails de l'apprentissage profond (*Deep Learning*), ou nous allons exposer la notion des réseaux de neurones artificiels et ses différents types et leurs principe de fonctionnement afin de mieux comprendre la notion de l'apprentissage profond. Ensuite nous allons présenter le principe de fonctionnement de *Deep Learning*, ses domaines d'applications pour pouvoir conclure le chapitre avec des exemples d'illustrations. [2.1]

2.2) Techniques de l'apprentissage

Les scientifiques et afin de créer une intelligence artificielle, utilisent un ensemble de techniques différentes regroupées sous une même discipline qu'on appelle l'apprentissage machine (ou *machine Learning*, en anglais). Donc dire intelligence artificielle revient automatiquement à dire "*Machine Learning*", ou l'apprentissage automatique.

2.2.1) L'apprentissage automatique

L'apprentissage automatique est considéré comme un sous-domaine de l'intelligence artificielle. Son objectif majeur est de comprendre la structure des données et de les faire intégrer dans des modèles simples que tout le monde peut comprendre et utiliser.

Bien que l'apprentissage automatique soit un domaine de l'informatique, il diffère de ces approches traditionnelles. Les algorithmes dans les approches informatiques sont des ensembles d'instructions programmées explicitement pour calculer et résoudre des problèmes bien définis, tandis que les algorithmes d'apprentissage automatique donnent aux ordinateurs la possibilité de s'entraîner sur un ensemble de données qui lui sont fournies comme entrées et utilisent l'analyse statistique pour produire des valeurs situées dans une plage spécifique comme sorties. Pour cela, l'apprentissage automatique facilite l'utilisation des ordinateurs dans la construction de modèles à partir de données d'échantillonnage afin d'automatiser les processus de prise de décision en fonction des données saisies.

De nos jours, l'apprentissage automatique est omniprésent dans toutes nos interactions avec les banques, nos achats en ligne et notre utilisation des médias sociaux. Les algorithmes d'apprentissage automatique entrent en jeu dans ses cas pour garantir la fluidité, l'efficacité et la sécurité de nos expériences. [2.2]

Les technologies les plus récentes sont toutes dotées de l'apprentissage automatique et leurs utilisateurs bénéficient des avantages de cette nouvelle discipline. Par exemple, la technologie de reconnaissance faciale permet aux plateformes des réseaux sociaux d'aider les utilisateurs à marquer et partager des photos d'amis. La technologie de reconnaissance optique des caractères est capable de convertir les images du texte en caractères mobiles. Les moteurs de recommandation, qui sont alimentés par l'apprentissage automatique, suggèrent à l'utilisateur les films ou émissions de télévision à regarder en fonction des ses préférences.

2.2.2) Types d'apprentissage

Généralement dans l'apprentissage machine, les tâches à résoudre sont classées en catégories. Ces catégories sont définies selon la façon dont l'apprentissage est reçu ou comment le retour d'informations est donné au système développé.

Il existe plusieurs méthodes d'apprentissage automatique, mais les plus largement utilisées sont l'apprentissage supervisé ou les algorithmes sont basés sur des données d'entrées et de sorties bien spécifiées par l'homme et l'apprentissage non supervisé qui ne fournit pas à l'algorithme des données étiquetées afin de lui pouvoir permettre de définir une structure et de découvrir une logique dans les données entrées. Auxquelles on ajoute l'apprentissage par renforcement et l'apprentissage profond. Explorons donc ces méthodes plus en détail.

❖ L'apprentissage supervisé

L'apprentissage supervisé est une variété de machine Learning qui utilise un ensemble de données d'apprentissage étiquetées afin de créer des modèles d'intelligence artificielle. Le but de cette méthode c'est le fait que l'algorithme soit capable «d'apprendre» en comparant sa sortie réelle avec les sorties «enseignées» pour trouver des erreurs et modifier le modèle en conséquence. Donc l'apprentissage supervisé utilise des modèles pour déterminer les valeurs d'étiquettes pour un ensemble de données non étiquetées. Pour pouvoir illustrer ce type d'apprentissage, on peut se baser sur l'exemple d'un enfant qui apprend à identifier des animaux en l'apprenant à partir d'un livre d'images. Dans le cadre de l'apprentissage supervisé, l'algorithme est entraîné par un ensemble de données qui est déjà étiqueté et qui a une sortie prédéterminée.

Il existe plusieurs algorithmes qui se basent sur la méthode d'apprentissage supervisée, dont on peut citer la classification multi classe et les machines à vecteurs de support.

❖ L'apprentissage non supervisé

Dans le cas de l'apprentissage non supervisé, on utilise un ensemble de données d'entrées non étiquetées, afin de laisser l'algorithme d'apprentissage trouver tout seul les points communs parmi cet ensemble de données. Les méthodes d'apprentissage automatique qui facilitent l'apprentissage non supervisé sont particulièrement utiles, vu que les données non étiquetées étant plus considérables que celles étiquetées.

On peut considérer que l'objectif initial de l'apprentissage non supervisé est aussi simple que de détecter les modèles cachés dans un ensemble de données, mais il peut aussi avoir un objectif d'apprentissage des caractéristiques, ce qui va rendre la machine intelligente capable de découvrir automatiquement les représentations nécessaires pour classer des données brutes.

Pour rester sur l'exemple précédent d'apprentissage des animaux pour un enfant, avec la méthode non supervisée, cet enfant apprend à identifier ces animaux en observant les couleurs, les dessins et les caractéristiques, plutôt que de mémoriser les noms avec l'aide d'un livre ou d'un enseignant. L'enfant recherchera les points communs entre les images et sera capable de les séparer en groupes, en attribuant à chaque groupe sa propre nouvelle étiquette.

❖ L'apprentissage par renforcement

L'apprentissage par renforcement fait référence à une classe de problèmes d'apprentissage automatique, qui consiste à apprendre à partir d'expériences successives, ce qu'il convient de faire de façon à trouver les meilleures solutions. Autrement dit, les machines intelligentes essaient plusieurs situations afin de pouvoir déterminer les actions les plus avantageuses, et ne se contentent pas de recevoir des instructions sur les actions à appliquer, ce qui distingue cette méthode des autres techniques d'apprentissage.

L'apprentissage par renforcement est un modèle d'apprentissage comportemental. L'algorithme dans ce cas reçoit les informations en analysant des données, pour pouvoir orienter l'utilisateur vers les meilleurs résultats. Dans ce type d'apprentissage, le système n'est pas entraîné à partir d'un ensemble de données mais il apprend par essais et erreurs, ce qui le diffère des autres types d'apprentissage supervisé.

L'apprentissage par renforcement est utilisé par exemple pour les voitures autonomes. L'entraînement d'une voiture autonome est un processus extrêmement complexe à cause des nombreux obstacles possibles. Si toutes les voitures étaient autonomes, les essais et les erreurs seraient plus faciles à surmonter, mais dans le monde réel, les facteurs humains sont souvent imprévisibles [2.3].

❖ L'apprentissage profond:

Le Deep Learning, ou apprentissage profond, est l'une des principales technologies de Machine Learning et d'intelligence artificielle. Découvrez en quoi consiste cette technologie, son fonctionnement, et ses différents secteurs d'application dans les sous titres suivants ou tout sera détaillé.

2.3) Apprentissage Profond (Deep Learning)

2.3.1) Définition et origines

Le Deep Learning ou apprentissage profond est un type d'intelligence artificielle, dérivé du machine Learning qui a été développé dans le but de créer des algorithmes capables d'apprendre et de s'améliorer de manière autonome, contrairement à la programmation où la machine se contente d'exécuter à la lettre des règles prédéterminées.

L'apprentissage en profondeur utilise une succession de couches d'unités de traitement non linéaire pour pouvoir extraire ou transformer les caractéristiques des données. La sortie d'une couche sert d'entrée de la couche suivante. Les algorithmes de l'apprentissage profond peuvent être supervisés et servir à classer les données, ou non supervisés et aider à effectuer une analyse de modèle. L'algorithme de *Deep Learning* absorbe des quantités de données énormes par rapport aux autres algorithmes d'apprentissages machine utilisés et développés actuellement, et il a été capable de battre les humains dans certaines tâches cognitives. Par exemple, la reconnaissance faciale par ordinateur et la reconnaissance vocale ont connu des progrès significatifs et cela grâce aux approches d'apprentissage approfondies.

Parmi ces approches d'apprentissage approfondies, on compte les réseaux de neurones artificiels sur lesquels reposent le Deep Learning et ainsi certaines technologies comme la reconnaissance d'image ou la vision robotique. Les réseaux de neurones artificiels sont inspirés par les neurones du cerveau humain. Ils sont constitués à base de plusieurs neurones artificiels connectés entre eux. Plus le nombre de neurones est considérable, plus le réseau est profond.

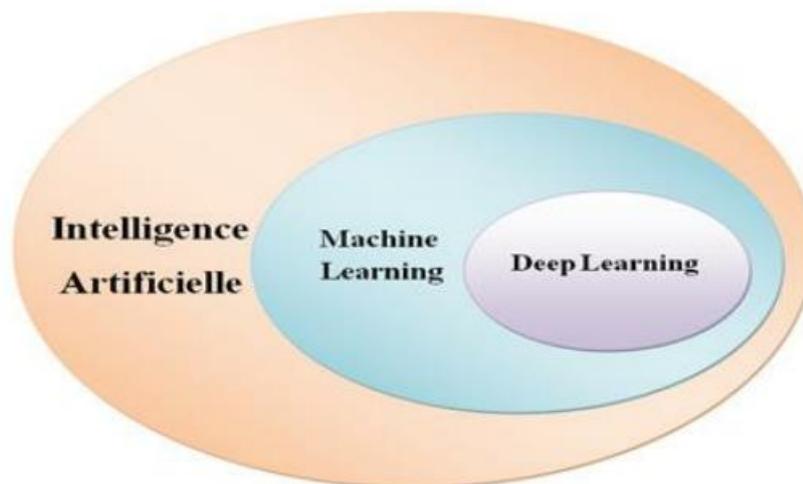


Figure 2.1: La relation entre l'IA, Machine Learning et Deep Learning

2.3.2) Les réseaux de neurones artificiels

2.3.2.1) Définition

Un réseau de neurones artificiels, ou (*Artificial Neural Network* en anglais), est un système informatique matériel et / ou logiciel s'inspirant du fonctionnement du cerveau humain pour apprendre. Il s'agit d'une variété de technologie Deep Learning, qui fait elle-même partie de la sous-catégorie d'intelligence artificielle et du Machine Learning. Ce genre de réseau est défini par un ensemble de couches de neurones qui sont fortement interconnectées entre elles.

2.3.2.2) Architecture d'un réseau de neurones artificiel

En général un réseau de neurones est constitué d'un ensemble de couches successives dont chacune prend ses entrées sur les sorties de la précédente c.-à-d. que cet ensemble est entièrement connectée. Chaque couche est un ensemble de neurones n'ayant pas de connexion entre eux et qui reçoivent des informations numériques en provenance de neurones voisins. L'ensemble de couches est composé d'une couche d'entrée qui lit les valeurs d'entrées, une couche de sortie qui fournit les résultats du système et entre ces deux se cache une à plusieurs couches dites cachées qui participent au transfert. Comme le montre la figure 2.2 suivante.

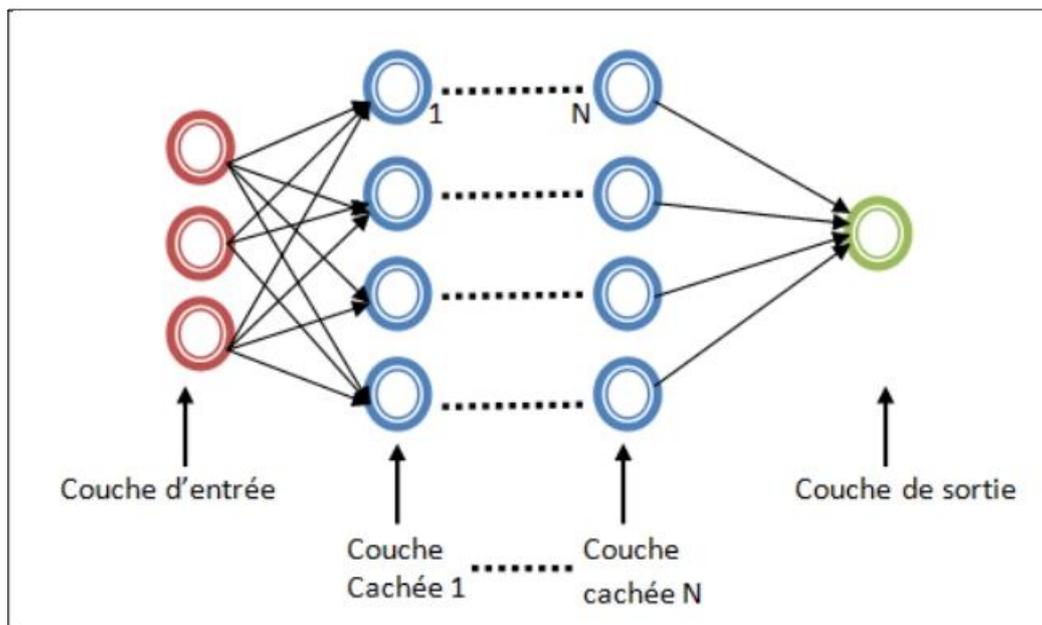


Figure 2.2 : Architecture de base d'un réseau de neurones artificiel

2.3.2.3) Principe de fonctionnement d'un réseau de neurones artificiel

Définissons maintenant le principe de fonctionnement d'un réseau de neurone en représentant les grandes étapes réalisées lors l'utilisation de ce réseau

- On attribue à chaque neurone constituant le réseau un coefficient qu'on appelle **poids**.
- Chaque donnée d'entrée (input) va être multipliée par ce coefficient et on va appliquer une fonction d'addition à l'ensemble des résultats. Cette fonction est appelée fonction de combinaison.
- Puis on va appliquer une fonction d'activation, en comparant la somme obtenue a un seuil déterminé. Si la somme est en dessous du seuil (souvent dans le cas ou la somme est négative), le neurone ne s'active pas, ce qui signifie que la donnée n'est pas intéressante. Si cette somme est au-dessus du seuil (souvent dans le cas ou la somme est positive), alors le neurone s'active et envoie l'information au neurone de la couche suivante (couches cachées), jusqu'à ce que la donnée ultime atteigne le dernier neurone. On aura alors un résultat final (output).

Ce principe est illustre dans la figure 2.3 suivante.

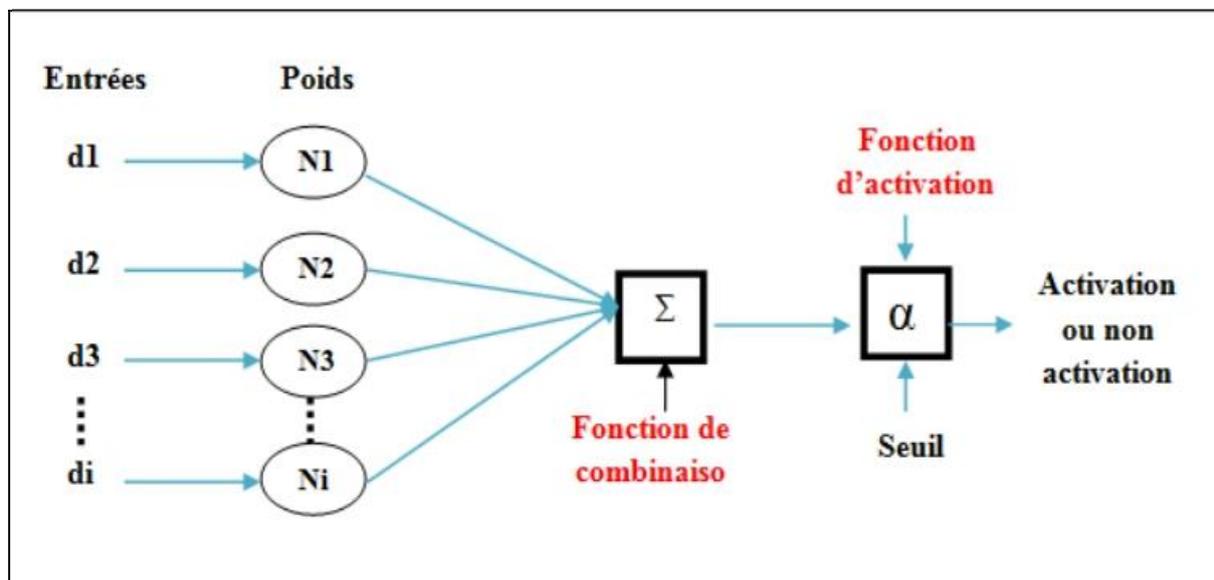


Figure 2.3: Principe de fonctionnement d'un réseau de neurones artificiel

2.3.2.4) Types des réseaux de neurones

Il existe différents types de réseaux de neurones, et ils sont classés en fonction du nombre de nœuds cachés du modèle ou encore du nombre d'entrées et de sorties de chaque nœud.

La propagation des informations entre les différents neurones peut varier et dépend du type de réseaux de neurones.

- **Les réseaux de neurones dit "feed-forward" (à propagation avant):** c'est la variante la plus simple, l'information ne se déplace que dans une seule direction, elle traverse directement l'entrée aux nœuds de traitement (couches cachées) puis aux sorties, avec absence de cycle ou de boucle dans le réseau.
- **Les réseaux de neurones récurrents:** Ce mode d'apprentissage est un peu plus complexes, ils comportent au moins un cycle dans leurs structures, ils sauvegardent les résultats produits par les nœuds de traitement et nourrissent le modèle à l'aide de ces résultats. Parmi ses applications, on trouve: la reconnaissance automatique de formes, la traduction automatique de la parole...etc.
- **Les réseaux de neurones convolutifs :** leur fonctionnement est inspiré par un processus biologique qui est le cortex visuel des animaux, les données à traiter

traversent une multitude de couches (de la couche d'entrée vers la couche de sortie). Ce type de réseaux s'applique dans différents domaines: la reconnaissance faciale, la reconnaissance d'images et vidéos, le traitement du langage naturel ...etc. (Ce type de réseau de neurones est bien détaillé dans le **chapitre 3**) [2.4].

2.3.3) Domaines d'applications du Deep Learning

Le Deep Learning touche un grand nombre de domaines d'applications, dont on peut citer :

- ***La reconnaissance faciale***

Un algorithme de Deep Learning apprend à détecter sur une photo les caractéristiques du visage tels que: les yeux, la bouche, le nez...etc. Et cela en fournissant à l'algorithme un ensemble d'images du visage qui vont être utilisées pour l'entraînement, et à force de les entraîner, il sera capable de détecter un visage sur une image.

- ***La détection d'objets***

Les algorithmes de détection d'objets sont capables maintenant d'identifier au pixel près un élément ou une personne sur une image qui contient beaucoup d'éléments (images complexes).

- ***La Traduction automatique***

La traduction automatique a fait d'immenses progrès grâce à l'utilisation des réseaux de neurones. Cependant, ce type d'IA a besoin de gros volumes de contenus traduits au préalable par des humains. Elle s'appuie sur l'apprentissage supervisé par lequel la machine fait une supposition puis reçoit la réponse correcte d'un humain, ce qui lui permet ensuite d'ajuster son traitement en conséquence. Cette méthode est efficace pour les langues très répandues, comme l'anglais ou le français, pour lesquelles il existe de nombreux documents parallèles.

- ***La reconnaissance d'images***

La reconnaissance d'images en fonction de ce qu'elles représentent (par exemple, des voitures, des êtres humains) est une tâche difficile pour un ordinateur, il doit d'abord effectuer une analyse afin d'extraire les caractéristiques de chaque image. Avec le *Deep Learning* l'information d'entrée (ici une image) est analysée couche après couche, le résultat de l'analyse d'une couche est transmis à la couche suivante. Les informations ainsi collectées sont cartographiées de façon à constituer un algorithme flexible, et donc l'ordinateur est capable, à l'aide de plusieurs opérations de déterminer si une image appartient à la catégorie voiture ou bien être humain. [2.5]

2.3.4) Exemples du Deep Learning

- ❖ ***Deep Learning: Exemple 1***

On va étudier un exemple concret de la reconnaissance d'images, à l'aide des réseaux de neurones, et qui consiste à reconnaître les photos d'un chat afin d'illustrer cette notion.

Tout d'abord, le réseau de neurones prend en entrée un ensemble d'images de chats différents mélangés avec des objets qui ne sont pas des chats, puis il va effectuer une analyse de chaque image afin de s'entraîner pour pratiquer le Deep Learning. Ces images sont ensuite converties en données et transférées sur le réseau. Ensuite les neurones artificiels attribuent un poids aux différents éléments. Enfin la couche finale de neurones va rassembler les différentes informations pour déduire le résultat s'il s'agit d'une image d'un chat ou non.

Le réseau de neurones va ensuite effectuer une comparaison entre cette réponse et les bonnes réponses proposées par les humains. Si les réponses sont pareilles, le réseau garde cette réussite en mémoire et s'en servira plus tard pour reconnaître les chats. Dans le cas contraire, le réseau ajuste le poids placé sur les différents neurones afin de corriger son erreur.

Ce processus représenté dans la **Figure 3.2** est répété des milliers de fois jusqu'à ce que le réseau arrive à reconnaître un chat sur une photo dans toutes les circonstances. [2.6]

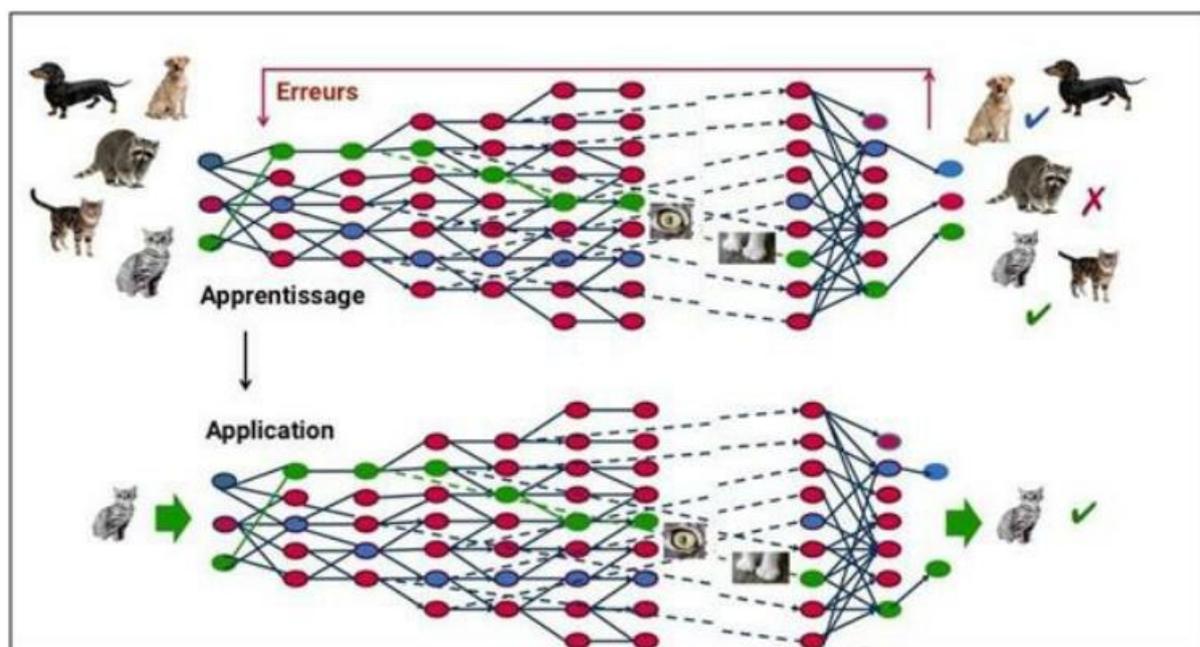


Figure 2.4: le processus d'auto-apprentissage pour l'identification d'un chat [2.7].

❖ Deep Learning: Exemple 2

Dans le Deep Learning, l'apprentissage se fait à travers une multitude de couches, ce qui permet à l'ordinateur de former une hiérarchie de concepts complexes à partir des concepts plus simples.

Considérons une image d'une personne. Les ordinateurs ne peuvent pas comprendre le sens d'une collection de pixels. Le mappage d'une collection de pixels vers un objet est

complicé. Avec l'utilisation du Deep Learning, le problème est divisé en une série de correspondances hiérarchiques. Chaque mappage est représenté par une couche spécifique.

Tout d'abord La couche d'entrée reçoit l'image à traiter. Ensuite, une extraction de caractéristiques est faite au niveau d'une série de couches dites cachées. Par exemple, à partir de pixels, la première couche cachée identifie les bords de l'image, après la deuxième couche, identifie à partir de ces bords les points et les contours, et à partir de ces derniers la troisième couche identifie des parties qui constituent l'objet. Enfin à partir des parties d'objets identifiés, la quatrième couche arrive à reconnaître des objets entiers [2.8]. Ce processus est illustré dans la **Figure 2.3**.

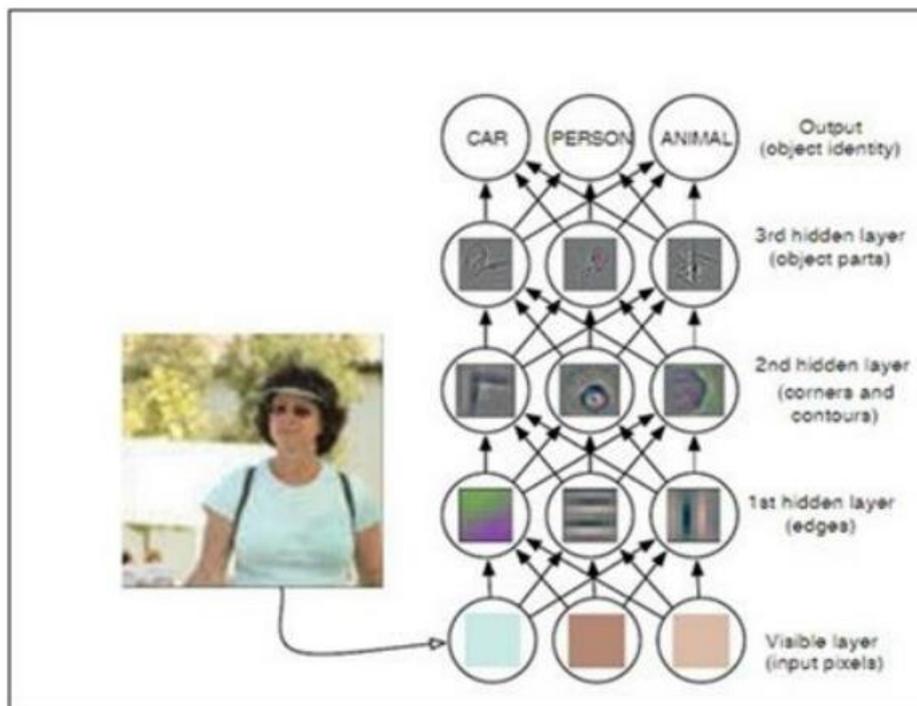


Figure 2.5 : le processus d'auto-apprentissage pour l'identification d'une personne.

3.4) Conclusion:

Nous nous sommes intéressés en premier lieu dans ce chapitre à la représentation de la *Machine Learning* comme concept général et ses différents types. Ensuite nous avons approfondi l'un de ses types qui est l'apprentissage profond (Deep learning) et à ce niveau nous avons abordé la notion de réseaux de neurones, sur laquelle se base le Deep learning et ses différents types. Nous avons enchaîné avec le principe de fonctionnement de ce type et ses différents domaines d'applications. Et vers la fin nous avons cité deux exemples d'illustrations pour mieux comprendre.

CHAPITRE III

La classification d'images avec le Deep learning.

3.1) Introduction:

De nos jours, le monde dans lequel nous vivons est un monde numérique, où un énorme volume d'informations est stocké, traité, indexé et recherché par des systèmes informatiques, ce qui rend la récupération des informations une tâche rapide et pas cher. Le domaine de la classification d'images a subi un progrès considérable, cela est dû aux nombreux travaux effectués et à la disponibilité des bases d'images internationales. Ces bases d'images ont permis aux chercheurs de signaler de manière crédible l'exécution de leurs approches dans ce domaine, auquel on ajoute la possibilité de les comparer à d'autres approches qu'elles utilisent les mêmes bases.

La classification automatique d'images représente une application de reconnaissance de formes qui permet d'attribuer automatiquement une classe à une image à l'aide d'un système de classification. Comme on peut trouver aussi, la classification d'objets, de scènes, de textures, la reconnaissance de visages, d'empreintes digitales...etc. Il existe deux principaux types d'apprentissage dépendant des informations disponibles sur les données à classer : l'apprentissage supervisé et l'apprentissage non supervisé.[3.1] Dans notre travail, nous allons nous intéresser aux méthodes de classification supervisées dont les classes sont définies au préalable.

Pour ce troisième chapitre consacré à la classification des images avec Deep Learning, nous aborderons en premier lieu les notions de bases définissant une image. Ensuite, nous allons présenter les deux types de méthodes qui permettent de les classer (approches supervisée et non supervisée), ainsi que les indicateurs de performance utilisés pour évaluer les modèles. Puis, le rôle de la machine learning et du *Deep Learning* dans la classification sera également souligné. Enfin, nous allons décrire et exposer les réseaux de neurones à convolution avec certains modèles.

3.2) Les motivations de la classification des images

La classification des images peut être définie comme étant la répartition systématique des images selon des classes connues au préalable. Donc classer une image c'est le fait de lui faire correspondre une classe, marquant ainsi sa parenté avec d'autres images.

Généralement, pour un humain la tâche de reconnaissance d'image est quelque chose d'aisé, vu qu'au fil de son existence il a acquis des connaissances qui lui permettent de s'adapter aux changements qui résultent des conditions différentes d'acquisition. Par exemple, la reconnaissance d'un objet dans ces différentes orientations, même si il est partiellement caché par un autre objet, est quelque chose de très facile pour un être humain.

Toutefois le développement des technologies d'acquisition d'images (microscopes, caméras, capteurs) et de stockage a engendré des quantités énormes d'informations et a multiplié les domaines d'application. C'est pour cette raison, que l'humain est devenu incapable d'analyser ce nombre important d'images, ainsi que le temps requis, le caractère répétitif de la tâche et la concentration nécessaire sont devenus problématiques. Cette tâche n'est pas forcément aisée aussi pour un système informatique qui considère une image comme étant un ensemble de valeurs numériques.

On peut déduire que l'objectif principal de la classification d'images est de construire un système capable d'attribuer une classe pour une image de façon automatique. Ainsi, ce système permet de réaliser une tâche d'expertise qui peut se montrer très coûteuse à acquérir pour un être humain. Cette difficulté d'acquisition est due aux contraintes physiques comme la fatigue, la concentration ou le temps important qu'un grand volume de données nécessite pour être analysé. [3.1]

3.3) Domaines d'applications de la classification d'images :

Les applications de la classification automatique d'images sont nombreuses et touchent à plusieurs domaines:

- Domaine médical : la reconnaissance de cellules et de tumeurs.
- Domaine urbain : la reconnaissance de panneaux de signalisation, la reconnaissance de piétons, la détection de véhicules, la reconnaissance de bâtiments.

- Domaine de la biométrie : la reconnaissance de visage, d'empreintes.
- Etc.

L'ensemble de ces applications possède un point commun, qui se résume à la nécessité de la mise en place d'une chaîne de traitement composée de plusieurs étapes, à partir des images disponibles, et cela afin de pouvoir fournir un décision en sortie. Pour chaque étape de la mise en place d'un tel système de classification il est nécessaire de chercher des méthodes appropriées pour une performance globale optimale, à savoir la phase d'extraction de caractéristiques et la phase d'apprentissage.

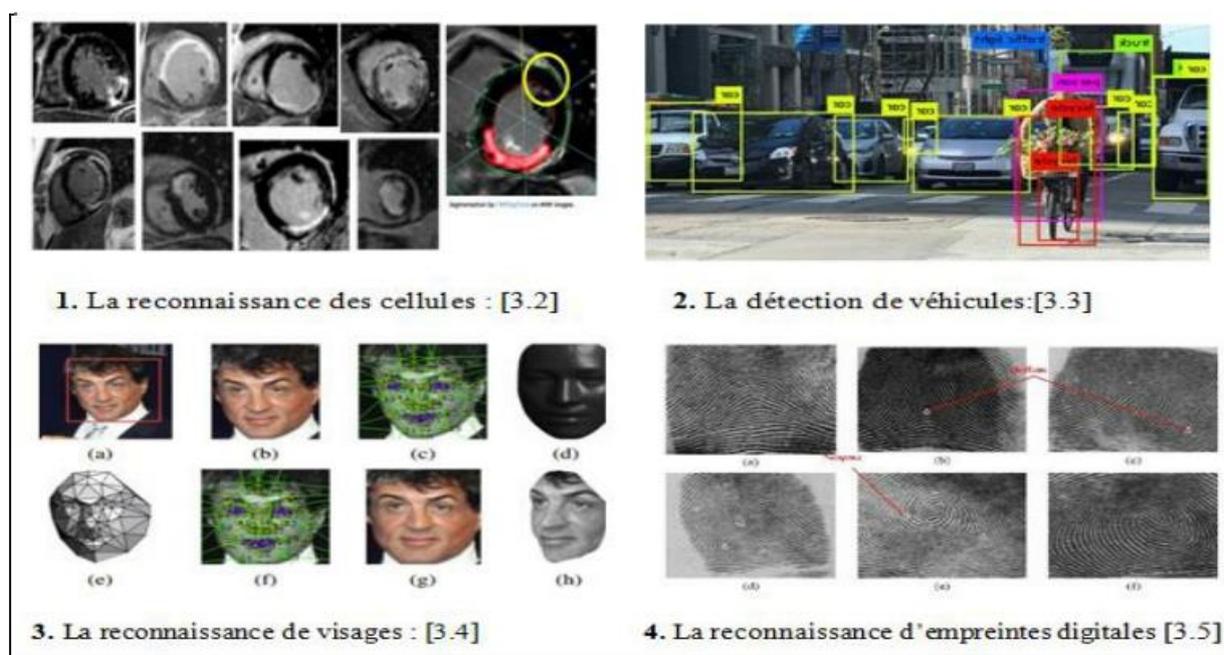


Figure 3.1 : Les Domaines d'application de la classification d'images

3.4) Notions de bases:

3.4.1) Définition d'une image:

Une image peut être définie mathématiquement comme une fonction qui associe à un couple (x, y) de coordonnées spatiales une valeur d'intensité lumineuse, c'est une reproduction analogique exacte d'une scène réelle. Elle est considérée comme un ensemble

de points auxquels on associe une grandeur physique telle que la luminance, la couleur, la dimension...etc.

3.4.2) Les différents types de format d'image:

- **Image couleur RVB :**

L'œil humain analyse les couleurs via des cellules photo réceptrices, et perçoit toute couleur comme une combinaison des 3 couleurs de base: rouge, vert, bleu, alors pour représenter la couleur d'un pixel, il est nécessaire de donner trois nombres qui correspondent au dosage des trois couleurs de base.

- **Image d'intensités ou image en niveaux de gris :**

C'est une matrice qui possède un ensemble de valeurs réelles comprises entre 0 et 1, qui servent à représenter les différents niveaux de gris. Le noir est désigné en 0 et le blanc en 1.

3.4.3) Caractéristiques de l'image:

L'image est un ensemble structuré d'information caractérisé par les paramètres suivants :

3.4.3.1) Le pixel:

Le pixel est une unité élémentaire qui constitue la base d'une image numérique. Il concrétise un point (x, y) du plan de l'image. Un pixel représente le niveau de gris ou de couleur prélevé de l'image réelle. La différence entre image monochrome et image couleur réside dans la quantité d'informations contenues dans chaque pixel.

3.4.3.2) Dimension et Résolution:

La dimension c'est la taille de l'image et elle correspond au nombre total de pixel dans celle ci. Comme on a déjà défini l'image comme étant une matrice, donc sa dimension peut être calculée comme suit:

Dimension d'une image = nombres de lignes * nombres de colonnes

Par contre, la résolution est la clarté ou la finesse de détails atteinte par un moniteur ou une imprimante dans la production d'images. On utilise aussi le mot résolution pour désigner le nombre de pixels par pouce (1 pouce = 2.54 centimètres). Elle est exprimée en "PPP" (points par pouce). Plus il y a de pixels (ou points) par pouce et plus il y aura d'information dans l'image donc plus précise.

3.4.3.3) Voisinage:

Le voisinage d'un pixel est représenté par l'ensemble des pixels qui entourent ce même pixel. Cet ensemble de voisinage définit la notion de l'assiette. La division du plan de l'image en forme rectangulaire ou hexagonale permet l'exploitation de la notion de voisinage.

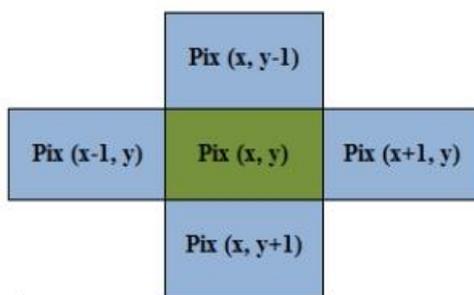


Figure 3.2: Voisinage à 4

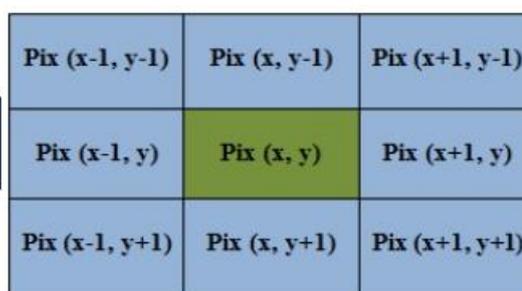


Figure 3.3: Voisinage à 8

3.4.3.4) Niveau de gris:

Elle indique la valeur d'intensité lumineuse d'un pixel qui varie de 0 qui représente le noir jusqu'à 255 qui est le blanc en passant par les nuances qui sont contenues dans l'intervalle [0,255], elle correspond en fait à la quantité de lumière réfléchi.

3.4.3.5) Luminance:

C'est une grandeur photométrique qui dépend de la sensation visuelle et donc de la sensibilité de l'œil, Elle représente le degré de luminosité des points de l'image, elle peut être défini aussi par le mot brillance qui correspond à l'éclat d'un objet. Elle est calculée ainsi:

$$\text{Luminance} = \frac{\text{L'intensité lumineuse d'une surface}}{\text{L'aire apparente de cette surface}}$$

3.4.3.6) Bruit:

Un bruit est un défaut parasite qui subit des dégradations dans la qualité de l'image, il représente la variation de l'intensité d'un pixel par rapport à ses voisins. Parmi ces défauts on a: la poussière, petits nuages, faible intensité électrique.....etc.

3.4.3.7) Contour:

Dans une image numérique, le contour désigne la frontière entre chaque objet de l'image, et la détection de ce contour a pour but de repérer les points de l'image numérique qui correspondent à un changement brutal de l'intensité lumineuse, et ça cause une élimination d'information et une réduction de la qualité de l'image.

3.5) Méthodes de classification:

Afin de pouvoir réaliser la classification d'images, plusieurs méthodes classiques ont été consacrées. Ces méthodes sont classées en deux grandes catégories :

1. Les méthodes de classification supervisées
2. Les méthodes de classification non supervisées.

3.5.1) Méthodes supervisées:

La classification d'images avec la méthode supervisée a comme objectif principal la définition d'un ensemble de règles, permettant de classer des objets dans des classes, à partir de certaines variables qui caractérisent ces objets. Pour cette méthode il faut posséder au départ un ensemble d'échantillons dit aussi d'apprentissage dont les classes sont connues. C'est-à-dire que chaque image possède une étiquette bien spécifique qui décrit sa classe d'appartenance. Cet échantillon est utilisé pour l'apprentissage des règles de classement comme première étape de classification, puis un deuxième échantillon indépendant, dit de validation ou de test est utilisé.

3.5.2) Méthodes non supervisées:

Ces méthodes procèdent d'une façon contraire que celles supervisées. C'est-à-dire ne nécessitent aucun apprentissage et aucune tâche préalable d'étiquetage manuel, autrement dit, les données sont non étiquetées. Alors, c'est au système d'extraire une règle d'appartenance de chaque image. Elles permettent de former à partir d'un nuage de points de n'importe quel espace un ensemble de groupes appelés Clusters.

3.6) Indicateurs de performance en classification

3.6.1 Matrice de confusion

La matrice de confusion appelée aussi matrice d'erreur, permet de mesurer la qualité du système de classification utilisé. Son rôle principal c'est d'effectuer plus rapidement une analyse approfondie des statistiques et aussi faciliter la lecture des résultats grâce à une visualisation claire des données. Elle est appliquée dans plusieurs applications de l'intelligence artificielle comme: les statistiques, le data mining, les modèles d'apprentissage automatique... Un des intérêts de la matrice de confusion est qu'elle montre rapidement si un système de classification parvient à classifier correctement ou non.

On peut aussi la définir d'une autre manière, comme étant un outil permettant de mesurer la concordance entre un ensemble d'éléments observés et un ensemble d'éléments de référence. Ici les éléments observés correspondent aux pixels issus de la classification et les éléments de référence à nos échantillons de vérification.[3.6].

Si on considère la matrice de confusion d'un modèle d'apprentissage automatique, on va aligner les prévisions sur la droite et en haut les données réelles. Les prévisions ou ce qu'on appelle aussi les résultats sont présentées dans les lignes, juste en dessous des données réelles. Les résultats de cette dernière peuvent être interprétés de la façon suivante:

- L'indication correcte d'une valeur positive comme étant un vrai positif (TP=True Positif).
- L'indication correcte d'une valeur négative comme étant un vrai négatif.(TN=True Négatif).
- L'indication correcte d'une valeur positive incorrecte comme étant un faux positif (FP=False Positif).
- L'indication correcte d'une valeur négative incorrecte comme étant un faux négatif (FN=False Négatif).

La figure suivante illustre un exemple de matrice de confusion, obtenu par la classification d'un ensemble d'images contenant (28 images), Divisé en deux classes C1 et C2 avec 14 images chacune. Cela en utilisant le modèle prédéfinis googleNet entraîné avec la méthode *cross validation*.

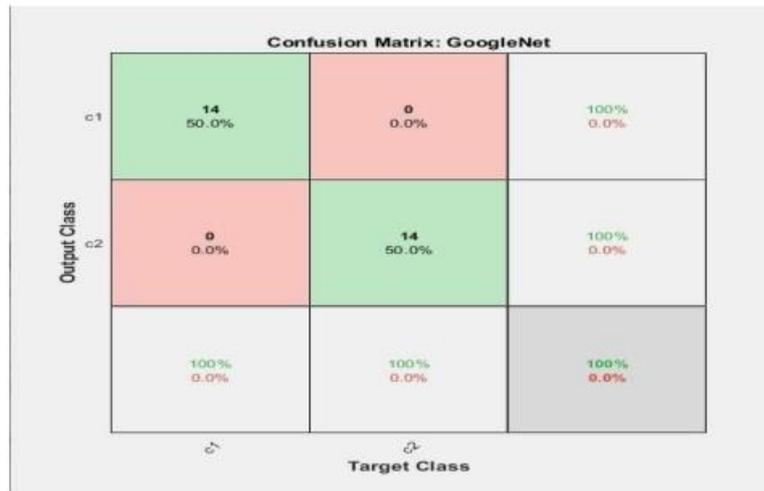


Figure 3.4: Exemple de Matrice de Confusion.

3.6.2 La courbe ROC (*Received Operating Characteristic*):

La courbe ROC consiste à un graphique qui représente l'ensemble de performances d'un modèle de classification pour tous les seuils de la classification. Cette courbe trace le taux de vrais positifs en fonction du taux de faux positifs.

Avec cet outil on peut avoir toutes les informations sur la performance du classifieur, et cela en utilisant l'aire sous la courbe. Plus elle se rapproche de 1, plus le classifieur est performant.

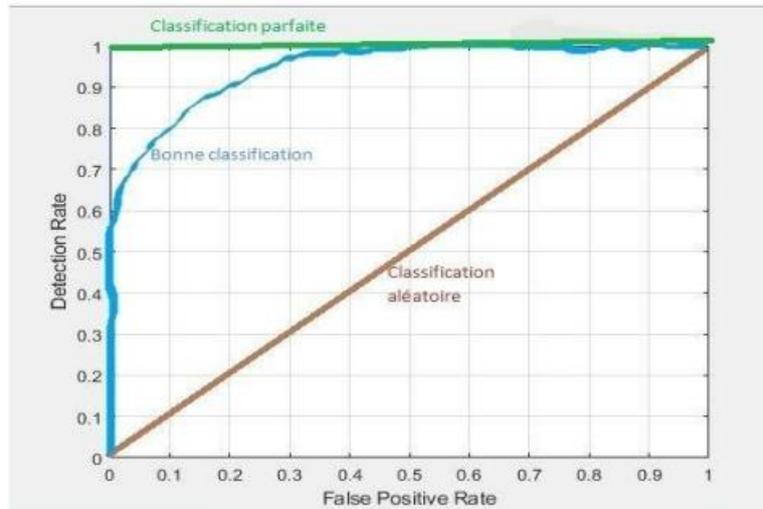


Figure 3.5: Exemple d'illustration de la courbe ROC.

3.6.3 La précision (Precision): La précision de la classification est le nombre total de prédictions correctes divisé par le nombre total de prédictions effectuées pour un ensemble de données.

$$\text{Precision} = \frac{TP}{TP + FP}$$

3.6.4 Le rappel (Recall): Proportion d'éléments bien classés par rapport au nombre d'éléments de la classe à prédire.

$$\text{Rappel} = \frac{TP}{TP + FN}$$

3.6.5 La justesse (Accuracy): il indique le pourcentage de bonnes prédictions.

$$\text{Accuracy} = \frac{TP + TN}{N} \quad \text{avec } N = TP + TN + FP + FN$$

3.6.6 La spécificité (Specificity): elle mesure la capacité d'un test à donner un résultat négatif lorsque l'hypothèse n'est pas vérifiée.

$$\text{Spécificité} = \frac{TN}{TN+FP}$$

3.6.7 F-Score: elle calcule la moyenne harmonique entre précision et rappel.

$$\text{F-Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

3.7) Classification des images et l'apprentissage machine:

La classification d'images avec les méthodes manuelles est une tâche très difficile à appliquer, d'où l'intervention du machine learning qui sert comme un système qui dirige tous les systèmes des grandes entreprises d'internet, dans le but d'effectuer différentes fonctionnalités telles que: le filtrage des contenus indésirables, l'affichage des informations concernant un utilisateur et hiérarchiser des réponses à une recherche.

Un système de classification ou de reconnaissance des formes est un système entraînable qui possède des entrées et des sorties, par exemple, il prend en entrée une image, un son ou un texte et une sortie qui peut définir la classe de l'objet dans l'image, le mot prononcé ou le thème du texte. L'apprentissage machine utilise une méthode supervisée, cela signifie qu'on lui fournit en entrée par exemple une photo d'un chat et on lui donne la sortie désirée pour un chat, et après plusieurs exemples d'apprentissage, la machine sera capable de fixer ses paramètres afin de rapprocher sa sortie de la sortie désirée. Donc après avoir traité des milliers d'exemples d'apprentissage, la machine devient capable de classifier correctement les données, avec la possibilité de généralisation, par exemple classifier des images de chats qu'elle a jamais analysées durant la phase d'apprentissage.

Les systèmes de reconnaissance des images classiques contiennent deux bloc: un extracteur de caractéristiques qui est programmé manuellement et qui sert à transformer le tableau de nombres représentant l'image en une série de nombres appelé vecteur de caractéristiques, suivi d'un classifieur entraînable simple qui prend le vecteur comme entrée

afin de calculer une somme pondérée des caractéristiques ou chaque nombre est multiplié par un poids (positif ou négatif) avant d'être sommé.

3.8) Classification des images et les réseaux de neurones:

L'approche classique de la reconnaissance des images malgré tous ces atouts, rencontre certains problèmes vu qu'un bon extracteur de caractéristiques est très difficile à construire. Ce dernier doit aussi être révisé pour chaque nouvelle application. Pour pallier ces problèmes une nouvelle approche est apparue, qui est l'apprentissage profond.

L'idée de cette approche est très simple, le système entraîné est constitué d'une série de modules, ou chacun désigne une étape de traitement. Chaque module est entraînable, et comporte un ensemble de paramètres ajustables semblable aux poids de classifieur linéaires. Le système est entraîné de bout en bout, de sorte à ajuster les paramètres de tous les modèles pour rapprocher la sortie produite par le système de la sortie désirée. La notion de profondeur vient de l'arrangement de ces modules en couches successives.

L'avantage des architectures profondes, c'est leur capacité d'apprendre à représenter le monde de manière hiérarchique. Dans ce cas, la construction d'un extracteur de caractéristiques à la main, n'est pas nécessaire, vu que toutes les couches sont entraînaibles. Un autre avantage est que chaque ensemble de couche possède un rôle bien spécifique, les premières couches sont utilisées pour extraire les caractéristiques simples (exemple présence de contours). Ces caractéristiques sont ensuite combinées par les couches suivantes pour former des concepts de plus en plus complexes et abstraits: assemblages de contours en motifs, de motifs en parties d'objets, de parties d'objets en objets.

3.9) Les réseaux de neurones convolutifs (CNN : Convolutional neural networks)

Comme on a vu dans le chapitre 2 l'apprentissage en profondeur est l'une des techniques d'apprentissage automatique les plus largement utilisées qui a remporté un énorme succès dans des applications telles que la détection d'anomalies, la reconnaissance d'images, la reconnaissance de formes et le traitement du langage naturel. Les architectures de Deep

learning se varient selon ces applications, on trouve alors trois catégories principales de réseaux de neurones, à savoir,

- Les réseaux de neurones convolutifs (la reconnaissance d'images).
- Les réseaux non surveillés pré-entraînés (la détection d'anomalies).
- Les réseaux de neurones récurrents / récurrents (le traitement du langage naturel, la synthèse vocale et la traduction automatique.).

Les progrès de l'IA et de l'apprentissage profond ont permis une évolution rapide dans les domaines de la vision par ordinateur et de l'analyse d'images. Tout cela est rendu possible par l'émergence et les progrès des réseaux de neurones convolutionnels.

Dans les titres à suivre on va élargir la notion de CNN, avec une définition, son architecture et principe de fonctionnement et les différents modèles prédéfinis (LeNet, AlexNet, VGG, GoogleNet et ResNet).

3.9.1 Définition d'un réseau neuronal convolutif

Un réseau neuronal convolutif est un algorithme d'apprentissage en profondeur qui peut reconnaître et classer les caractéristiques des images pour la vision par ordinateur. Il s'agit d'un réseau neuronal multicouche conçu pour analyser les entrées visuelles et effectuer des tâches telles que la classification d'images, la segmentation et la détection d'objets.[3.7]

3.9.2 Architecture et fonctionnement d'un CNN

L'architecture de CNN est inspirée par l'organisation et la fonctionnalité du cortex visuel et conçue pour imiter le modèle de connectivité des neurones dans le cerveau humain.

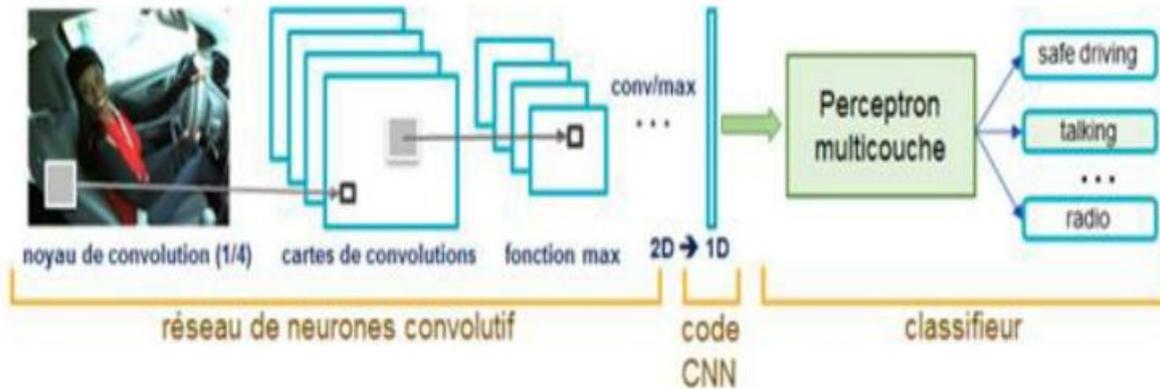


Figure 3.6: Architecture de base d'un réseau neuronal convolutif [3.8]

Les réseaux de neurones convolutifs sont à ce jour les modèles les plus performants pour classer des images. Ils comportent deux parties bien distinctes: la partie convolutive et la partie classification.

3.9.2.1 Partie convolution

Son rôle est d'extraire les caractéristiques des images. Une succession de filtres est appliquée à l'image d'entrée, pour produire de nouvelles images en sortie appelées cartes de convolutions. Certains filtres utilisent l'opération de maximum local (Pooling) pour redimensionner l'image en réduisant sa dimension, tout en préservant ses caractéristiques importantes. Finalement les cartes de convolutions sont chaînées dans un vecteur de caractéristiques appelé code CNN.

Dans la partie convolution on distingue 3 types de couches :

- Les couches de convolutions
- Les couches de correction *ReLU*
- Les couches de *Pooling*

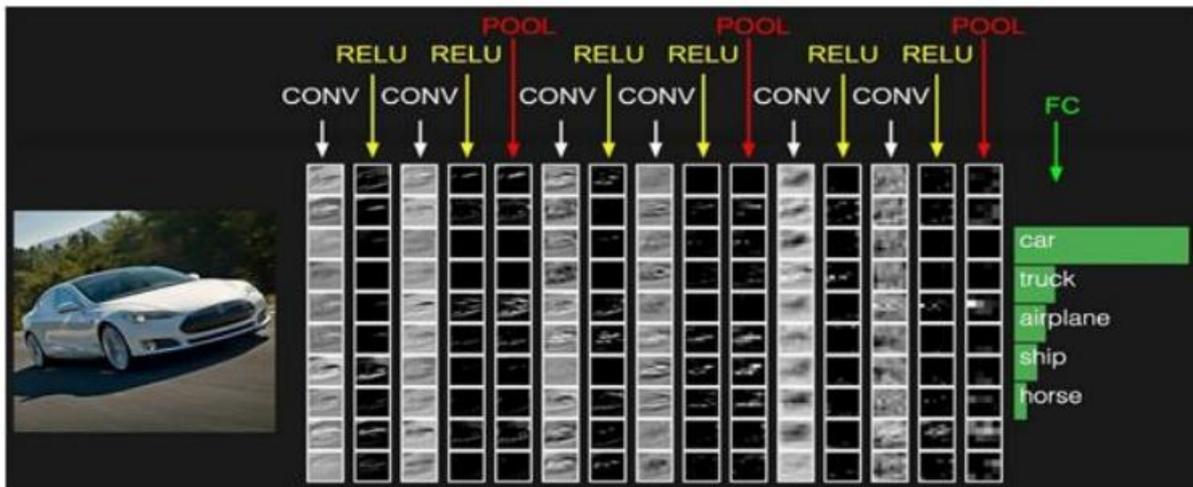


Figure 3.7: La partie convolution d'un CNN [3.8]

- Couche de convolution

- ✓ ·Notion de convolution

La convolution est une opération mathématique simple généralement utilisée pour le traitement et la reconnaissance d'images. Sur une image, son effet s'assimile à un filtrage dont voici le fonctionnement [3.9]

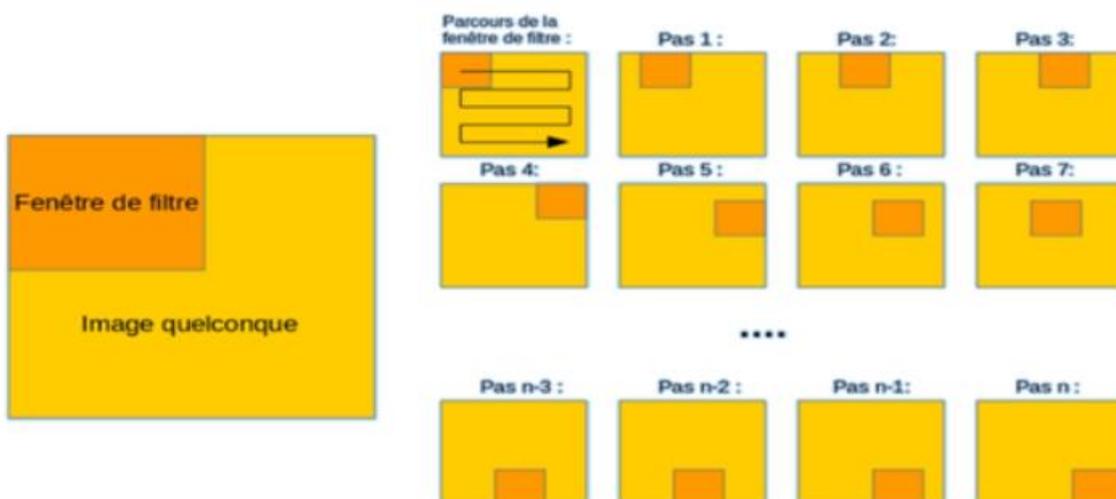


Figure 3.8 : Parcours de la fenêtre de filtre sur l'image [3.9]

1. Dans un premier temps, on définit la taille de la fenêtre de filtre située en haut à gauche.
2. La fenêtre de filtre, représentant la feature, se déplace progressivement de la gauche vers la droite d'un certain nombre de cases défini au préalable (le pas) jusqu'à arriver au bout de l'image.
3. À chaque portion d'image rencontrée, un calcul de convolution s'effectue permettant d'obtenir en sortie une carte d'activation ou *featuremap* qui indique où sont localisées les *features* (les caractéristiques) dans l'image : plus la *featuremap* est élevée, plus la portion de l'image balayée ressemble à la feature.

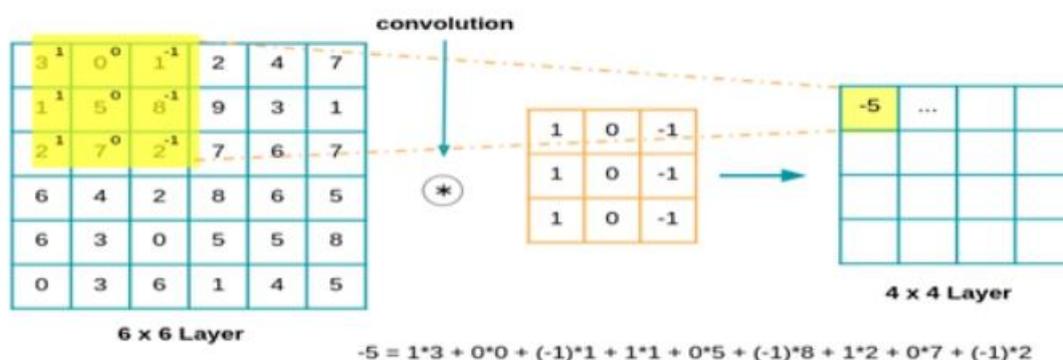


Figure 3.9: Le principe de convolution [3.10]

La couche de convolution reçoit donc en entrée plusieurs images, et calcule la convolution de chacune d'entre elles avec chaque filtre. Les filtres correspondent exactement aux caractéristiques que l'on souhaite retrouver dans les images.

Pour chaque paire (image, filtre), une carte d'activation ou *featuremap*, lui est associée et qui sert à indiquer où se trouvent les features dans l'image : plus la valeur est élevée, plus l'endroit correspondant dans l'image ressemble à la *feature*.

- Couches de correction (*RELU*)

Il est possible d'améliorer l'efficacité du traitement en insérant entre les couches de traitement une couche qui va appliquer une fonction mathématique (fonction d'activation) sur les signaux de sortie [3.1]

On fait passer les cartes de convolutions à travers une couche d'activation non linéaire telle que *Rectified Linear Unit* (ReLU), qui consiste à remplacer les nombres négatifs des images filtrées par des zéros.[3.11]



Figure 3.10: Fonctionnement de la fonction *RELU* [3.12].

- Couche *Pooling* (ou sous-échantillonnage)

Ce type de couche est souvent placé entre deux couches de convolution, elle consiste à un ré-échantillonnage des données; elle prend plusieurs *featuremap* en entrée et applique à chacune d'entre elle la fonction de *Pooling*, qui permet de réduire la taille des images en conservant leurs caractéristiques importantes.

Pour cela, l'image est découpée en plusieurs cellules régulières. Puis dans chaque cellule on garde la valeur maximale. Généralement, on utilise des cellules carrées, de petites tailles pour ne pas perdre beaucoup d'informations.

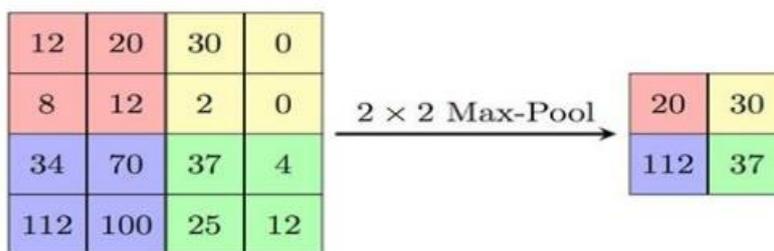


Figure 3.11: Le principe de Pooling avec un filtre 2*2 et un pas de 2 [3.13]

A la sortie de cette couche, on obtient en le même nombre de *feature maps* qu'en entrée, mais cette fois de taille réduite.

La couche de *Pooling* est utilisée pour minimiser le nombre de calculs et de paramètres dans le réseau. Afin de pouvoir contrôler *l'overfitting* (sur-apprentissage) et ainsi améliorer l'efficacité du réseau, une couche de Pooling est fréquemment insérée périodiquement entre deux couches de convolution successives.

A la fin du réseau, pour pouvoir indiquer la probabilité qu'une entité spécifique appartient à une certaine classe, un MLP (Multi Layer perceptron) souvent appelé *Fully-Connected* est ajouté.

3.9.2.2 Partie classification

Elle est constituée de couches entièrement connectées (perceptron multicouche). Elle prend la sortie de la partie convolutive qui est le code CNN comme entrée, afin de combiner ses caractéristiques pour classer l'image.

- **Couche *Fully-connected* (entièrement connecté)**

Elle constitue toujours la dernière couche d'un réseau de neurones convolutif ou non. Elle reçoit un vecteur en entrée contenant les pixels aplatis de toutes les images filtrées, corrigées et réduites par le Pooling et produit un nouveau vecteur en sortie. Pour cela, elle applique une combinaison linéaire puis éventuellement une fonction d'activation aux valeurs reçues en entrée.

La dernière couche *Fully-connected* permet de classifier l'image en entrée du réseau : elle renvoie un vecteur de taille N , où N est le nombre de classes dans notre problème de classification d'images. Chaque élément du vecteur indique la probabilité pour l'image en entrée d'appartenir à une classe. [3.12].

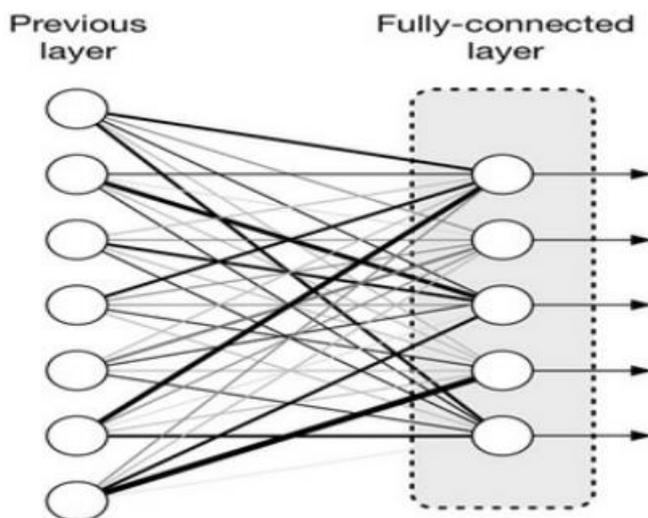


Figure 3.12: La couche *Fullyconnected* [3.14]

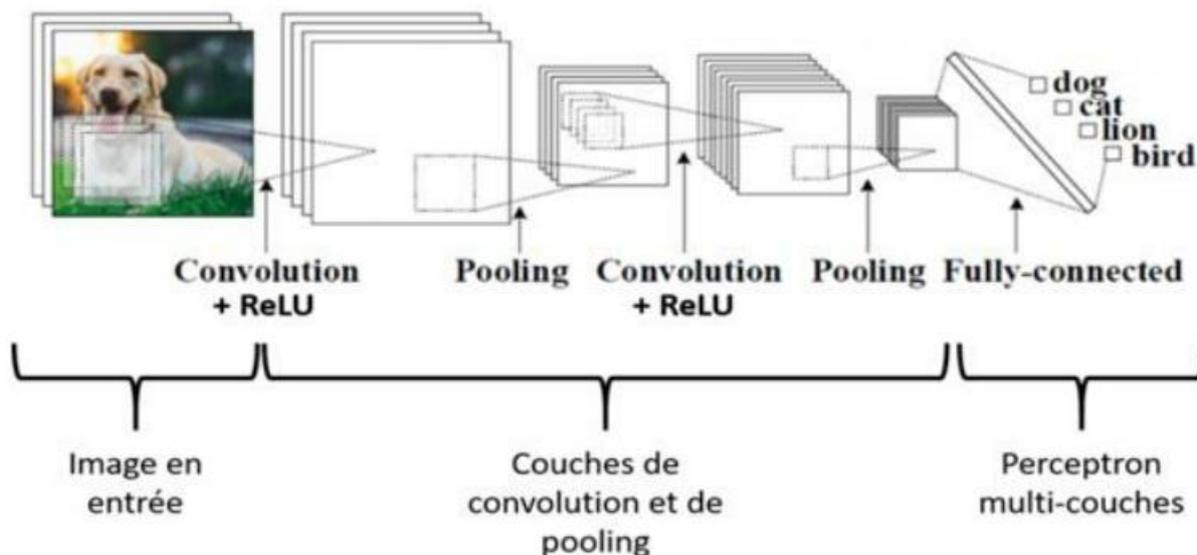


Figure 3.13: Le fonctionnement d'un réseau de neurones à convolution [3.14]

3.9.3. Modèles CNN prédéfinis

Comme on l'a mentionné précédemment, un réseau neuronal convolutif est un type spécial de réseaux neuronaux multicouches, conçu pour reconnaître des modèles visuels directement à partir d'images pixel avec un prétraitement minimal. -

Un concours logiciel annuel ILSVRC (ImageNet Large Scale Visual Recognition Challenge) est organisé par le projet ImageNet où les programmes logiciels sont en concurrence dans le but de classer et détecter avec une grande précision les objets et les scènes. Ce projet est basé sur une grande base de données d'images ImageNet, conçue pour être utilisée dans la recherche de logiciels permettant la reconnaissance d'objets visuels. Cette section est consacrée aux différentes architectures CNN des principaux concurrents de l'ILSVRC.

✓ *LeNet*

LeNet est proposé par l'inventeur Yann LeCun en 1998. Ce réseau était dédié à la reconnaissance des photos noir et blanc et de chiffres. Il est composé uniquement de quelques couches et peu de filtres, en raison des limitations de l'ordinateur à cette époque. Il s'agit d'un petit réseau qui contient les modules de base du Deep Learning.

L'architecture CNN de LeNet est composée de sept couches dont trois sont des couches convolutives, deux sont des couches de sous-échantillonnage et deux sont des couches entièrement connectées.

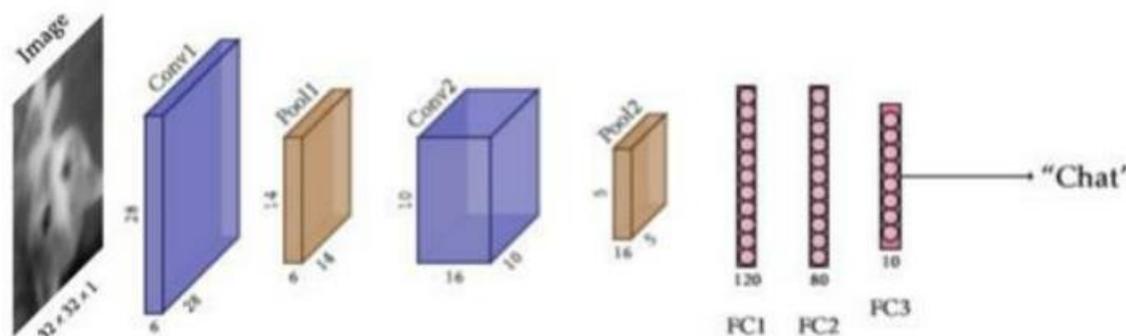


Figure 3.14: Visualisation de l'Architecture du réseau LeNet-5 [3.15]

✓ *AlexNet*

AlexNet a été retenu dans ILSVRC 2012. Il résout le problème de la classification des images où l'entrée est une image de l'une des 1000 classes différentes (par exemple, chats, chiens, etc.) et la sortie est un vecteur de 1000 éléments. [3.16]

AlexNet contient 5 couches convolutives et 3 couches entièrement connectées. La couche Relu est appliquée après les couches de convolution et entièrement connectée. La couche dropout est appliquée avant les première et deuxième couches entièrement connectées.

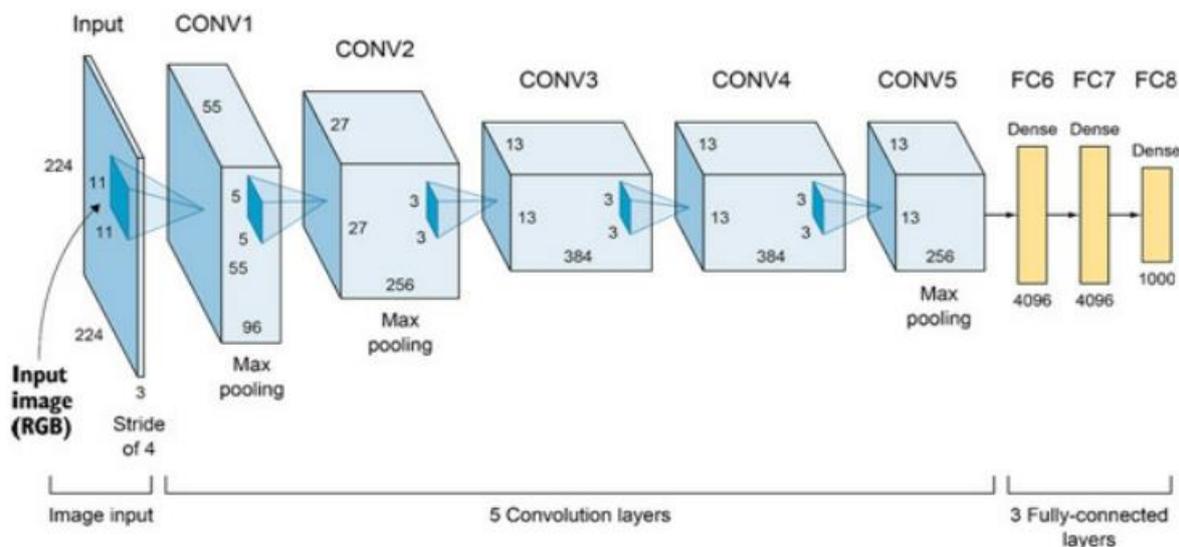


Figure 3.15: Visualisation de l'architecture du réseau AlexNet [3.17]

✓ *VGG16*

VGG est une architecture de réseau de neurones populaire qui est également basée sur les CNN, a été appliquée au Défi ImageNet en 2014. Le réseau a atteint une précision de test de 92,7% dans le top 5 sur l'ensemble de données ImageNet .

Les principales améliorations de VGG, par rapport à AlexNet, incluent l'utilisation de grands filtres de la taille d'un noyau. VGGNet se compose de 16 couches convolutives.

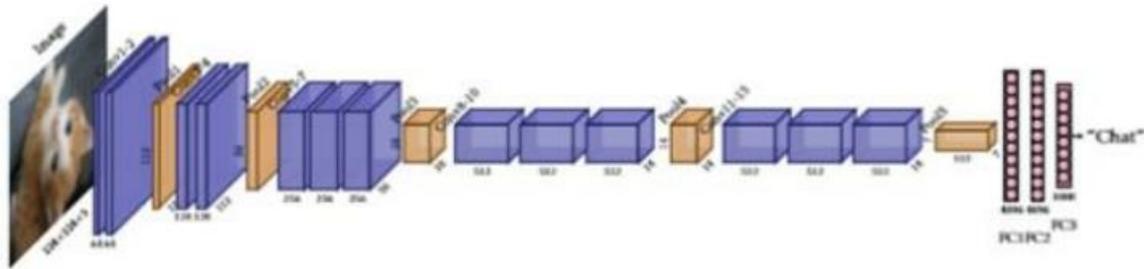


Figure 3.16: Visualisation de l'architecture du réseau VGG16.[3.15]

✓ *GoogleNet*

GoogleNet est le lauréat de l'ILSVRC 2014, avec un taux d'erreur de 6,67% dans le top 5. Le réseau utilise un réseau CNN inspiré par LeNet. La principale caractéristique de cette architecture est la meilleure utilisation des ressources informatiques à l'intérieur du réseau. Ceci a été réalisé grâce à une conception soigneusement conçue qui permet d'augmenter la profondeur et la largeur du réseau tout en maintenant le budget de calcul constant. [3.18]

L'architecture de GoogleNet consistait en un réseau CNN profond de 22 couches mais réduisait le nombre de paramètres de 60 millions (AlexNet) à 4 millions.

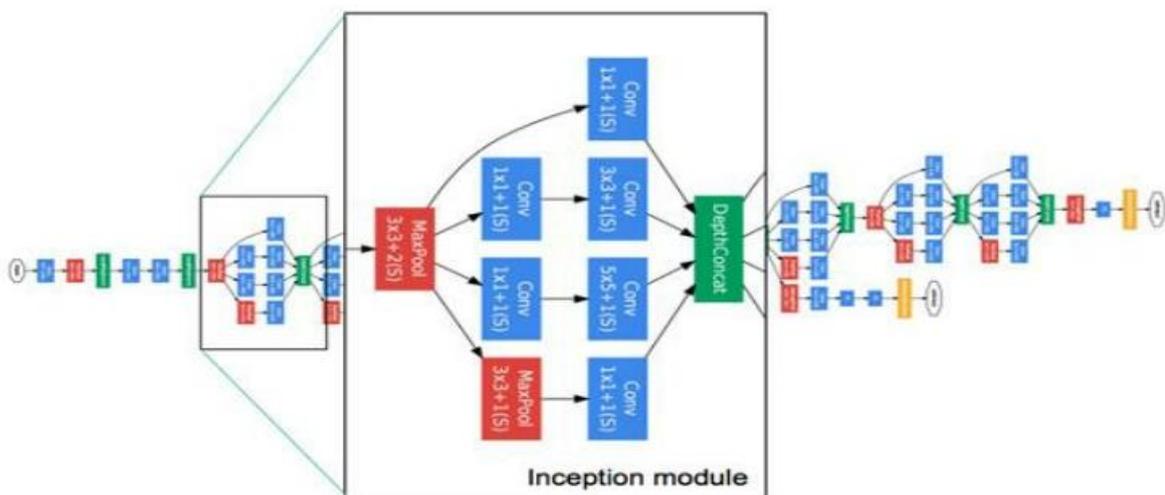


Figure 3.17: Visualisation de l'architecture du réseau GoogleNet (1). [3.19]

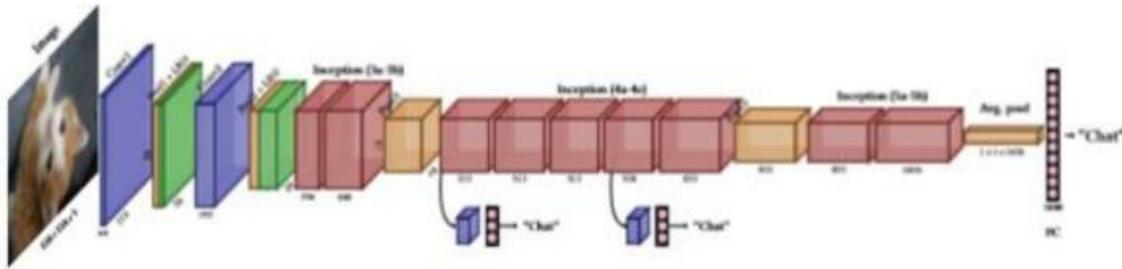


Figure 3.18: Visualisation de l'architecture du réseau GoogleNet (2). [3.15]

✓ ResNet

C'est l'un des réseaux de neurones profonds les plus puissants qui a obtenu des résultats de performances fantastiques dans le défi de classification *ILSVRC 2015*. ResNet a réalisé d'excellentes performances de généralisation sur d'autres tâches de reconnaissance et a remporté la première place sur la détection ImageNet, la localisation ImageNet, la détection COCO et la segmentation COCO dans les concours *ILSVRC* et *COCO 2015*.

Il existe de nombreuses variantes de l'architecture ResNet, c'est-à-dire le même concept mais avec un nombre de couches différent. Nous avons *ResNet-18*, *ResNet-34*, *ResNet-50*, *ResNet-101*, *ResNet-110*, *ResNet-152*, *ResNet-164*, *ResNet-1202*.

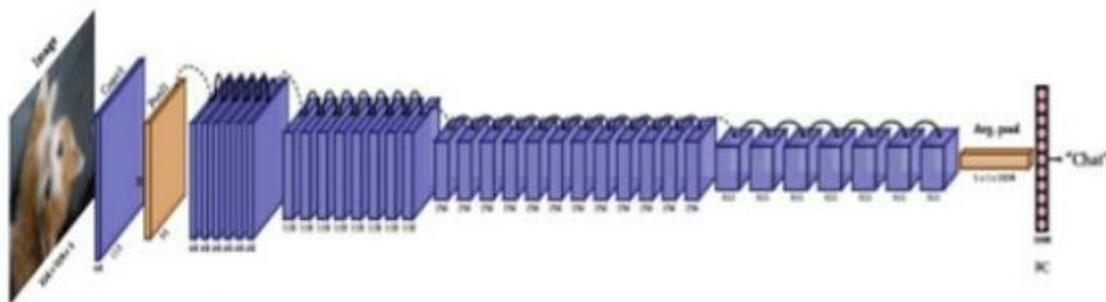


Figure 3.19: Visualisation de l'architecture du réseau ResNet. [3.15]

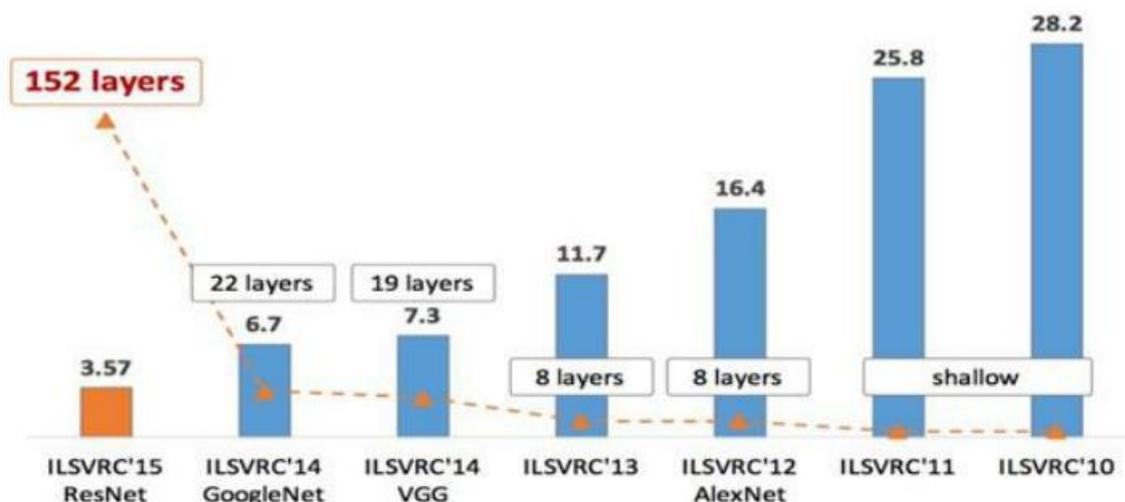


Figure 3.20 : Classification des gagnants du Défi de reconnaissance visuelle à grande échelle ImageNet (ILSVRC). [3.20]

3.9) Conclusion

Nous avons consacré ce chapitre à la présentation des notions de classification d'images et de ses différentes caractéristiques, et domaines d'application, ainsi qu'à l'utilisation de la machine Learning et du Deep Learning dans ce domaine. Nous avons également abordé la notion des réseaux de neurones convolutifs, plus précisément, nous avons donné sa définition, son architecture et son principe de fonctionnement où nous avons cité les différentes couches qui constituent ce réseau. Enfin, nous avons présenté les différents modèles prédéfinis du CNN, tout en détaillant chaque modèle (LeNet, AlexNet, VGG, GoogleNet et ResNet).

Chapitre IV

Implémentation

4.1) Introduction

L'identification des ravageurs est l'une des techniques les plus efficaces pour leur gestion. En déterminant le type de ceux-ci dans une ferme, les épidémies d'insectes et les maladies des cultures peuvent être évitées. Comme nous l'avons vu dans les chapitres précédents, les réseaux de neurones à convolution profonde sont devenus les méthodes les plus largement utilisées pour la classification d'images.

Dans notre travail, nous souhaitons pouvoir classer les images d'insectes nuisibles, selon des classes déjà connues afin de faciliter leurs identifications. Par conséquent, ce chapitre présentera notre solution proposée au problème de l'identification des ravageurs basée sur des réseaux de neurones convolutifs profonds. Pour cela, nous présenterons tout d'abord les outils logiciels et matériels avec lesquels nous avons implémenté notre solution. Puis nous allons décrire l'ensemble de données que nous allons traiter au cours de notre travail. Nous définissons par la suite les architectures et les paramètres des modèles CNNs utilisés dans l'implémentation et citer les différents indicateurs de performances que nous allons utiliser. Pour pouvoir enfin arriver à la partie résultats et leurs discussions, où nous allons présenter les résultats de la classification avec les différents modèles utilisés et dans les différents cas (avec et sans augmentation de données).

4.2) Les Outils logiciels utilisés.

Afin de réaliser notre implémentation et de pouvoir classer l'ensemble des images d'insectes, nous avons utilisé un outil logiciel qui est le logiciel Matlab.

4.2.1) Définition de Matlab

MATLAB est à la fois un langage de programmation et un environnement pour faire du calcul numérique, il permet de faire plusieurs tâches telles que: des manipulations sur des matrices, le traçage des fonctions, la visualisation de données, le traitement d'images, l'implémentation de nouveaux algorithmes, et faire des interfaces graphiques...etc.

4.2.2) *Environnement de développement :*

Pour pouvoir implémenter les différentes fonctions et fichiers de Matlab, on utilise un ensemble d'outils qui constituent l'environnement de développement du logiciel.

Les principaux outils sont :

- La fenêtre de commande (Command Window)
- Un historique des commandes (Command History)
- Un éditeur et un débogueur
- Des navigateurs pour visualiser l'aide, l'espace de travail, des fichiers, et le chemin de recherche.

4.2.3) *Boîte à outils Matlab (Toolbox) :*

MATLAB contient une multitude de boîtes à outils (toolbox) spécifiques à des domaines variés. Elles constituent un ensemble de fonctions spécifiques à un domaine d'application.

On peut citer:

- Résolution d'équations aux dérivées partielles
- Optimisation
- Traitement du signal
- Traitement de l'image
- Classification d'images
- Logique floue
- Réseaux de neurones
- ...etc.

Dans notre cas on a utilisé la boîte à outils spécifique au domaine de classification d'images, qui offre différents packages qui permettent l'implémentation des modèles d'apprentissage Alexnet et GoogleNet sous Matlab.

4.2.4) Matlab et le Deep learning

On peut utiliser Matlab pour appliquer les techniques d'apprentissage en profondeur pour créer, modifier, associer une étiquette à une donnée et analyser des architectures du Deep Learning à l'aide d'applications et d'outils de visualisation.

4.3) Configuration matérielle utilisée.

La configuration matérielle utilisée pour l'entraînement et l'évaluation des modèles CNN est la suivante:

Un Pc HP avec les caractéristiques suivantes:

- processeur intel Core i7-4700MQ CPU @ 2.40GHz.
- Carte graphique intel HD 4600
- RAM d'une taille de 8 GO
- Disque dur (HDD) d'une taille de 600 GO
- Système d'exploitation Windows 10 professionnel 64 bit

Un Pc Acer avec les caractéristiques suivantes:

- Processeur intel Core i5-4200U 1.6GHz with Turbo Boost up to 2.6GHz
- Carte graphique intel HD 4400
- RAM d'une taille de 4 GO
- Disque dur (HDD) d'une taille de 1000 GO
- Système d'exploitation Windows 10 professionnel 64 bit

4.4) Description de l'ensemble de données

L'ensemble de données utilisé dans le cadre de ce mémoire est collecté à partir d'un site web sur les ravageurs [4.1]. Cet ensemble contient six classes de ravageurs qui sont :

Classe 1: Papillon (Lépidoptères).

Classe 2: Mouche (Diptères).

Classe 3: Abeille (Hyménoptères).

Classe 4: Sauterelle (Orthoptères).

Classe 5: Fourmis blanches (Isoptères).

Classe 6: Cafard (Blattodea).



Figure 4.1 : Classe 1: Papillon (Lépidoptères). [4.2]



Figure 4.2 : Classe 2: Mouche (Diptères).[4.3]



Figure 4.3 : Classe 3: Abeille (Hyménoptères). [4.4]**Figure 4.4 : Classe 4: Sauterelle (Orthoptères). [4.5]****Figure 4.5 : Classe 5: Fourmis blanche (Isoptères). [4.6]****Figure 4.6 : Classe 6: Cafard (Blattodea). [4.7]**

L'ensemble de données contient 3528 images dont:

731 images pour la classe 1: Papillon.

525 images pour la classe 2: Mouche.

618 images pour la classe 3: Abeille.

741 images pour la classe 4: Sauterelle.

578 images pour la classe 5: Fourmies blanche.

335 images pour la classe 6: Cafard.

Pour tous les projets de Machine Learning, une étape d'évaluation des performances de ses modèles doit être effectuée. Fréquemment, la base de données est divisée en données d'entraînement, de validation et de test. Mais dans le cas de projets avec un nombre limité de données, cette méthode n'est pas vraiment recommandée. Vu que, la réservation d'une partie de la base de données pour la validation reviendrait à réduire la quantité déjà faible de données dont on dispose. Et même si on faisait ce sacrifice, les données de test seraient trop peu pour être représentatives de l'ensemble de données. Pour pallier ces problèmes, on utilise la validation croisée (cross-validation en anglais (CV)).

La validation croisée est une méthode statistique qui permet d'évaluer la capacité de généralisation d'un modèle. Lorsqu'on parle de cross-validation, on fait référence généralement au k-fold cross-validation, qui est sa variante la plus populaire. Avec cette méthode, nous profitons de toutes les données dont on dispose en les divisant en k parties égales (folds) sur lesquelles on effectue l'entraînement et le test d'un modèle pendant k itérations. A chaque itération, le modèle est entraîné sur k-1 folds et est testé sur le fold restant.

La figure suivante aide à mieux comprendre le principe de fonctionnement de CV.

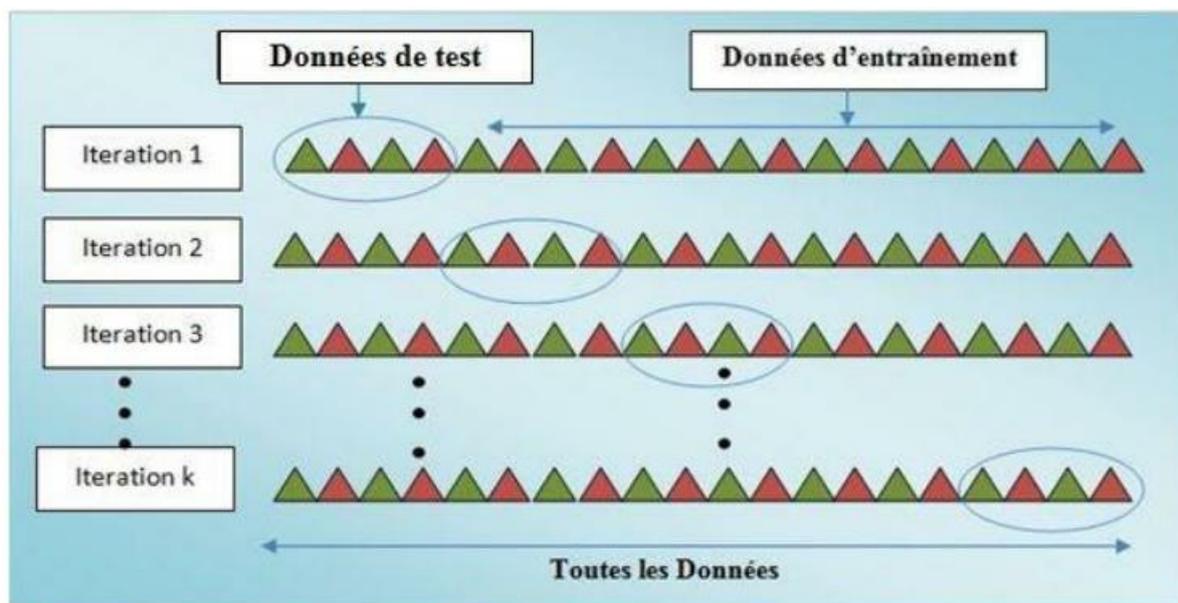


Figure 4.7: Fonctionnement de la méthode k-fold CV.

Sur cette figure, on prend l'exemple d'un dataset (ou jeu de données) ayant 20 triangles, qui a été divisé en 5 folds, chaque fold contient 4 triangles. Lors de la première itération, le premier fold sert de données de test, pendant que le reste des folds sont utilisés pour l'entraînement du modèle. A la seconde itération, le modèle est de nouveau entraîné, mais cette fois-ci, sur les données des folds 1, 3, 4 et 5 le second fold servant de données de test, et ainsi de suite...

Avec cette méthode, on s'assure que chaque triangle de notre dataset a servi une fois au moins au test et à l'entraînement.

Dans le cadre de notre mémoire, on travaille avec la méthode de cross validation, l'ensemble de ces images est divisé en 5 parties égales (folds) sur lesquelles on effectue l'entraînement et le test d'un modèle pendant 5 itérations.

Chaque fold contient 705 images. Dans chaque itération, un modèle est entraîné sur 4 folds et est testé sur le fold restant. Donc l'entraînement se fait sur 80% de l'ensemble des images et le test se fait sur les 20% restants.

4.5) Architectures et paramètres des modèles CNNs utilisés dans l'implémentation

Au cours de notre travail et afin de pouvoir réaliser la classification automatique de l'ensemble d'images, nous avons adopté deux modèles CNNs à savoir AlexNet et GoogleNet. Ces deux réseaux pré-entraînés sont conçus pour résoudre le problème de la classification des images où l'entrée est une image de l'une des 1000 classes différentes et la sortie est un vecteur de 1000 éléments. Donc leurs 3 dernières couches sont configurées pour 1000 classes.

Dans notre cas nous travaillons sur 6 classes, donc pour utiliser ces réseaux, nous les avons adapté à notre nouveau cas, c'est pour cela nous avons effectué les changements suivants :

A) Architecture d'AlexNet et de GoogleNet

- AlexNet

Nous avons gardé toutes les couches des réseaux telle quelles à l'exception des trois dernières. Ces trois couches doivent être adaptées pour le nouveau problème de classification. Alors ces trois dernières on les a remplacées par une couche entièrement connectée, une couche softmax et une couche de sortie de classification. Nous avons spécifié les options de la nouvelle couche entièrement connectée en fonction des nouvelles données en la définissant pour qu'elle ait la même taille que le nombre de classes dans la nouvelle situation, alors nous l'avons configuré pour 6 classes. Comme il est mentionné dans la partie du code suivante :

```
net=alexnet;
% extraire les couche du reseau sauf les 3 dernieres
layersTransfer = net.Layers(1:end-3);

% nombre de classes
numClasses = numel(categories(imdsTrain.Labels));
% Nouvelle couche
layers = [
layersTransfer
fullyConnectedLayer(numClasses, 'WeightLearnRateFactor', 20, 'BiasLearnRateFactor', 20)
softmaxLayer
classificationLayer];
```

La figure suivante représente une visualisation de notre architecture du réseau AlexNet après avoir modifié les dernières couches.

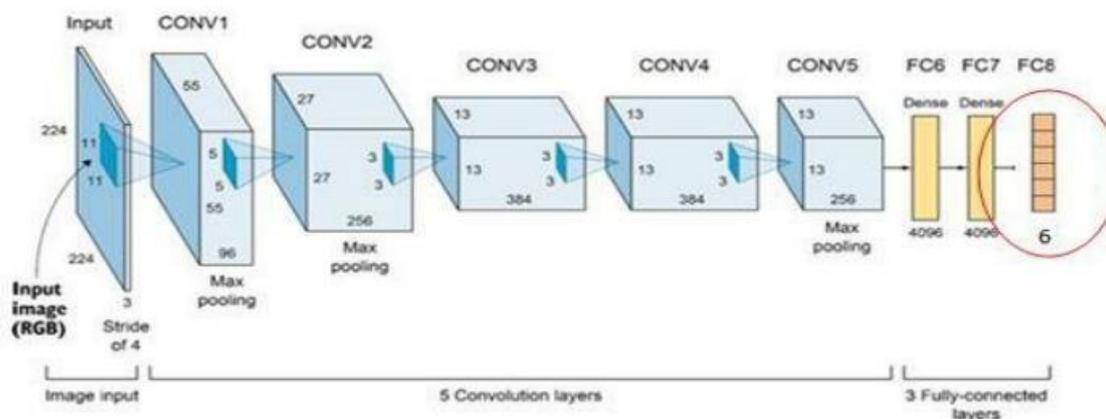


Figure 4.8 : Visualisation de l'architecture du réseau AlexNet après modifications.

- GoogleNet

Le rôle des couches convolutives du réseau est d'extraire des caractéristiques d'image que la dernière couche apprenable et la couche de classification finale utilisent pour classer l'image d'entrée. Dans GoogleNet les deux couches, 'loss3-classifier' et 'output', contiennent des informations sur la manière de combiner les fonctionnalités extraites par le réseau en probabilités de classe, valeur de perte et étiquettes prédites. Afin d'adapter ce réseau à l'ensemble d'images, nous avons remplacé ces deux couches par de nouvelles couches adaptées à cet ensemble. Comme il est mentionné dans la partie du code suivante.

```
net=googlenet;
lgraph = layerGraph(net);

% Number of categories
numClasses = numel(categories(imdsTrain.Labels));

% New Learnable Layer

newFCLayer = fullyConnectedLayer(numClasses, 'Name', 'new_fc', 'WeightLearnRateFactor', 10, 'BiasLearnR
lgraph = replaceLayer(lgraph, 'loss3-classifier', newFCLayer);
newClassLayer = classificationLayer('Name', 'new_classoutput');
lgraph = replaceLayer(lgraph, 'output', newClassLayer);
```

La figure suivante représente une visualisation de l'architecture du réseau **GoogleNet** après avoir modifié les dernières couches.

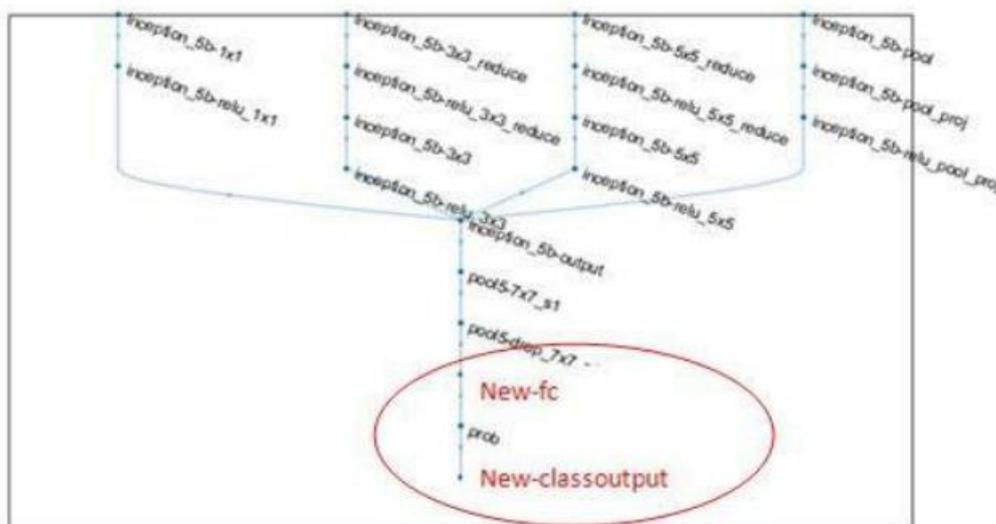


Figure 4.9 : Visualisation de l'architecture du réseau GoogleNet après modifications.

B) Les paramètres d'AlexNet et de GoogleNet

Le **tableau 4.1** représente les différents paramètres des modèles de CNNs utilisés dans le cadre de ce mémoire à savoir AlexNet et GoogleNet.

Modèle \ Paramètres	Epoch	Input images sizes	Learning rate
AlexNet	10	227	1e-4
GoogleNet	10	224	1e-4

Tableau 4.1 : Tableau représentant les paramètres des modèles AlexNet et GoogleNet

Epochs: Le nombre d'époques est un hyper paramètre qui définit le nombre de fois où l'algorithme d'apprentissage fonctionnera sur l'ensemble des données d'apprentissage.

Input images sizes: c'est la taille des images entrées. Ces tailles diffèrent d'un modèle à un autre.

Learning rate: ou taux d'apprentissage en français. Dans le domaine de l'apprentissage automatique, le taux d'apprentissage consiste à un paramètre de réglage, qui détermine la taille du pas à chaque itération tout en visant le minimum d'une fonction de perte. Il représente métaphoriquement la vitesse à laquelle un modèle d'apprentissage automatique apprend.

4.6) Les indicateurs de performance

Dans le but d'évaluer les performances des deux modèles (AlexNet et GoogleNet) que nous allons utiliser pour l'implémentation, nous allons calculer la précision, le rappel, la justesse et la spécificité. En plus de ceux-ci, nous allons présenter la matrice de confusion et la courbe ROC déjà présentées au chapitre 3.

- **La précision (*Precision*) :** La précision de la classification est le nombre total de prédictions correctes divisé par le nombre total de prédictions effectuées pour un ensemble de données.

$$Precision = \frac{TP}{TP + FP}$$

- **Le rappel (*Recall*) :** Proportion d'éléments bien classés par rapport au nombre d'éléments de la classe à prédire.

$$Rappel = \frac{TP}{TP + FN}$$

- **La justesse (Accuracy) :** Elle indique le pourcentage de bonnes prédictions.

$$\text{Accuracy} = \frac{TP+TN}{N} \quad \text{avec } N=TP+TN+FP+FN$$

- **La spécificité(Specificity):** Elle mesure la capacité d'un test à donner un résultat négatif lorsque l'hypothèse n'est pas vérifiée.

$$\text{Spécificité} = \frac{TN}{TN+FP}$$

La signification des symboles

- TP=True Positif. L'indication correcte d'une valeur positive comme étant un vrai positif
- TN=True Negatif. L'indication correcte d'une valeur négative comme étant un vrai négatif.
- FP=False Positif. L'indication correcte d'une valeur positive incorrecte comme étant un faux positif.
- FN=False Negatif. L'indication correcte d'une valeur négative incorrecte comme étant un faux négatif.

4.7) Résultats et Discussion

Dans cette section nous allons présenter et discuter les résultats obtenus suite à l'application de AlexNet et GoogleNet avec et sans augmentation des données sur la base d'images de ravageurs. Nous allons également donner une comparaison des performances de ces deux modèles.

➤ *Classification sans la technique d'augmentation des données*

Comme première étape du processus de formation (Training process), nous avons formé AlexNet et GoogleNet sans la technique d'augmentation de données. Les figures 4.10, 4.11, 4.12, 4.13 représentent les résultats obtenus.

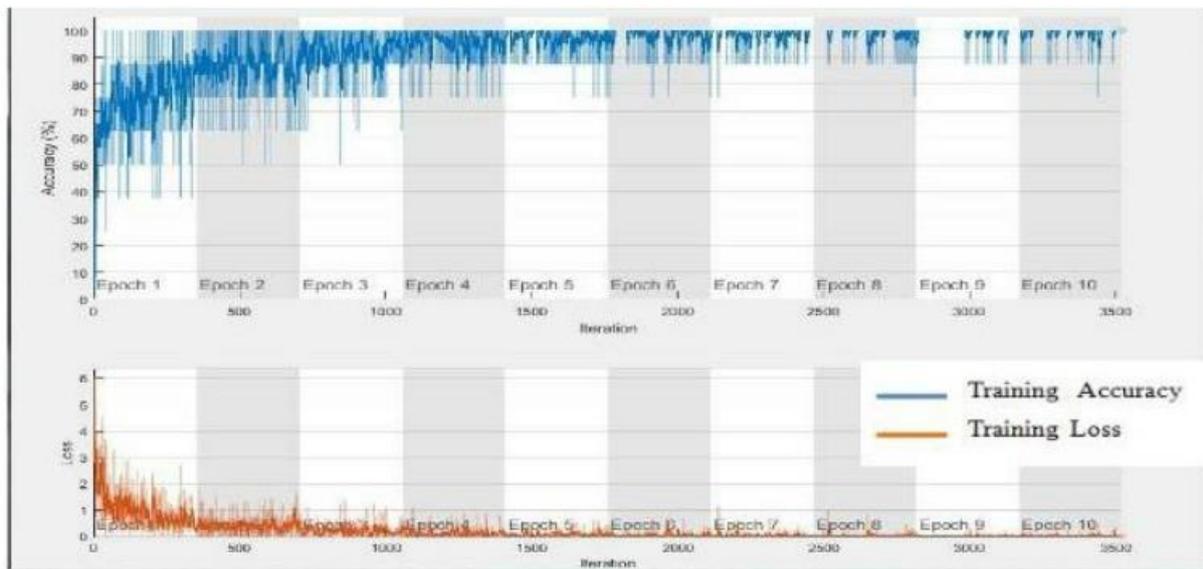


Figure 4.10 : Progression de l'apprentissage du réseau AlexNet sans augmentation des données.

Confusion Matrix: alexnet

Output Class	classe1	643 18.2%	13 0.4%	2 0.1%	29 0.8%	9 0.3%	13 0.4%	90.7% 9.3%
	classe2	8 0.2%	477 13.5%	2 0.1%	5 0.1%	12 0.3%	11 0.3%	92.6% 7.4%
	classe3	4 0.1%	3 0.1%	600 17.0%	1 0.0%	1 0.0%	2 0.1%	98.2% 1.8%
	classe4	54 1.5%	19 0.5%	11 0.3%	691 19.6%	18 0.5%	28 0.8%	84.2% 15.8%
	classe5	6 0.2%	5 0.1%	0 0.0%	6 0.2%	513 14.5%	19 0.5%	93.4% 6.6%
	classe6	16 0.5%	8 0.2%	3 0.1%	9 0.3%	25 0.7%	262 7.4%	81.1% 18.9%
			88.0% 12.0%	90.9% 9.1%	97.1% 2.9%	93.3% 6.7%	88.8% 11.2%	78.2% 21.8%
		classe1	classe2	classe3	classe4	classe5	classe6	
		Target Class						

Figure 4.11 : Matrice de confusion d'AlexNet pour l'ensemble de test.

La figure 4.10 représente les courbes de précision et de perte d'entraînement du modèle AlexNet avec un nombre d'épochs égal à 10. D'après cette figure nous pouvons observer que la précision d'entraînement est en train d'augmenter avec le nombre d'épochs, tandis que la perte d'entraînement diminue avec le nombre d'épochs.

La figure 4.11 représente la matrice de confusion d'AlexNet sans augmentation des données pour l'ensemble de test. Cette matrice montre les bonnes prédictions sur sa diagonale tout en donnant le nombre et le pourcentage de bonnes prédictions. D'une manière générale, AlexNet a obtenu un taux de précision de 90.3% avec un taux d'erreur de 9.7%.

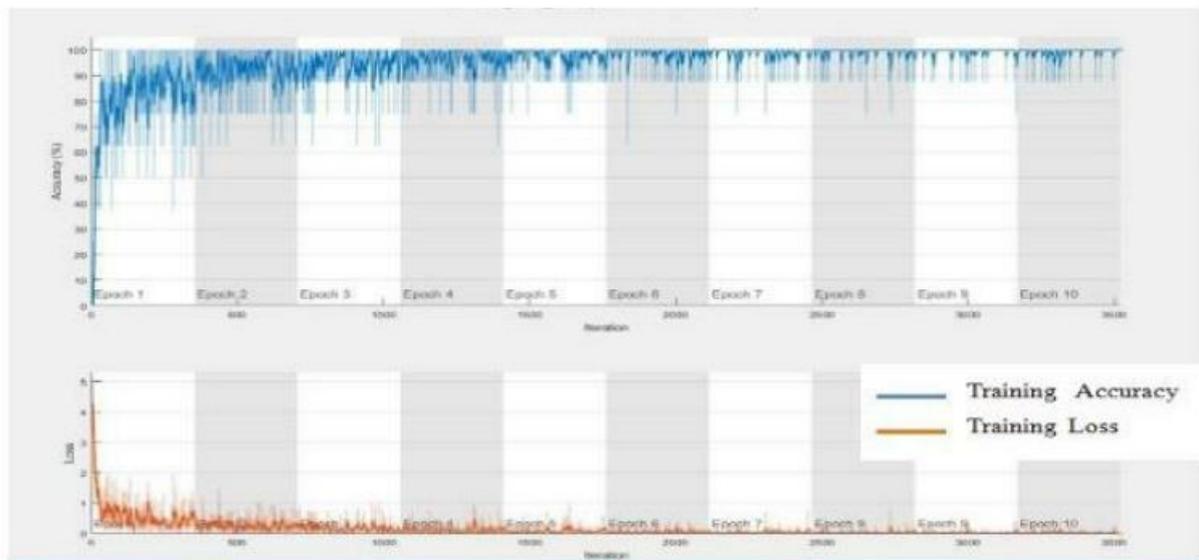


Figure 4.12 : Progression de l'apprentissage du réseau GoogleNet sans augmentation des données.

Confusion Matrix: GoogleNet

Output Class	classe1	classe2	classe3	classe4	classe5	classe6	
classe1	693 19.6%	11 0.3%	5 0.1%	23 0.7%	5 0.1%	6 0.2%	93.3% 6.7%
classe2	8 0.2%	489 13.9%	5 0.1%	7 0.2%	5 0.1%	7 0.2%	93.9% 6.1%
classe3	3 0.1%	6 0.2%	602 17.1%	3 0.1%	0 0.0%	3 0.1%	97.6% 2.4%
classe4	16 0.5%	9 0.3%	2 0.1%	697 19.8%	6 0.2%	13 0.4%	93.8% 6.2%
classe5	2 0.1%	3 0.1%	2 0.1%	3 0.1%	538 15.2%	16 0.5%	95.4% 4.6%
classe6	9 0.3%	7 0.2%	2 0.1%	8 0.2%	24 0.7%	290 8.2%	85.3% 14.7%
	94.8% 5.2%	93.1% 6.9%	97.4% 2.6%	94.1% 5.9%	93.1% 6.9%	86.6% 13.4%	93.8% 6.2%
	classe1	classe2	classe3	classe4	classe5	classe6	

Target Class

Figure 4.13 : Matrice de confusion de GoogleNet pour l'ensemble de test.

La figure 4.13 représente les courbes de précision et de perte d'entraînement du modèle **GoogleNet** avec un nombre d'époques égal à 10. D'après cette figure nous pouvons voir que la précision d'entraînement est en train d'augmenter avec le nombre d'époques, tandis que la perte d'entraînement diminue avec le nombre d'époques.

La figure 4.14 représente la matrice de confusion de **GoogleNet** sans augmentation des données pour l'ensemble de test. Cette matrice montre que **GoogleNet** a obtenu un taux de précision de 93.8% avec un taux d'erreur de 6.2%.

D'après ces figures, on remarque qu'avec l'utilisation de **GoogleNet** le taux d'erreur est réduit à 6.2% donc avec 3.5% par rapport à **AlexNet**.

➤ *Classification avec la technique d'augmentation des données*

Les algorithmes du Deep learning raffolent aujourd'hui de gros volumes d'images labellisés, et pour améliorer leurs performances on utilise la technique de *Data Augmentation*, qui consiste à faire subir des transformations à l'ensemble des images avant de les utiliser pour l'entraînement, tout en conservant leur structure. Pour chaque observation, on crée plusieurs variations. Ces transformations peuvent être classées en deux catégories: un ensemble de transformations affines comme le flip horizontal et/ou vertical, la rotation ou non affines comme la variation de luminosité et de contraste, le redimensionnement...etc. Ainsi le volume de données est artificiellement multiplié.

Comme une deuxième étape du processus de formation, nous avons formé **AlexNet** et **GoogleNet** cette fois avec la technique d'augmentation des données. Les **figures 4.14, 4.15, 4.16, 4.17** représentent les résultats obtenus.

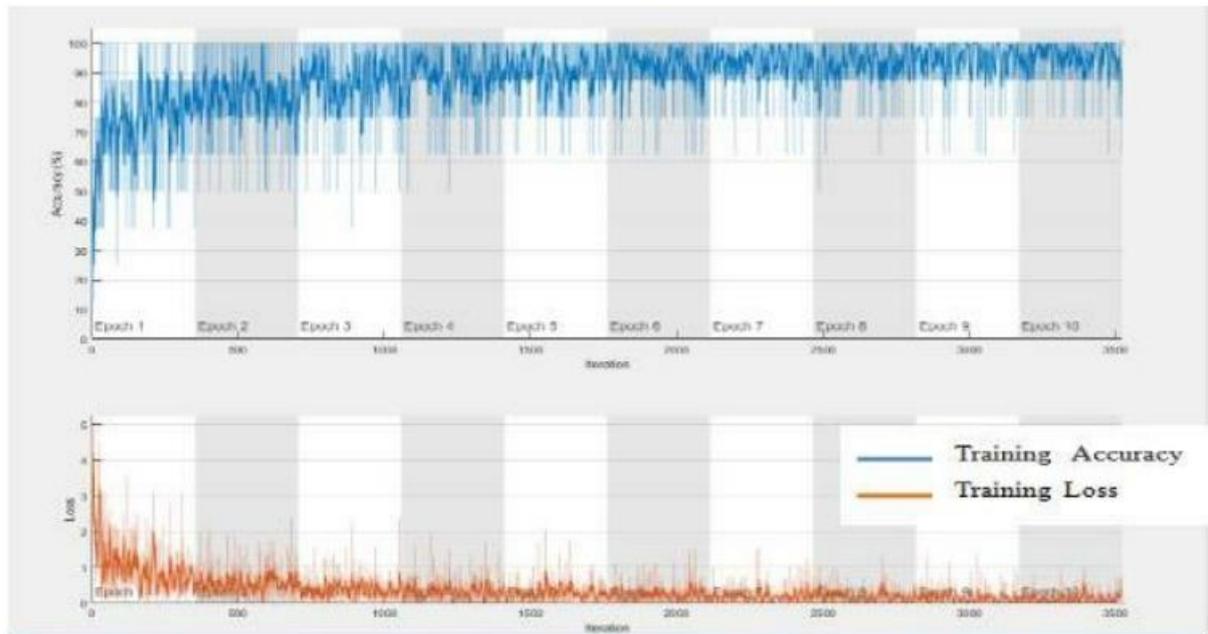


Figure 4.14 : Progression de l'apprentissage du réseau AlexNet avec augmentation des données.

Confusion Matrix: alexnetaug

Output Class	classe1	classe2	classe3	classe4	classe6	classe5	
classe1	662 18.8%	17 0.5%	12 0.3%	40 1.1%	13 0.4%	7 0.2%	88.1% 11.9%
classe2	6 0.2%	464 13.2%	2 0.1%	6 0.2%	5 0.1%	3 0.1%	95.5% 4.5%
classe3	7 0.2%	7 0.2%	597 16.9%	4 0.1%	2 0.1%	1 0.0%	96.6% 3.4%
classe4	33 0.9%	20 0.6%	5 0.1%	672 19.0%	16 0.5%	4 0.1%	89.6% 10.4%
classe6	13 0.4%	11 0.3%	0 0.0%	11 0.3%	278 7.9%	18 0.5%	84.0% 16.0%
classe5	10 0.3%	6 0.2%	2 0.1%	8 0.2%	21 0.6%	545 15.4%	92.1% 7.9%
	90.6% 9.4%	88.4% 11.6%	96.6% 3.4%	90.7% 9.3%	83.0% 17.0%	94.3% 5.7%	91.2% 8.8%
	classe1	classe2	classe3	classe4	classe6	classe5	

Target Class

Figure 4.15 : Matrice de confusion d'AlexNet pour l'ensemble de test.

La figure 4.14 Représente les courbes la précision et la perte d'entraînement du modèle AlexNet avec la technique d'augmentation des données avec un nombre d'epochs égal à 10. D'après cette figure nous pouvons observer que la précision d'apprentissage est en train d'augmenter pour atteindre 100% après 10 epochs, tandis que la perte diminue au fur et à mesure.

À partir de la figure 4.15 qui représente la matrice de confusion d'AlexNet pour l'ensemble de test en appliquant la technique d'augmentation des données, on remarque qu'il y a une amélioration de la précision de classification du modèle (91.2%) et du taux d'erreur (8.8%) par rapport à la première formation (sans la technique d'augmentation des données) où la précision était de 90.3 % et le taux d'erreur était de 9.7%..

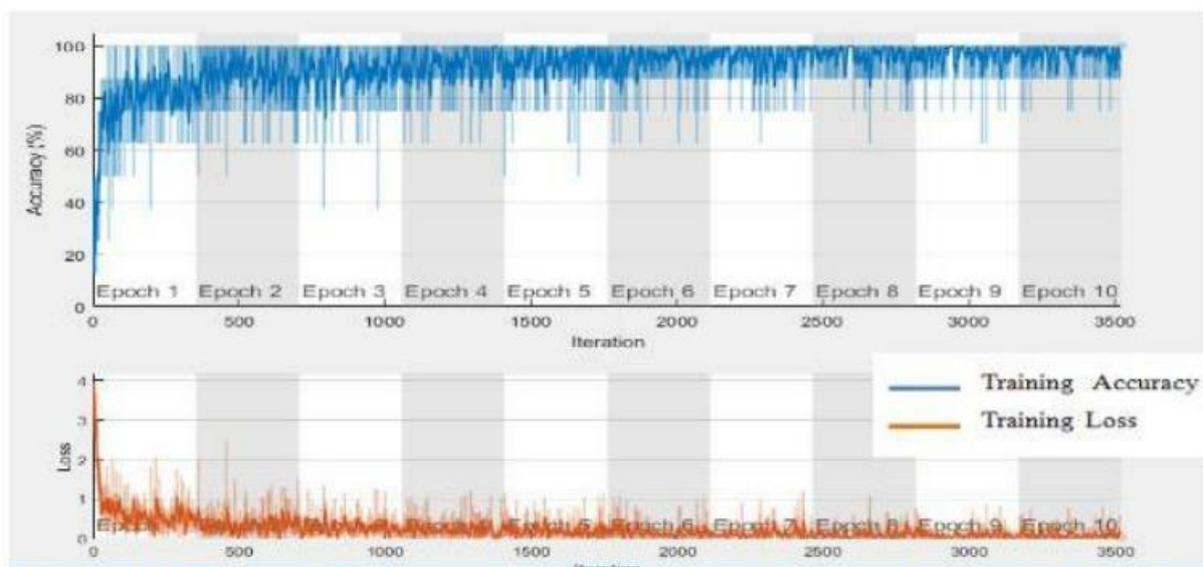


Figure 4.16 : Progression de l'apprentissage du réseau GoogleNet avec augmentation des données

Confusion Matrix: GoogleNetaug

	classe1	classe2	classe3	classe4	classe6	classe5		
Output Class	classe1	699 19.8%	7 0.2%	5 0.1%	24 0.7%	8 0.2%	8 0.2%	93.1% 6.9%
	classe2	3 0.1%	497 14.1%	6 0.2%	7 0.2%	3 0.1%	6 0.2%	95.2% 4.8%
	classe3	3 0.1%	2 0.1%	598 17.0%	2 0.1%	2 0.1%	0 0.0%	98.5% 1.5%
	classe4	20 0.6%	11 0.3%	5 0.1%	697 19.8%	16 0.5%	2 0.1%	92.8% 7.2%
	classe6	5 0.1%	3 0.1%	3 0.1%	7 0.2%	292 8.3%	24 0.7%	87.4% 12.6%
	classe5	1 0.0%	5 0.1%	1 0.0%	4 0.1%	14 0.4%	538 15.2%	95.6% 4.4%
		95.6% 4.4%	94.7% 5.3%	96.8% 3.2%	94.1% 5.9%	87.2% 12.8%	93.1% 6.9%	94.1% 5.9%
	classe1	classe2	classe3	classe4	classe6	classe5		
	Target Class							

Figure 4.17 : Matrice de confusion de GoogleNet pour les données de test.

La figure 4.16 Représente la courbe de précision et de perte d'entraînement de GoogleNet avec augmentation de données. D'après cette figure nous pouvons observer que la précision d'apprentissage augmente à travers les epochs, tandis que la perte diminue au fur et à mesure.

Selon la figure 4.17 qui représente la matrice de confusion de GoogleNet pour l'ensemble de test en appliquant la technique d'augmentation des données, on voit qu'il ya une amélioration dans les résultats obtenus par ce modèle avec un pourcentage de précision de 94.1 et un taux d'erreur de 5.9 par rapport à la première formation (c.à.d. avant l'augmentation des données).

On remarque dans ce cas aussi, qu'avec l'utilisation de GoogleNet le taux d'erreur est réduit par rapport à AlexNet avec un pourcentage de 2.9% (de 8.8% à 5.9%).

Le **Tableau 4.2** suivant va résumer les valeurs des taux de précision et d'erreur des modèles pour l'ensemble de test avec et sans augmentation.

	Avec ou sans augmentation	Précision	Taux d'erreur
AlexNet	Sans augmentation	90.3%	9.7%
	Avec augmentation	91.2%	8.8%
GoogleNet	Sans augmentation	93.8%	6.2%
	Avec augmentation	94.1%	5.9%

Tableau 4.2: Les valeurs des taux de précision et d'erreur des modèles avec et sans augmentation.

**❖ Comparaison des résultats des deux modèles (AlexNet, GoogleNet)
avant et après augmentation**

Dans cette partie nous présentons d'autres résultats obtenus par les deux modèles en utilisant d'autres paramètres de performance qui sont : la précision, le rappel, la justesse et la spécificité.

Le **Tableau 4.3** présente un récapitulatif des toutes les valeurs de ces indicateurs de performances pour AlexNet et GoogleNet avec et sans augmentation des données.

	Avec ou sans augmentation	Précision	Rappel	Justesse	Spécificité
AlexNet	Sans augmentation	88%	89%	96%	96%
	Avec augmentation	90%	90%	96%	97%
GoogleNet	Sans augmentation	91%	92%	97%	95%
	Avec augmentation	93%	93%	97%	98%

Tableau 4.3: Tableau comparatif des valeurs des indicateurs de performance de GoogleNet et AlexNet.

➤ **Classification sans la technique d'augmentation de s données.**

D'après le **Tableau 4.3** on constate qu'après avoir calculé les indicateurs de performance afin d'évaluer l'efficacité du modèle AlexNet sans augmentation des données, on a abouti à une précision égale à 88%, un rappel de 89%, une justesse de 96% et une spécificité égale à 96%. De l'autre côté, le modèle GoogleNet a donné une précision de 91%, un rappel de 92%, une justesse de 97% et une spécificité égale à 95%.

De là, on remarque que l'utilisation de GoogleNet permet d'obtenir des valeurs plus satisfaisantes pour les différents indicateurs de performances, une meilleure précision, un meilleur rappel, et une justesse et une spécificité améliorées par rapport à AlexNet.

Cela on peut le distinguer plus clairement des **figures 4.18** et **4.19** qui représentent respectivement la courbe ROC d' AlexNet et la courbe ROC de GoogleNet de l'ensemble de test sans l'augmentation de données pour les deux cas, vu qu'on a déjà présenté la courbe ROC comme étant un graphique qui représente l'ensemble de performances d'un modèle de classification pour tous les seuils de la classification. Avec cet outil on peut avoir toutes les informations sur la performance du classifieur, et cela en utilisant l'aire sous la courbe. Plus elle se rapproche de 1, plus le classifieur est performant.

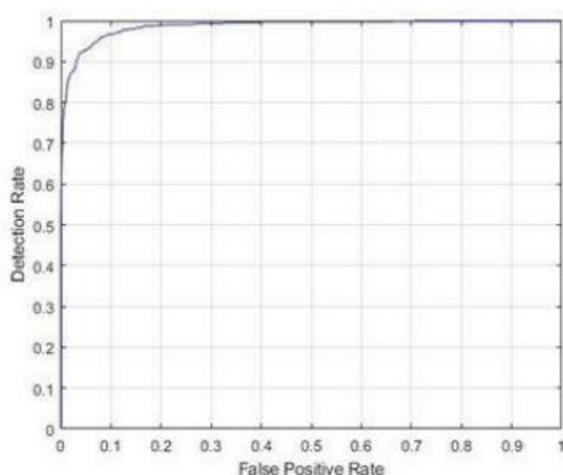


Figure 4. 18 : Courbe ROC d'AlexNet sans augmentation des données de l'ensemble de test.

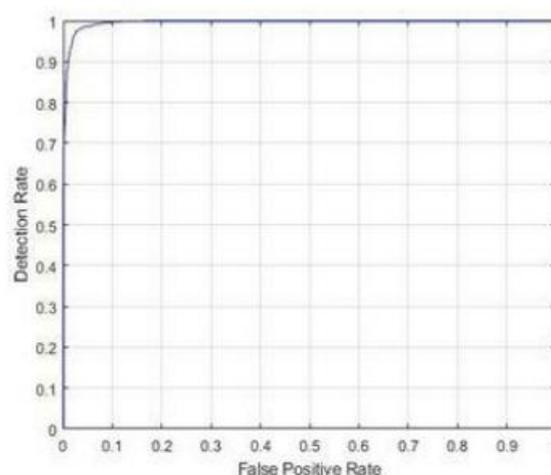


Figure 4. 19 : Courbe ROC de GoogleNet sans augmentation des données de l'ensemble de test.

D'après les deux figures ci-dessus on remarque que l'aire sous la courbe de GoogleNet est plus grande que celle sous la courbe d'AlexNet, donc celle de GoogleNet est plus proche à 1 que celle d'AlexNet. De là, on peut dire que le modèle GoogleNet est plus performant que AlexNet.

➤ **Classification avec la technique d'augmentation des données.**

D'après le **Tableau 4.3** on constate que, après avoir calculé les indicateurs de performance afin d'évaluer l'efficacité du modèle AlexNet avec augmentation des données, on a abouti à une précision à 90%, un rappel de 90%, une justesse de 96% et une spécificité égale à 97%. Par contre, le modèle GoogleNet a donné une précision de 93%, un rappel de 93%, une justesse de 97% et une spécificité égale à 98%.

De là, on remarque que l'utilisation de GoogleNet en appliquant l'augmentation de données permet d'obtenir toujours des valeurs plus satisfaisantes pour les différents indicateurs de performances par rapport à AlexNet.

Cela on peut le distinguer plus clairement des **figures 4.20** et **4.21** qui représentent respectivement la courbe ROC d'AlexNet et la courbe ROC de GoogleNet de l'ensemble de test avec augmentation des données.

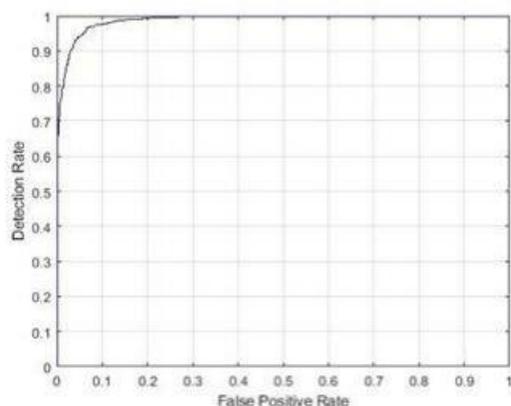


Figure 4. 20 : Courbe ROC d'AlexNet avec augmentation des données de l'ensemble de test.

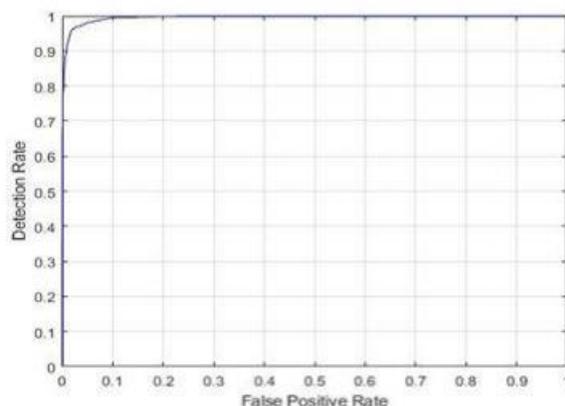


Figure 4. 21 : Courbe ROC de GoogleNet avec augmentation des données de l'ensemble de test.

D'après les deux figures ci-dessus on remarque que l'aire sous la courbe de GoogleNet est plus grande que celle sous la courbe d'AlexNet, donc celle de GoogleNet est plus proche de 1 que celle d'AlexNet. De là, on peut dire que le modèle GoogleNet avec augmentation est plus performant qu'AlexNet avec augmentation.

En résumé:

Après avoir effectué l'étude sur l'ensemble des résultats obtenus par l'application des modèles de Deep Learning : AlexNet et GoogleNet avec et sans la technique d'augmentation des données sur l'ensemble d'images de ravageurs, les points essentiels suivants peuvent être tirés:

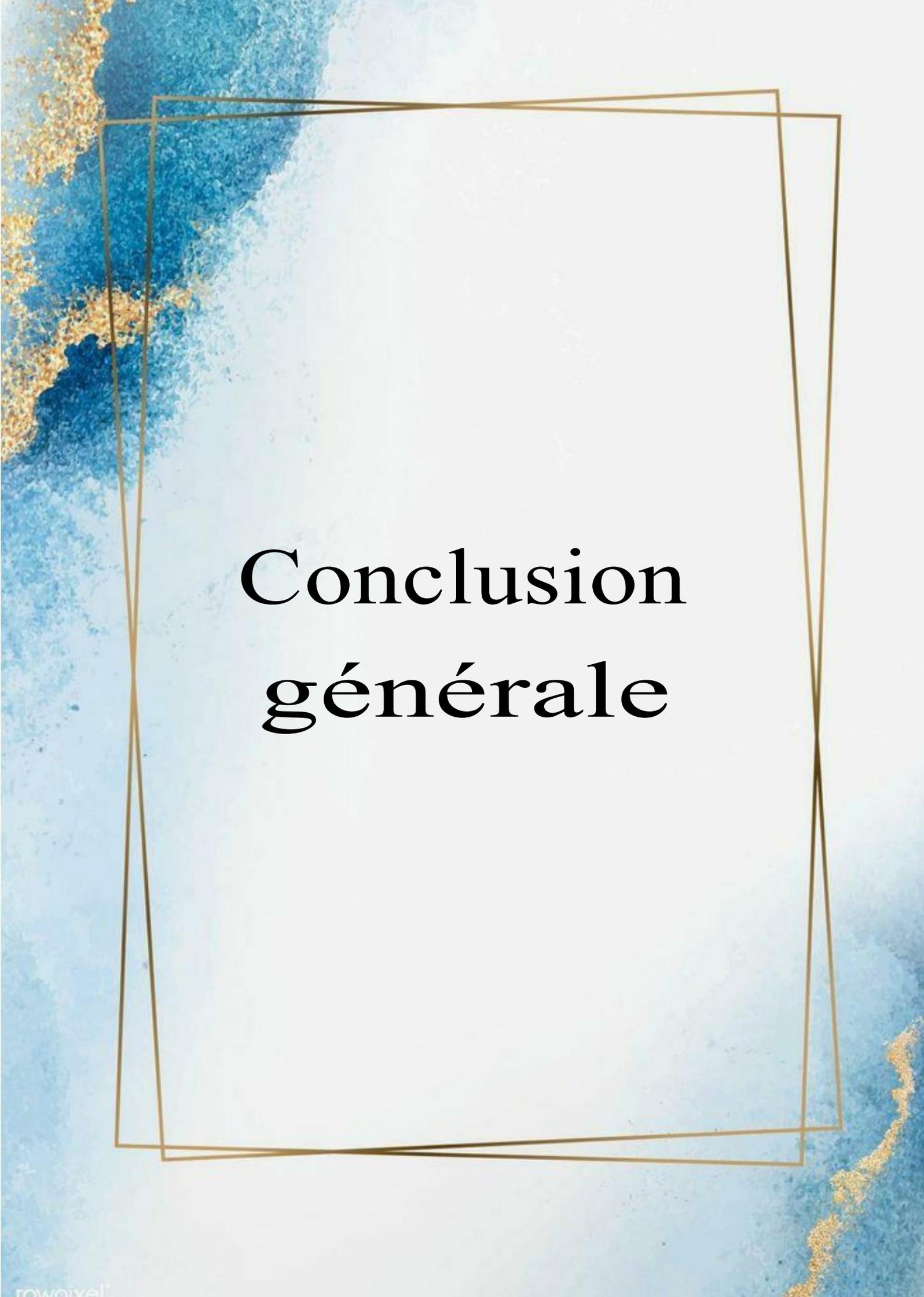
- L'utilisation de GoogleNet nous a permis d'obtenir un taux d'erreur de classification d'images réduit par rapport à l'utilisation d'AlexNet avec et sans l'utilisation de la technique d'augmentation des données.
- L'application de la technique d'augmentation de données sur l'ensemble des images utilisées dans les deux modèles de Deep Learning, permet aussi de réduire le taux

d'erreur ou de perte et cela d'une valeur plus élevée que leurs utilisations sans augmentation.

- Plus la quantité de données utilisées comme entrées des modèles est grande, plus on obtient des résultats de classification meilleurs (Taux d'erreur réduit et valeurs des indices de performances élevées). Donc, un modèle de Deep Learning est plus performant s'il est entraîné sur un énorme flux de données.

4.8) Conclusion

Dans ce dernier chapitre nous avons proposé une approche d'identification de ravageurs basée sur les réseaux de neurones convolutionnels. Cette dernière se base sur deux modèles CNN à savoir AlexNet et GoogleNet. Les résultats obtenus ont montré que GoogleNet est plus performant qu'AlexNet avec une précision de 94.1% avec la technique d'augmentation des données et de 93.8% sans la technique d'augmentation des données. Les expérimentations ont montré l'efficacité de notre approche dans l'identification de ravageurs.

The background features a soft, painterly texture in shades of light blue and white, with irregular splatters of a darker blue and shimmering gold particles. A thin, gold-colored double-line frame is centered on the page, enclosing the text.

Conclusion générale

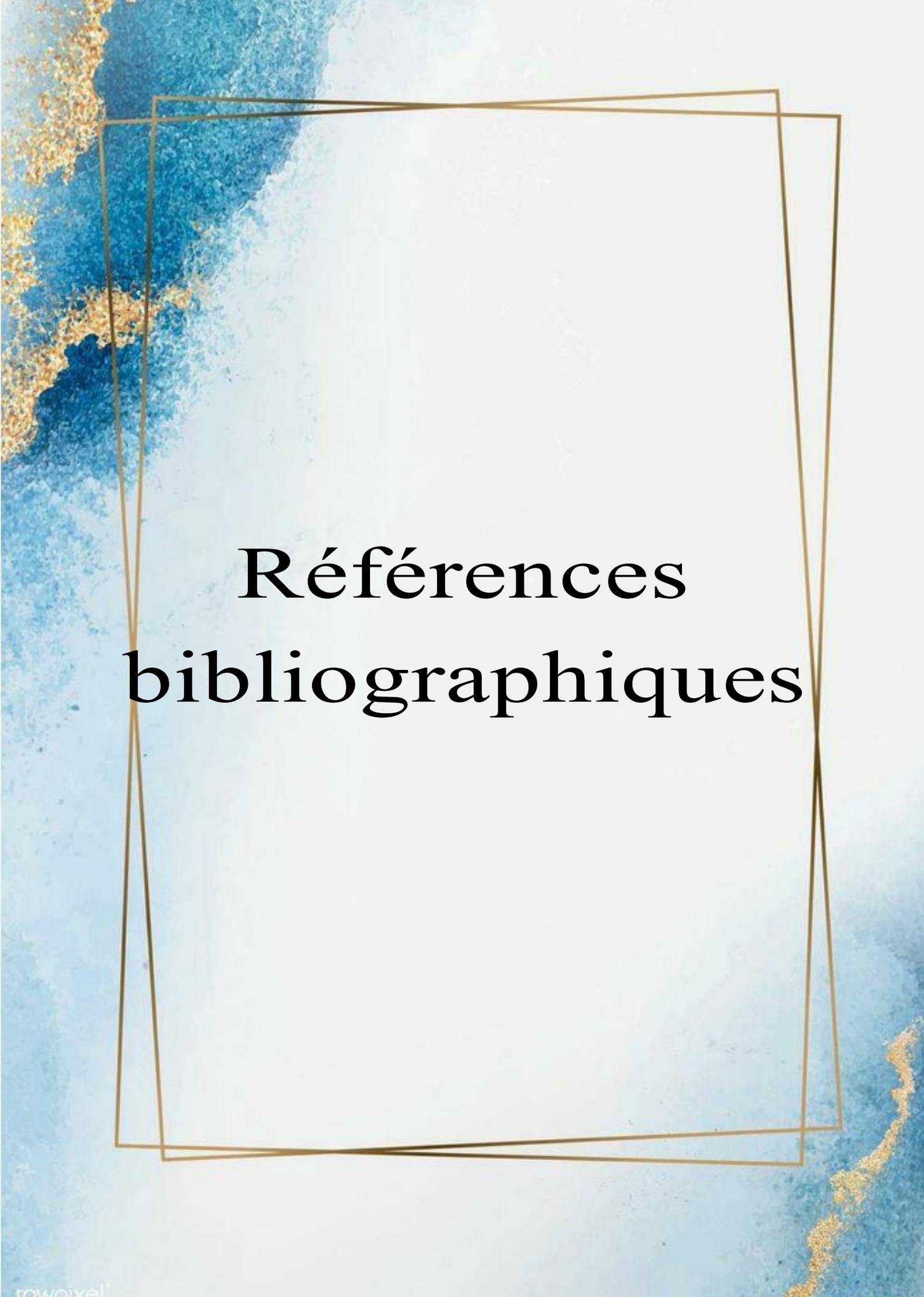
Conclusion générale

Au cours de la réalisation de notre mémoire qui porte sur la classification d'images d'insectes nuisibles, nous avons utilisé le Deep Learning, la méthode d'apprentissage qui a montré de grandes performances ces dernières années dans ce domaine et nous avons adopté la méthode CNNs comme méthode de classification. Nous avons discuté des notions fondamentales des réseaux de neurones en général et des réseaux de neurones convolutionnels en particulier. Nous avons introduit les CNNs en présentant les différents types de couches utilisées dans la classification d'images. Le choix de cette méthode est justifié par la simplicité et l'efficacité de cette dernière.

Dans ce projet nous avons utilisé deux modèles prédéfinis des réseaux de neurones convolutionnels le GoogleNet et l'Alexnet, que nous avons exécutés une première fois sans appliquer la technique d'augmentation des données et une seconde fois en l'appliquant, afin d'augmenter le nombre de nos images d'entrée et avoir des résultats meilleurs des modèles. Ce travail a été réalisé afin de pouvoir établir une comparaison entre les résultats des deux modèles et dans les deux formations (avec et sans augmentation) pour déduire le modèle le plus performant en termes de réduction de taux d'erreur et d'augmentation des valeurs des indicateurs de performance. Ainsi, nous avons constaté que le GoogleNet était plus performant que l'AlexNet et le rendu de l'utilisation de l'augmentation était meilleur (une précision de 94.1% avec la technique d'augmentation des données et de 93.8% sans la technique d'augmentation des données). Les expérimentations ont montré l'efficacité de notre approche dans l'identification de ravageurs.

Notre travail n'est que dans sa version initiale, nous pouvons dire qu'il reste ouvert pour effectuer des améliorations afin d'obtenir des résultats beaucoup plus satisfaisants dont nous pouvons recommandés :

- L'utilisation d'une base de données d'images plus importante que celle utilisée.
- L'utilisation des réseaux de neurones convolutionnels déployé sur un GPU au lieu d'un CPU, pour palier au problème de temps d'exécution couteux que nous avons rencontré lors de la phase d'implémentation.
- L'utilisation d'autres modèles de classification qui ont montré leurs efficacités dans ce domaine comme le ResNet.



Références bibliographiques

Références Bibliographiques

• Chapitre 1 : Principes et concepts de base de l'intelligence artificielle

[1.1] : Mémoire master sous l'intitulé : L'intelligence artificielle principe, outils et

Objectifs réalisé par Zara Islem de UNIVERSITE BADJI MOKHTAR ANNABA 2019.

[1.2] : Mémoire de fin d'études de la 2ème année de Master : L'INTELLIGENCE ARTIFICIELLE AU SERVICE DE LA RECHERCHE CLINIQUE (2018 – 2019) université de Lille.

[1.3] : <http://www.onisep.fr/Pres-de-chez-vous/Hauts-de-France/Amiens/Informations-metiers/Le-numerique-et-l-intelligence-artificielle/L-intelligence-artificielle-et-ses-domaines-d-application>

[1.4] : <http://tpeai.e-monsite.com/pages/ii-les-limites-de-l-intelligence-artificielle.html>

L'ensemble de livres lus afin de réaliser le chapitre 1.

- ✓ Artificial Intelligence: The Very Idea – 1 janvier 1985 de John Haugeland (Auteur)
- ✓ Artificial Intelligence: Can Computers Think? -06 janvier 1978 de Richard Bellman (Auteur).
- ✓ The Age of Intelligent Machines – 19 mars 1992 de Ray Kurzweil (Auteur)
- ✓ Artificial Intelligence – 1 janvier 1991 de Elaine Rich (Auteur), Kevin Knight (Auteur)
- ✓ Introduction to Artificial Intelligence – 1 janvier 1985 d'Eugene Charniak (Auteur), D. McDermott (Auteur).
- ✓ Artificial Intelligence – 30 avril 1992 de Winston (Auteur).
- ✓ Computational Intelligence: A Logical Approach - 8 janvier 1998 de David Poole (Auteur), Alan Mackworth (Auteur), Randy Goebel (Auteur).
- ✓ Artificial Intelligence: A New Synthesis-15 avril 1998 de Nils J. Nilsson (Auteur).

• Chapitre 2: Outils et algorithmes mis en œuvre pour l'IA

[2.1]: https://www.college-de-france.fr/media/yann-lecun/UPL4485925235409209505_Intelligence_Artificielle_Y._LeCun.pdf

[2.2] : <https://www.oracle.com/ca-fr/artificial-intelligence/what-is-machine-learning.html>

[2.3] : Intelligence artificielle 3e édition – 10 décembre 2010 de Stuart Russel (Auteur), Peter Norvig (Auteur).

[2.4] : "Backpropagation appliquée à la reconnaissance de code postal manuscrite", *Neural Computation*, 1989. Y. LeCun, B. Boser, JS Denker, D. Henderson, RE Howard, W. Hubbard, et al. (Auteurs)

[2.5]:<https://www.ionos.fr/digitalguide/web-marketing/search-engine-marketing/deep-learning/>

[2.6] : <https://www.lebigdata.fr/deep-learning-definition>

[2.7]:<https://www.futura-sciences.com/tech/definitions/intelligence-artificielle-deep-learning-17262/>

[2.8]:<https://www.technologies-ebusiness.com/enjeux-et-tendances/introduction-autour-deep-learning-application-aux-iot-partie-2>

• **Chapitre 3: La classification d'images avec le Deep Learning**

[3.1] : « Classification des images avec les réseaux de neurones convolutionnels ». Mémoire de fin d'études pour l'obtention du diplôme de Master en informatique. 03/07/2017

[3.2] : <https://www.creatis.insa-lyon.fr/site7/fr/node/46474>

[3.3]:<https://www.datagenius.fr/single-post/reconnaissance-d-image-intelligence-artificielle-ai-compare>

[3.4]:<https://www.usine-digitale.fr/article/facebook-va-plus-loin-dans-la-reconnaissance-faciale-avec-deepface.N249640>

[3.5]:https://www.memoireonline.com/04/12/5741/m_Conception-et-realisation-dun-systeme-de-vote-electronique-pour-le-parlement-cas-du-senat-c9.html

[3.6] : « Développement et automatisation de méthodes de classification à partir de séries temporelles d'images de télédétection - Application aux changements d'occupation des sols et à l'estimation du bilan carbone ». thèse en vue de l'obtention du DOCTORAT de l'université de TOULOUSE. Présentée par Antoine Masse le 11 octobre 2013.

[3.7]:<https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-architecture-for-giving-pathways-future/>

[3.8]:<https://www.anakeyn.com/2020/01/13/classification-de-pages-web-via-deep-learning-reseau-de-neurones-convolutif/>

[3.9]:<https://www.natural-solutions.eu/blog/la-reconnaissance-dimage-avec-les-reseaux-de-neurones-convolutifs#:~:text=Qu'est%20ce%20que%20la,%C3%A0%20la%20reconnaissance%20d'image.>

[3.10] :<https://datakeen.co/3-deep-learning-architectures-explained-in-human-language/>

[3.11]:<https://datasciencetoday.net/index.php/fr/deep-learning/173-les-reseaux-de-neurones-convolutifs>

[3.12]:<https://datasciencetoday.net/index.php/fr/deep-learning/173-les-reseaux-de-neurones-convolutifs>

[3.13]:<https://medium.com/@yannicksergeobam/comprendre-les-reseaux-de-neurones-convolutifs-cnn-d5f14d963714>

[3.14]:<https://medium.com/@yannicksergeobam/comprendre-les-reseaux-de-neurones-convolutifs-cnn-d5f14d963714>

[3.15] : Classification de données massives de télédétection . (*Classification of big remote sensing data*). Nicolas Audebert: Bretagne Loire University, Rennes, France (2018)

[3.16] : Reconnaissance d'iris par les réseaux de neurones convolutionnels ». Mémoire Présenté en vue de l'obtention du diplôme de MASTER. 2019.

[3.17]:<https://livebook.manning.com/book/grokking-deep-learning-for-computer-vision/chapter-5/v-3/34>

[3.18] : « Going Deeper with Convolutions » [Christian Szegedy](#), [Wei Liu](#), [YangqingJia](#), [Pierre Sermanet](#), [Scott Reed](#), [DragomirAnguelov](#), [DumitruErhan](#), [Vincent Vanhoucke](#), [Andrew Rabinovich](#) / 7 Sep 2014

[3.19]: <https://medium.com/@RaghavPrabhu/cnn-architectures-lenet-alexnet-vgg-googlenet-and-resnet-7c81c017b848>

[3.20] : <https://tariq-hasan.github.io/concepts/computer-vision-cnn-architectures/>

- **Chapitre 4: Implémentation**

[4.1] : <https://www.insectimages.org/>

[4.2] : <https://www.insectimages.org/browse/taxthumb.cfm?order=131>

[4.3] : <https://www.insectimages.org/browse/taxthumb.cfm?order=58>

[4.4] : <https://www.insectimages.org/browse/taxthumb.cfm?order=98>

[4.5] : <https://www.insectimages.org/browse/taxthumb.cfm?order=159>

[4.6] : <https://www.insectimages.org/browse/taxthumb.cfm?order=121>

[4.7] : <https://www.insectimages.org/browse/taxthumb.cfm?order=369>