

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieure et de la Recherche



جامعة بجاية
Tasdawit n Bgayet
Université de Béjaïa



Scientifique

Université Abderrahmane Mira

Faculté de la Technologie

*Département d'Automatique, Télécommunication et
d'Electronique*

Filière : Télécommunications.

Spécialité : Réseaux et Télécommunications

Thème

**Identification du mélanome en se basant sur les
attributs texture couleur**

Réalisé par :

Saadou Nafiou

Amadou Tidjani

Encadré par :

Mr Reda Kasmi

Année Universitaire : 2024/2025

Populaire et Démocratique Algérienne République
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Déclaration sur l'honneur
Engagement pour respecter les règles d'authenticité scientifique dans
l'élaboration d'un travail de recherche

Arrêté ministériel n° 1082 du 27 décembre 2020 (*)
fixant les règles relatives à la prévention et la lutte contre le plagiat

Je soussigné,

Nom : Saadou Nafion

Prénom : Amadou Tidjani

Matricule : 19198NER6730

Spécialité et/ou Option : Télécommunication / Réseaux et Télécommunication

Département : Automatique Télécommunication Electronique

Faculté : Technologie

Année universitaire : 2024/2025

et chargé de préparer un mémoire de (Licence, Master, Autres à préciser) : Master

Intitulé : Identification du mélanome en se basant
sur les attributs texture couleur

déclare sur l'honneur, m'engager à respecter les règles scientifiques, méthodologiques,
et les normes de déontologie professionnelle et de l'authenticité académique requises dans
l'élaboration du projet de fin de cycle cité ci-dessus.

Fait à Béjaïa le
09/07/2025

Signature de l'intéressé

(*) Lu et approuvé

.....

(*) Arrêté ministériel disponible sur le site www.univ-bejaia.dz/formation (rubrique textes réglementaires)

Remerciement

Tout d'abord, je tiens à remercier DIEU de m'avoir donné la force et le courage de mener à bien ce travail.

Je tiens à remercier mon encadrant Monsieur Reda Kasmi, pour sa patience, et surtout pour sa confiance, ses remarques et ses conseils, et sa bienveillance.

Mes vifs remerciements vont également aux membres de jury pour l'honneur qu'ils ont fait en acceptant de juger ce mémoire de fin d'étude et de l'enrichir par leur proposition.

Je remercie toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

Sommaire

Table des matières

Liste des tableaux

Liste des figures

Liste des abréviations

1	Introduction générale	1
1.1	Contexte Médical	1
1.2	Problématique.....	1
1.3	Objectif du travail.....	1
1.4	Organisation du mémoire.....	1
2	Chapitre I : Synthèse Bibliographique.....	3
2.1	Le mélanome (aspects médicaux).....	3
2.1.1	Le mélanome	3
2.1.2	Le développement du mélanome.....	4
2.1.3	L'importance du diagnostic précoce.....	6
2.2	Intelligence Artificielle, Apprentissage Profond, Imagerie Médicale, Apprentissage Profond en Imagerie Médicale	7
2.2.1	Intelligence Artificielle	7
2.2.2	Apprentissage Profond (Deep Learning)	8
2.2.3	Imagerie Médicale	8
2.2.4	Apprentissage Profond en Imagerie Médicale	9
2.2.5	Les bases de données pour la dermatologie (focus sur Melanoma Cancer Dataset)	10
2.2.6	Réseaux des Neurones	10
2.2.7	Réseaux de neurones convolutif(CNN)	11
2.2.8	Les architectures populaires de CNN(LeNet, AlexNet, ResNet, MobileNet,...)	15
2.2.9	Travaux récents sur la classification des lésions cutanées.....	20
3	Chapitre II : Données, Outils et Méthodologie	24
3.1	Présentation de la base de données Melanoma Cancer Dataset	24

3.2	Prétraitement sur la base de données	24
3.3	Présentation des architectures utilisées	24
3.3.1	Modèle CNN personnalisé	24
3.3.2	MobileNetV2.....	25
3.3.3	ResNet50.....	25
3.4	Environnement de travail.....	25
3.5	Métriques et protocole d'évaluation	26
4	Chapitre III : Résultats et Discussion.....	28
4.1	Résultats de l'entraînement	28
4.2	Discussion sur les résultats obtenus	30
5	Conclusion générale.....	32
6	Références bibliographiques.....	33

Liste des tableaux

Tableau I.1 :	Résumé de l'architecture AlexNet.....	15
Tableau I.2 :	Résumé de l'architecture VGGNet.....	16
Tableau III.1 :	Tableau de comparaison des performances	29

Liste des figures

Chapitre I

Figure I.1 :	Le mélanome	2
Figure I.2 :	Développement de mélanome	3
Figure I.3 :	Types de mélanome	3
Figure I.4 :	IA, ML, DP	6
Figure I.5 :	Structure d'une neurone artificiel.....	9
Figure I.6 :	Architecture standard d'un CNN	9
Figure I.7 :	Principe de la convolution.....	10
Figure I.8 :	Les types de Pooling.....	11

Figure I.9 : Histoire évolutive des CNNs montrant les innovations architecturales ...	13
Figure I.10 : L'architecture LeNet	14
Figure I.11 : L'architecture AlexNet.....	14
Figure I.12 : L'architecture du VGGNet	15
Figure I.13 : L'architecture de ResNet.....	16
Figure I.14 : L'architecture du MobileNet	16
Figure I.15 : L'architecture de GoogleNet.....	17
Figure I.16 : Application FotoSkin.....	19
Figure I.17 : Application Doctor Mole	19
Figure I.18 : Application SkinVision.....	20
Chapitre II	
Figure II.1 : Interface de Kaggle	23
Figure II.2 : Les formules de métriques clés utilisés.....	24
Figure II.3 : Matrice de confusion.....	24
Chapitre III	
Figure III.1 : Résultats d'entraînement du CNN personnalisé	25
Figure III.2 : Résultats d'entraînement de MobileNetV2	26
Figure III.3 : Résultats d'entraînement de ResNet50.....	26
Figure III.4 : Matrices de confusion.....	27

Liste des abréviations

CNN : Réseau de Neurones Convolutifs (Convolutional Neural Network)

DL : Apprentissage Profond (Deep Learning)

IA : Intelligence Artificielle

ML : Apprentissage Automatique (Machine Learning)

ReLU : Rectified Linear Unit

FC : Couche Entièrement Connectée (Fully Connected)

POOL : Couche de Pooling

CONV : Couche de Convolution

AUC : Area Under the Curve

TP : True Positives

TN : True Negatives

FP : False Positives

FN : False Negatives

SSM : Superficial Spreading Melanoma

IRM : Imagerie par Résonance Magnétique

TEP : Tomographie par Émission de Positons

GPU : Graphics Processing Unit

CPU : Central Processing Unit

RAM : Random Access Memory

Résumé :

Ce mémoire étudie l'identification du mélanome à partir d'images dermatoscopiques en utilisant des réseaux de neurones convolutifs (CNN). Trois architectures ont été testées : un CNN simple, MobileNetV2 et ResNet50, sur la base Melanoma Cancer Dataset (13 879 images). Les résultats obtenus montrent des performances limitées (précision de 48% pour ResNet50, 88.7% pour MobileNetV2 mais avec une faible AUC de 0.522), révélant des difficultés majeures liées au déséquilibre des classes, aux limites du prétraitement des images et aux contraintes matérielles. Bien que ces modèles n'atteignent pas une fiabilité suffisante pour une application clinique directe, cette analyse met en lumière les défis techniques et propose des pistes d'amélioration : rééquilibrage des données, fine-tuning des modèles pré-entraînés et intégration de méthodes de segmentation avancées. Ce travail souligne l'importance d'une approche critique dans l'application de l'IA au diagnostic médical.

Mots clés : Mélanome, CNN, Limites de L'IA, Déséquilibre de classes, Prétraitement d'images , Diagnostic médical.

Abstract :

This thesis explores melanoma detection in dermoscopic images using convolutional neural networks (CNNs). Three architectures were evaluated: a basic CNN, MobileNetV2, and ResNet50, trained on the **Melanoma Cancer Dataset** (13,879 images). The results revealed significant limitations (48% precision for ResNet50, 88.7% accuracy for MobileNetV2 but with a low AUC of 0.522), highlighting challenges such as class imbalance, insufficient image preprocessing, and hardware constraints. While the models lack reliability for direct clinical use, this study identifies key technical hurdles and suggests improvements: data rebalancing, fine-tuning of pre-trained models, and advanced segmentation techniques. The work emphasizes the need for a critical approach when applying AI to medical diagnosis.

Keywords : Melanoma, CNN, AI limitations, Class imbalance, Image processing, Medical diagnosis.

1 Introduction générale

1.1 Contexte Médical

La peau, organe principal du système tégumentaire, protège le corps et peut être atteinte de lésions bénignes ou malignes (cancéreuses). Les cancers cutanés se divisent en deux catégories : les mélanomes et les non-mélanomes (carcinomes épidermoïdes et basocellulaires). Leur développement dépend de facteurs génétiques (phototype, albinisme, vitiligo) et de l'exposition aux UV[1-[2]. La stadification de ces cancers permet d'évaluer leur progression et d'adapter le traitement, soulignant l'importance d'un diagnostic précoce pour améliorer les chances de guérison.

Le diagnostic repose d'abord sur un examen visuel analysant la morphologie des lésions (taille, forme, couleur, etc.), mais cette évaluation reste subjective en raison des variations de perception et des propriétés optiques de la peau[3]. Pour améliorer la précision, la dermoscopie permet une visualisation détaillée des structures cutanées invisibles à l'œil nu, facilitant la détection précoce des mélanomes [4]. D'autres techniques avancées (microscopie confocale, spectroscopie, imagerie multi-spectrale) offrent une analyse plus fine des tissus, mais leur complexité et leur temps d'interprétation limitent encore leur utilisation en routine clinique[4].

1.2 Problématique

Le diagnostic manuel des lésions cutanées, bien que reposant sur des méthodes éprouvées comme l'examen visuel et la dermoscopie, présente des limites significatives. La subjectivité de l'interprétation des critères morphologiques (couleur, forme, asymétrie), la ressemblance entre certaines lésions bénignes et malignes aux stades précoces, ainsi que la variabilité inter-observateur rendent le diagnostic précoce particulièrement complexe. Dans ce contexte, comment les nouvelles technologies d'imagerie et d'intelligence artificielle peuvent-elles pallier ces difficultés pour améliorer la précision et la rapidité du diagnostic des cancers de la peau ?[5]

1.3 Objectif du travail

Dans ce travail, nous avons exploité un réseau de neurones convolutif (CNN), une architecture d'apprentissage profond particulièrement efficace pour l'extraction de caractéristiques et la classification d'images, grâce à sa capacité à traiter des données visuelles avec une haute précision. Notre objectif principal était d'appliquer ces réseaux neuronaux convolutifs à l'analyse d'images dermatoscopiques, afin de développer un système automatisé pour la classification et la détection précoce du mélanome.

1.4 Organisation du mémoire

Ce mémoire est organisé de la sorte :

- Chapitre 1 (synthèse bibliographique) consacré au mélanome (aspects médicaux) et l'importance du diagnostic précoce , les bases de données pour la dermatologie, l'apprentissage profond en imagerie médicale, les CNNs et les travaux récents sur la classification des lésions cutanés
- Chapitre 2 (Matériel et Méthodologie) présente la base de données **Melanoma Cancer Dataset** utilisée, les prétraitement sur la base de données , les architectures avec outils utilisés et la méthodologie d'évaluation
- Chapitre 3 (Résultats et Discussion) est la partie expérimentale de ce mémoire suivi par une discussion des résultats et outils que nous avons utilisé dans notre travail. Le mémoire se termine avec une conclusion et quelques perspectives pour des travaux futures

2 Chapitre I : Synthèse Bibliographique

2.1 Le mélanome (aspects médicaux)

La peau, en tant qu'organe le plus exposé aux agressions extérieures, est continuellement confrontée à des risques susceptibles de provoquer des maladies et des infections cutanées. Une lésion cutanée se caractérise par une zone de peau présentant une croissance ou un aspect anormal en comparaison avec les tissus voisins. Parmi les principaux facteurs de risque, l'exposition aux rayons ultraviolets (UV) émis par le soleil joue un rôle crucial. Ces radiations peuvent entraîner des modifications cellulaires, pouvant parfois déboucher sur le développement d'un cancer de la peau, souvent reconnaissable par l'apparition d'une tache pigmentée irrégulière. Le mélanome, qui est la forme la plus agressive de cancer cutané, est responsable d'un nombre significatif de décès, en particulier parmi les populations à la peau claire. Au cours des dernières décennies, son incidence a connu une augmentation constante : il est désormais classé parmi les cancers les plus fréquents dans les pays occidentaux, avec un taux de nouveaux cas doublant tous les dix ans depuis 50 ans[6].

On distingue trois principaux types de cancer de la peau [6]:

- le mélanome qui se développe à partir des mélanocytes. il se propage de manière horizontale (sur l'épiderme) ou aussi de manière verticale (dans les couches profondes de la peau) .
- le carcinome spinocellulaire qui se développe à partir de l'épiderme. Il se propage très vite à la surface de la peau et peut aussi atteindre d'autres organes. le carcinome spinocellulaire est le deuxième cancer de la peau le plus mortel.
- le carcinome basocellulaire qui se développe à partir des cellules basales. Cette forme de cancer est la plus fréquente mais rarement métastatique et moins dangereux que les deux autres.

2.1.1 Le mélanome

2.1.1.1 Définition :

Le mélanome est une tumeur maligne des mélanocytes, les cellules responsables de la pigmentation cutanée. Dans plus de 90 % des cas, il se développe au niveau de la peau, mais peut aussi apparaître, plus rarement, sur les muqueuses, sous les ongles ou dans les yeux. Sur la peau, il peut se manifester sous diverses formes (tache marron, bouton, etc.). Représentant 3 % des cancers humains, c'est celui dont la fréquence augmente le plus, avec une incidence croissante de 10 % par an depuis 50 ans. Il touche principalement les peaux claires, et l'exposition solaire est un facteur de risque majeur. Il peut se présenter comme une lésion pigmentée noire récente sur une peau saine. Bien qu'il soit le cancer de la peau le plus mortel

(responsable de 75 % des décès liés à ce type de cancer), une détection et un traitement précoces permettent une guérison par simple excision. En revanche, sans intervention rapide, il peut s'étendre rapidement à d'autres organes et entraîner le décès [7].



Figure I.1 : Le mélanome

2.1.2 Le développement du mélanome :

Dans 80 % des cas, les mélanomes cutanés se développent sur une peau saine, sans lésion préexistante, se manifestant sous la forme d'une nouvelle tache pigmentée ressemblant à un grain de beauté. Les 20 % restants proviennent d'un nævus (grain de beauté) préexistant. À ses débuts, les cellules cancéreuses sont peu nombreuses et restent confinées à la surface de la peau (épiderme), formant ce qu'on appelle un mélanome in situ, qui s'étend initialement de manière horizontale. En l'absence de traitement, la tumeur peut progresser en profondeur, franchissant la membrane basale pour envahir le derme, ce qui marque le passage à un mélanome invasif. À ce stade, les cellules cancéreuses peuvent se propager via les vaisseaux sanguins ou lymphatiques, entraînant la formation de métastases dans d'autres organes [7].

La figure ci-dessous illustre un cas où le mélanome se développe à partir d'un grain de beauté (nævus) préexistant :



Grain de beauté

Mélanome

Figure II.2: Développement du mélanome

2.1.2.1 Les types de mélanomes :

Il existe quatre grands types des mélanomes en fonction de leur aspect clinique et leur mode progression (Figure I.3) [8]:



a) Superficiel Extensif

b) Nodulaire

c) Dubreuilh

d) Lentigineux

Figure I.3 : Types de mélanome

- Le mélanome à extension superficielle (SSM : Superficial Spreading Melanoma) [Cabarro, 2006] :

Le mélanome superficiel extensif est le type le plus courant, représentant environ 70 % de l'ensemble des mélanomes. Il se caractérise par une extension horizontale sur la surface cutanée et apparaît généralement sous la forme d'une lésion plate et mince aux contours irréguliers. Ses couleurs peuvent varier, englobant différentes nuances de rouge, bleu, brun, noir, gris et blanc. Parfois, il peut également se développer à partir d'un grain de beauté préexistant. (Figure I.3.a)

- mélanome nodulaire [Cabarro, 2006] :

Le mélanome nodulaire est le deuxième type le plus courant, représentant environ 15 à 20 % de l'ensemble des mélanomes. Contrairement aux autres formes, il se développe rapidement en profondeur dans la peau plutôt que de s'étendre à la surface. Il se manifeste généralement sous la forme d'une masse surélevée, qui est souvent noire, mais peut également être rouge, rose ou de la même couleur que la peau. Son évolution rapide en fait l'un des mélanomes les plus agressifs (Figure I.3.b).

- mélanome lentigineux des extrémités (lentigo malin) [Cabarro, 2006] :

Le mélanome lentigineux des extrémités affecte principalement les individus à la peau foncée, en particulier ceux d'origine africaine, asiatique ou hispanique. Contrairement à d'autres formes de mélanome, il n'est pas lié à l'exposition au soleil. Représentant moins de 5 % des cas, il se manifeste généralement par une petite tache plane, brun foncé ou noire, qui est souvent discrète dans ses phases initiales (Figure I.3.c).

- Le mélanome sur mélanose de Dubreuilh [Cabarro, 2006] :

Il affecte principalement les personnes âgées et représente entre 10 et 15 % de l'ensemble des mélanomes. Il se manifeste généralement par une lésion aux

contours irréguliers, de couleur ocre ou brun clair au départ, qui s'assombrit progressivement au fur et à mesure de son évolution, affichant des nuances variées de brun et de noir. Ce type de mélanome se développe souvent sur des zones régulièrement exposées au soleil, comme le visage, et évolue lentement avant de devenir invasif (Figure I.3.d).

2.1.3 L'importance du diagnostic précoce

L'incidence du mélanome chez les populations blanches vivant dans des régions ensoleillées double environ tous les dix ans. Dans la plupart des pays européens, on estime qu'elle se situe entre 5 et 10 nouveaux cas pour 100 000 habitants chaque année [Lens, 2004]. Toutefois, ce chiffre est bien plus élevé en Australie, où il atteint 40 nouveaux cas pour 100 000 habitants par an parmi les populations blanches. À l'inverse, l'incidence du mélanome demeure très faible dans les pays à majorité noire ou asiatique [Garbe, 2001].

En Algérie, une étude rétrospective menée entre 1985 et 1995 sur 116 patients atteints de mélanome a mis en évidence une incidence plus élevée dans la région ouest du pays, atteignant 53,17 %. Cette pathologie se révèle plus fréquente chez les personnes âgées de 40 à 50 ans, représentant 33,80 % des cas, et concerne davantage les femmes, qui composent 55,70 % des patients. La forme nodulaire se présente comme la plus courante (74,31 %), principalement localisée sur les membres inférieurs (55,70 %), et une majorité des cas (82,07 %) se développent de novo. La consultation tardive est également observable, comme en témoigne la proportion élevée de mélanomes au stade de Clark V, qui s'élève à 47,91 % [Boudghene, 1990].

La peau est l'organe humain le plus souvent affecté par le cancer. En vu de ce nombre et de la dangerosité de ce cancer, d'où l'intérêt d'un diagnostic précoce et un traitement chirurgical efficace pour améliorer le pronostic. Le diagnostic précoce du mélanome permet d'atteindre un taux de réussite chirurgicale curative de 80 % lorsque la tumeur est localisée. En revanche, lorsque des métastases se développent, le taux de survie à cinq ans diminue de manière significative, ne s'élevant qu'à 10 %. L'IA et l'imagerie médicale avancée facilitent l'identification automatique des mélanomes, et peuvent prédire la réponse à certains traitements.

2.2 Intelligence Artificielle, Apprentissage Profond, Imagerie Médicale, Apprentissage Profond en Imagerie Médicale

2.2.1 Intelligence Artificielle

L'être humain excelle dans la polyvalence et la prise de décision complexe, mais montre des limites dans le traitement de grandes quantités de données numériques. À l'inverse, les ordinateurs réalisent des calculs complexes rapidement mais peinent à imiter les tâches humaines simples sans modélisation préalable. Initialement conçus pour imiter le cerveau humain, les réseaux de neurones artificiels ont évolué vers des applications spécialisées, s'éloignant de leur inspiration biologique. Aujourd'hui, ils sont largement utilisés en vision par ordinateur, reconnaissance vocale, traduction automatique, réseaux sociaux, jeux et diagnostic médical [9]. L'intelligence artificielle, grâce à des algorithmes sophistiqués et au machine learning, permet désormais aux machines d'accomplir des tâches complexes autrefois réservées à l'homme [10].

Pour accomplir ses tâches, une machine doit apprendre à interpréter son environnement. C'est là qu'intervient le Machine Learning (apprentissage automatique), un ensemble de méthodes qui permettent aux systèmes d'acquérir des connaissances à partir de données d'entraînement. Grâce à ces données, ce qui permet aux algorithmes de généraliser et de prendre des décisions face à de nouvelles informations.

Parmi les différentes approches du Machine Learning, le Deep Learning (apprentissage profond) se distingue en tant que sous-domaine puissant, reposant sur des réseaux de neurones artificiels.

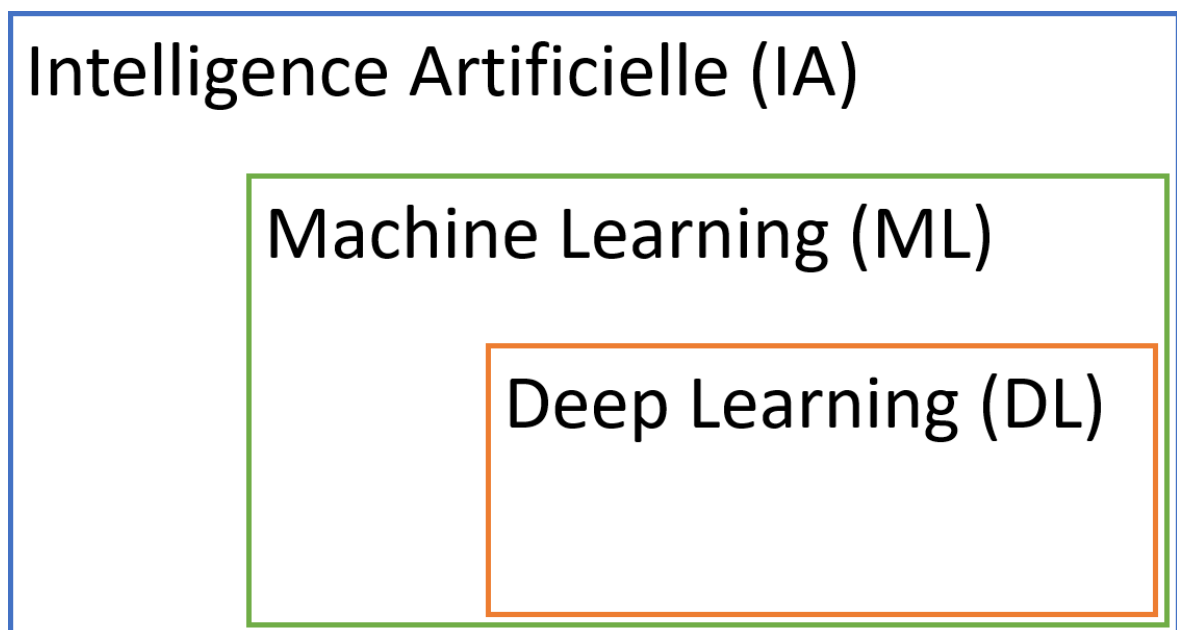


Figure I.4 : IA, ML, DP

2.2.2 Apprentissage Profond (Deep Learning)

D'après Yann LeCun [11], « L'apprentissage profond ou le Deep Learning permet aux modèles de calcul composé de plusieurs couches d'apprendre des représentations de données avec plusieurs niveaux d'abstraction ». L'apprentissage profond (DL), une sous-catégorie de l'apprentissage automatique, se concentre sur l'apprentissage de représentations à partir de données brutes. Cela permet à l'algorithme de découvrir automatiquement les motifs et caractéristiques sous-jacents nécessaires à des tâches de détection ou de classification. Contrairement aux méthodes traditionnelles, qui exigent une extraction manuelle des caractéristiques, le DL traite directement les données d'entrée (qu'il s'agisse d'images, de textes, de sons, etc.) à l'aide d'architectures de réseaux neuronaux profonds. Ces réseaux sont capables de hiérarchiser et d'optimiser ces représentations de manière autonome. Bien que ce concept puisse être appliqué à n'importe quelle technologie d'apprentissage automatique, il excelle particulièrement dans des domaines complexes tels que la vision par ordinateur, le traitement du langage naturel ou la reconnaissance vocale. Son approche end-to-end (de bout en bout) permet d'éliminer le besoin d'intervention humaine dans la sélection des caractéristiques [11].

L'apprentissage en profondeur permet à un modèle informatique d'apprendre directement à partir d'images, de textes ou de sons afin d'effectuer des tâches de classification, souvent avec une précision dépassant celle des performances humaines. Ces modèles sont formés à l'aide de vastes volumes de données étiquetées et d'architectures de réseaux neuronaux multicouches, qui extraient progressivement des caractéristiques de plus en plus complexes. Il s'emploie dans plusieurs domaines et contextes d'expertise parmi lesquelles on peut citer :

- Le diagnostic et la recherche médicale
- La détection des maladies
- La reconnaissance vocale ou faciale
- La robotique

2.2.3 Imagerie Médicale

Le diagnostic médical a été facilité par l'introduction de nouvelles technologies et principalement l'imagerie médicale et ses outils; qu'ils soient électroniques ou informatiques pour mesurer et évaluer le corps humain dans le but de détecter des maladies au stade le plus précoce que possible [18]. L'imagerie médicale englobe l'ensemble des techniques et technologies utilisées pour acquérir, traiter et interpréter des images du corps humain dans le but de poser des diagnostics, de mettre en œuvre des traitements ou de suivre l'évolution des patients. Elle inclut plusieurs modalités [12] :

- Radiographie (rayons X)
- Échographie (ondes ultrasonores)
- IRM (Imagerie par Résonance Magnétique)
- Scanner (tomodensitométrie)
- Médecine nucléaire (TEP, scintigraphie)

Ces méthodes, non invasives ou peu invasives, révèlent la structure, le fonctionnement ou les anomalies des tissus et organes, jouant un rôle clé dans la détection précoce, la planification chirurgicale et la recherche médicale.

2.2.4 Apprentissage Profond en Imagerie Médicale

L'intelligence artificielle (IA) offre des perspectives prometteuses dans le secteur de la santé, en optimisant la prise en charge des patients et en allégeant la charge de travail des médecins, tout en préservant une qualité diagnostique élevée. L'explosion des données de santé pave la voie à des applications concrètes, telles que le dépistage précoce du cancer. En raison de la marge d'erreur inhérente à l'évaluation humaine, les outils d'aide au diagnostic basés sur l'IA pourraient contribuer à réduire ces erreurs et à détecter des indices subtils en analysant d'importantes bases de données médicales. Pour traiter efficacement ces volumes de données, l'IA s'appuie sur des techniques avancées telles que le Deep Learning (apprentissage profond) et le Machine Learning (apprentissage automatique), permettant ainsi des analyses complexes et précises [13].

Les applications d'apprentissage en profondeur pour l'analyse d'images médicales ont émergé lors d'ateliers et de conférences avant de se diffuser dans les revues scientifiques. À partir de 2015 et 2016, le nombre de publications sur ce sujet a connu une croissance rapide, devenant un thème dominant dans les grandes conférences, comme en témoigne le premier numéro spécial dédié dans IEEE Transactions on Medical Imaging en mai 2016. Aujourd'hui, une conférence entièrement consacrée à ce domaine, le Workshop on Machine Learning in Medical Imaging (MLMI), se tient même en parallèle de la référence du secteur, Medical Image Computing and Computer Assisted Intervention (MICCAI), soulignant l'importance croissante de l'apprentissage automatique en imagerie médicale[15-16]. Les prestataires de soins de santé génèrent et capturent d'immenses volumes de données, incluant des signaux biologiques et des images médicales. Face à cette masse d'informations hétérogènes, l'apprentissage automatique s'impose comme une solution efficace pour intégrer, analyser et extraire des prédictions pertinentes à partir de ces vastes ensembles de données [17].

L'apprentissage profond représente une véritable révolution dans l'analyse automatique des images. Cependant, cette approche requiert une quantité considérable de données de référence, car la précision de la classification dépend largement de la qualité et de la taille de l'ensemble de données utilisées pour l'entraînement. Il est nécessaire de former un classificateur avec des milliers, voire des millions d'images annotées par des experts. Par la suite, la machine est capable de comprendre ces données par elle-même, sans nécessiter d'intervention humaine. Grâce à sa capacité d'apprentissage continu sur de nouvelles images, le système peut diagnostiquer certaines maladies avec une précision supérieure à celle de la plupart des médecins.

2.2.5 Les bases de données pour la dermatologie (focus sur Melanoma Cancer Dataset)

L'un des défis majeurs en intelligence artificielle appliquée à la dermatologie est la disponibilité de jeux de données annotés de qualité. Parmi les bases de données publiques utilisées pour l'entraînement des modèles de classification, on peut citer [19]:

- ❖ ISIC Archive
- ❖ Melanoma Cancer Dataset
- ❖ HAM10000
- ❖ DermaMNIST
- ❖ SD-198
- ❖ PAD-UFES-20
- ❖ Derm7pt Dataset

Dans notre approche, Melanoma Cancer Dataset sert de base pour l'entraînement initial des modèles, avant un affinage (fine-tuning) sur des données locales ou complémentaires. Son utilisation permet d'établir une classification hiérarchique plus précise, en combinant ses annotations avec des métadonnées cliniques pour améliorer l'interprétabilité des résultats.

Cette base, bien qu'imparfaite, reste un pilier pour le développement de solutions IA en dermatologie, notamment pour le déploiement clinique de systèmes de diagnostic assisté.

Nous allons présenter dans le chapitre 2 ses caractéristiques, ses avantages et ses limites.

2.2.6 Réseaux des Neurones

Les réseaux de neurones (RN) sont des systèmes composés de neurones (nœuds) interconnectés par des synapses qui transmettent des signaux entre eux, permettant ainsi de détecter automatiquement des relations non linéaires complexes entre de nombreuses variables, sans nécessiter d'intervention humaine [21].

Un réseau de neurones est typiquement structuré en couches successives, où chaque couche (i) reçoit ses entrées des sorties de la couche précédente. Chacune de ces couches contient N_i neurones, connectés aux N_{i-1} neurones de la couche antérieure via des poids (W) qui déterminent la force de ces connexions. Mathématiquement, cette opération équivaut à multiplier le vecteur d'entrée par une matrice de transformation, où chaque neurone effectue une combinaison linéaire pondérée de ses entrées. Ainsi, l'empilement des couches revient à enchaîner plusieurs transformations matricielles, permettant au réseau de modéliser des relations de plus en plus complexes à travers cette composition de fonctions [22].

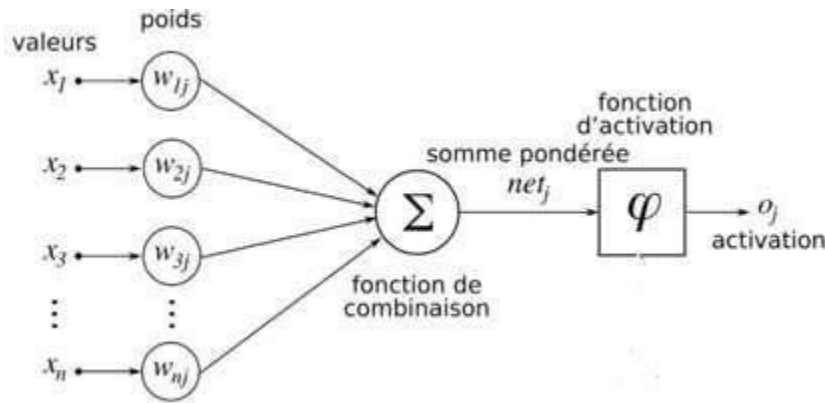


Figure I.5: Structure d'un neurone artificiel [22]

Les réseaux de neurones (RN) se déclinent en plusieurs architectures, chacune conçue pour répondre à des besoins spécifiques. Leur classification repose généralement sur des critères tels que la profondeur (nombre de couches entre l'entrée et la sortie), la complexité des couches cachées, ou encore la configuration des connexions entre neurones[23]. Dans le domaine de la détection d'objets, diverses architectures et méthodes d'apprentissage ont été expérimentées, parmi lesquelles les réseaux de neurones convolutionnels (CNN) se sont particulièrement distingués par leurs performances. C'est d'ailleurs cette architecture que nous avons adoptée comme fondement pour développer notre propre réseau détecteur.

2.2.7 Réseaux de neurones convolutionnel(CNN)

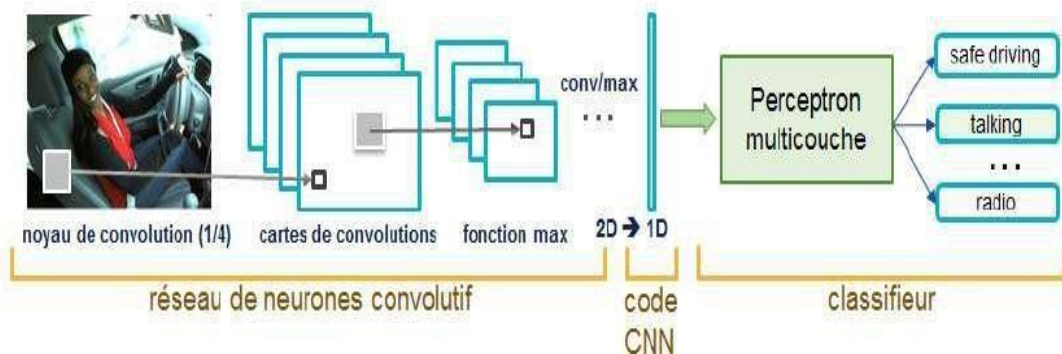


Figure I.6: Architecture standard d'un CNN

Les Réseaux de Neurones Convolutionnels (CNN, de l'anglais Convolutional Neural Network) sont actuellement les modèles les plus performants pour classer les pixels d'une image. Ils se composent de deux parties distinctes : une image en entrée, représentée sous forme de matrice (2D pour les niveaux de gris, avec une troisième dimension de profondeur 3 pour les couleurs RVB), est traitée par une partie convolutive qui extrait des caractéristiques via une série de filtres (noyaux de convolution), générant des cartes de convolution. Certains filtres réduisent la résolution par un maximum local, et les cartes résultantes sont finalement aplaties et concaténées en un vecteur de caractéristiques appelé code CNN[20]. Le code CNN

en sortie de la partie convolutive est ensuite transmis à une deuxième partie, composée de couches entièrement connectées (un perceptron multicouche), dont le rôle est de combiner ces caractéristiques pour classer les pixels ou les images elles-mêmes. La sortie finale est une couche comportant un neurone par catégorie, dont les valeurs numériques sont généralement normalisées (par exemple via une fonction softmax) pour produire une distribution de probabilité entre 0 et 1, avec une somme égale à 1, indiquant la probabilité d'appartenance à chaque classe (voire la figure I.5) .

Une architecture CNN est constituée d'un empilement de couches de traitement distinctes[24] :

- la couche de convolution (CONV) : qui applique des filtres sur des champs récepteurs locaux pour extraire des caractéristiques ;
- la couche de pooling (POOL) : qui réduit la dimension spatiale des cartes de caractéristiques par sous-échantillonnage (max-pooling ou average-pooling) ;
- la couche de correction (ReLU) : qui introduit une non-linéarité via la fonction d'activation ReLU (Rectified Linear Unit) ;
- la couche entièrement connectée (FC) : similaire à un perceptron multicouche, qui combine les caractéristiques pour la classification ;
- et enfin la fonction de perte (LOSS) : qui calcule l'erreur entre les prédictions et les véritables labels pour guider l'apprentissage.

❖ La couche de convolution (CONV) :

La couche de convolution constitue l'élément central des réseaux de neurones convolutifs. Son objectif est d'identifier la présence d'un ensemble de caractéristiques au sein des images qui lui sont présentées en entrée[26]. Pour ce faire, on effectue un filtrage par convolution : le principe consiste à faire glisser une fenêtre, représentant la caractéristique, sur l'image et à calculer le produit de convolution entre cette caractéristique et chaque section de l'image considérée, ce qui permet de générer une carte de caractéristiques (featuresmap) en sortie[26].

Dans les CNN, un filtre et une caractéristique désignent la même chose : un motif à détecter. La couche convolutive applique chaque filtre (ou noyau) à toutes les images d'entrée via un produit de convolution. Chaque filtre, appris automatiquement, cible une particularité visuelle spécifique. Le résultat est une carte d'activation par filtre, où les valeurs élevées indiquent une forte correspondance entre la zone analysée et la caractéristique recherchée. Ces cartes révèlent ainsi spatialement où et avec quelle intensité chaque motif apparaît dans l'image originale [25]. Il est nécessaire de définir certains paramètres qui influencent le calcul de la taille des images à la sortie de chaque couche (la carte des caractéristiques). Ces paramètres incluent :

- Le 'Pas': détermine le nombre de pixels dont le filtre se décale à chaque opération sur la matrice d'entrée. Un Pas 'élevé' diminue la précision spatiale mais augmente l'efficacité, tandis qu'un 'Pas' faible

préserve plus de détails au prix d'une complexité accrue. Ce paramètre influence directement la taille en sortie de la couche convolutive[27].

- 'La marge (à 0)' ou 'zéro padding' : parfois, il est commode de mettre des zéros à la frontière du volume d'entrée. La taille de ce 'zero-padding' est le troisième hyper paramètre. Cette marge permet de contrôler la dimension spatiale du volume de sortie. En particulier, il est parfois souhaitable de conserver la même surface que celle du volume d'entrée [20].
- La 'Profondeur' de la couche : fait référence au nombre de noyaux de convolution, ou au nombre de neurones associés à un même champ récepteur[20].

La figure ci-dessous représente le principe de convolution :

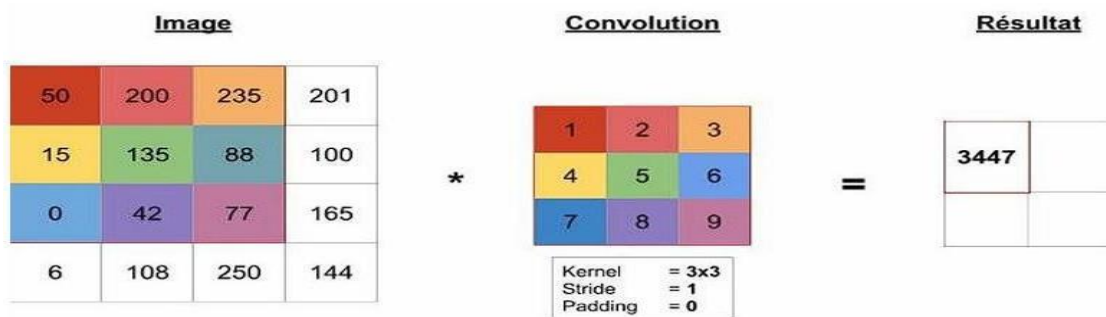


Figure I.7: Principe de la convolution [28]

❖ La couche de pooling (POOL) :

La couche de pooling, souvent insérée entre deux couches de convolution, a pour rôle de réduire la taille spatiale des cartes de caractéristiques tout en conservant leurs informations essentielles. Elle prend en entrée plusieurs cartes de caractéristiques et applique à chacune une opération de sous-échantillonnage, généralement un max-pooling (conservation de la valeur maximale par cellule) ou un average-pooling (moyenne des valeurs). En pratique, on utilise des fenêtres de petite taille (souvent 2x2 ou 3x3) avec un pas (stride) de 2, ce qui diminue la résolution spatiale de moitié tout en évitant une perte excessive d'information. Cette réduction permet de limiter le nombre de paramètres et de calculs, améliorant ainsi l'efficacité du réseau et réduisant les risques de sur-apprentissage. De plus, le pooling introduit une certaine invariance spatiale, rendant le modèle moins sensible aux variations de position ou d'orientation des caractéristiques détectées. En sortie, le nombre de cartes de caractéristiques reste identique à l'entrée, mais leur dimension est considérablement réduite, favorisant une meilleure généralisation du réseau[25].

La figure ci-dessous représente la différence entre max pooling et average pooling :

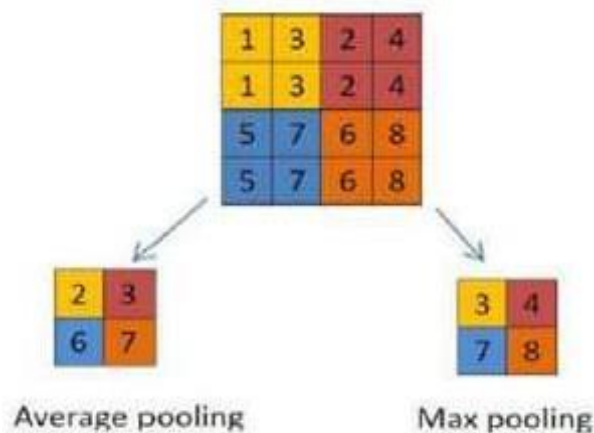


Figure I.8: Les types de Pooling [29]

❖ La couche de correction (ReLU) :

La convolution étant une opération linéaire, l'ajout d'une fonction d'activation non linéaire après chaque couche convolutive est essentiel pour permettre au réseau d'apprendre des représentations complexes. La fonction ReLU (Rectified Linear Unit), définie par $f(x) = \max(0, x)$, est la plus couramment utilisée : elle remplace toutes les valeurs négatives par zéro tout en conservant les positives intactes. Cette non-linéarité simple mais efficace améliore significativement les performances du réseau en introduisant des transformations non linéaires sans compliquer excessivement les calculs. Contrairement aux fonctions saturantes (comme sigmoïde ou tanh), ReLU évite le problème du gradient vanishing, accélère la convergence et préserve la sensibilité des neurones aux caractéristiques pertinentes, tout en maintenant l'efficacité computationnelle. Son implémentation légère et son effet régularisateur (sparsification des activations) en font un choix privilégié dans les architectures CNN modernes[22].

❖ La couche entièrement connectée (FC) :

La couche entièrement connectée (FC), présente en fin de réseau (CNN ou autre), transforme un vecteur d'entrée en un vecteur de sortie via une combinaison linéaire pondérée suivie d'une fonction d'activation. Dans un CNN appliqué à la segmentation d'images, sa dernière couche FC produit un vecteur de taille N (nombre de classes), où chaque valeur représente la probabilité d'appartenance d'un pixel à une classe donnée. Par exemple, pour une segmentation binaire ($N=2$), un vecteur $[0.9, 0.1]$ indique 90% de probabilité pour la classe 1. Cette couche opère par multiplication matricielle entre les poids appris (matrice de connexions) et les caractéristiques d'entrée, suivie d'une normalisation (softmax pour $N>2$, sigmoïde pour $N=2$) pour obtenir des probabilités. Son nom "entièrement connectée" vient de ce que chaque neurone d'entrée influence tous les neurones de sortie, permettant une décision globale basée sur l'ensemble des caractéristiques extraites par le réseau [25].

❖ La fonction de perte (Loss) :

La fonction de perte, située en fin de réseau, mesure et pénalise l'écart entre les prédictions du modèle et les valeurs réelles du jeu de données d'entraînement. Elle joue un rôle crucial dans l'apprentissage en quantifiant l'erreur à minimiser. Les fonctions de perte les plus courantes incluent l'entropie croisée (particulièrement adaptée aux problèmes de classification, car elle compare la distribution de probabilité en sortie avec la vérité terrain) et l'erreur quadratique moyenne (souvent utilisée en régression pour mesurer la différence entre valeurs prédites et observées). Le choix de la fonction de perte influence directement la façon dont le réseau ajuste ses paramètres lors de la rétropropagation du gradient [20].

❖ La couche de normalisation :

La normalisation des entrées (centrage-réduction pour une moyenne nulle et un écart-type unitaire) est une pratique établie pour accélérer la convergence des réseaux de neurones. Étendant ce principe à toutes les couches, la batch normalisation standardise dynamiquement les activations entre les couches pendant l'entraînement. Positionnée entre deux couches successives, cette technique normalise en temps réel les sorties de chaque couche (en les recentrant et les re-échelonnant par batch) avant leur propagation vers la couche suivante. Elle stabilise et accélère l'apprentissage en réduisant les variations internes (covariate shift), tout en agissant comme un régularisateur léger. Cette approche permet notamment d'utiliser des taux d'apprentissage plus élevés et atténue les problèmes d'initialisation des poids [10].

❖ La couche Dropout

Le Dropout est une technique de régularisation qui consiste à désactiver aléatoirement une proportion définie de neurones à chaque itération d'entraînement. Pendant la phase de propagation avant, les neurones sélectionnés (par tirage aléatoire) voient leur activation forcée à zéro, ce qui revient à les exclure temporairement du réseau. Cette suppression aléatoire empêche les mises à jour des poids associés lors de la rétropropagation, brisant ainsi les dépendances fortuites entre neurones et évitant la co-adaptation excessive. Le Dropout force ainsi le réseau à apprendre des caractéristiques redondantes et robustes, améliorant sa capacité de généralisation en réduisant le sur-apprentissage. En test, tous les neurones sont réactivés, mais leurs poids sont pondérés par le taux de Dropout pour compenser l'augmentation artificielle de l'activité.

2.2.8 Les architectures populaires de CNN(LeNet, AlexNet, ResNet, MobileNet,...)

La figure I.7 illustre l'apparition des différentes méthodes en fonction des années

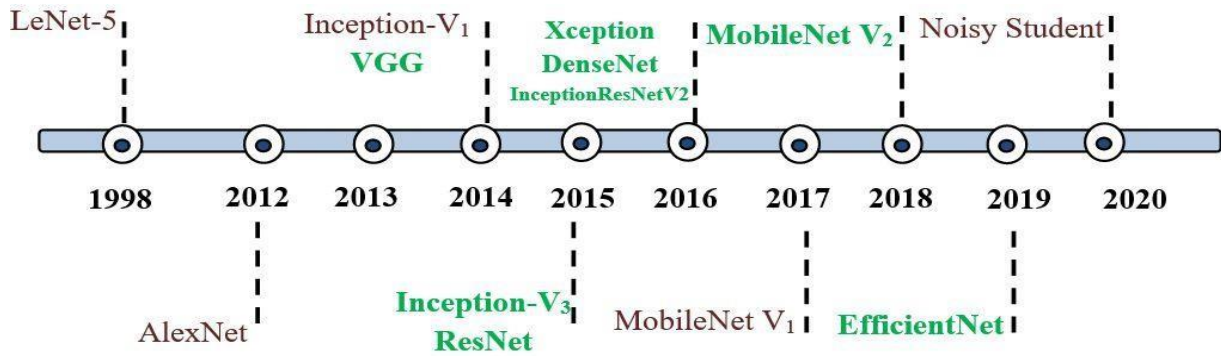


Figure I.9: Histoire évolutive des CNNs montrant les innovations architecturales [30]

Plusieurs architectures CNNs existent, les plus populaires sont :

- **LeNet**

L'architecture LeNet, version la plus efficace et célèbre développée par Yann LeCun, a marqué l'histoire comme le premier véritable réseau de neurones convolutifs (CNN) opérationnel. Conçue initialement pour la reconnaissance de chiffres manuscrits et de codes postaux, cette architecture pionnière combine intelligemment 4 couches alternées de convolution (CONV) et de pooling pour l'extraction hiérarchique de caractéristiques, suivies de 3 couches entièrement connectées pour la classification. Sa grande innovation réside dans sa capacité à apprendre automatiquement des caractéristiques pertinentes directement à partir des pixels bruts, éliminant le besoin d'une extraction manuelle de features. Par son utilisation astucieuse des opérations convolutives et de sous-échantillonnage, LeNet parvient à réduire drastiquement le nombre de paramètres tout en conservant une excellente capacité de reconnaissance, posant ainsi les bases des CNN modernes. Son succès pratique sur des tâches réelles de vision par ordinateur a démontré pour la première fois le potentiel des approches d'apprentissage profond [30].

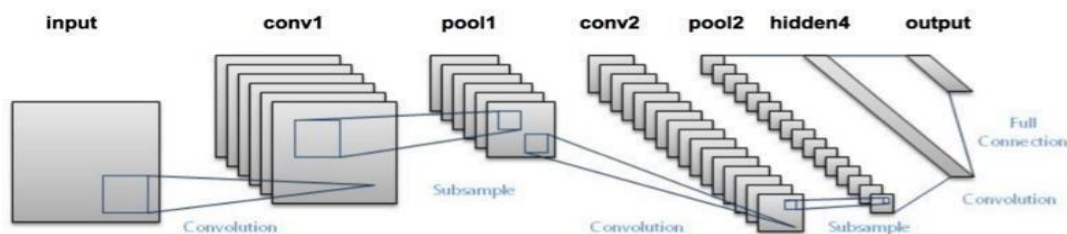


Figure I.10: L'architecture LeNet[31]

- **AlexNet**

AlexNet, développé par Alex Krizhevsky, Ilya Sutskever et Geoff Hinton, est la première architecture CNN à avoir marqué l'histoire du deep learning en révolutionnant la vision par ordinateur. Présenté lors du challenge ImageNet ILSVRC 2012, ce modèle a surpassé de loin tous ses concurrents, établissant un nouveau standard de performance. Son architecture pionnière combine 5 couches convolutives utilisant des ReLU (introduites comme innovation majeure pour accélérer l'apprentissage), 3 couches de max-pooling pour la réduction dimensionnelle, et se termine par 3 couches fully-connected pour la classification. AlexNet a non seulement démontré l'efficacité des CNN sur des tâches complexes, mais a aussi popularisé des techniques clés comme l'utilisation des GPU pour l'entraînement et le dropout pour la régularisation.

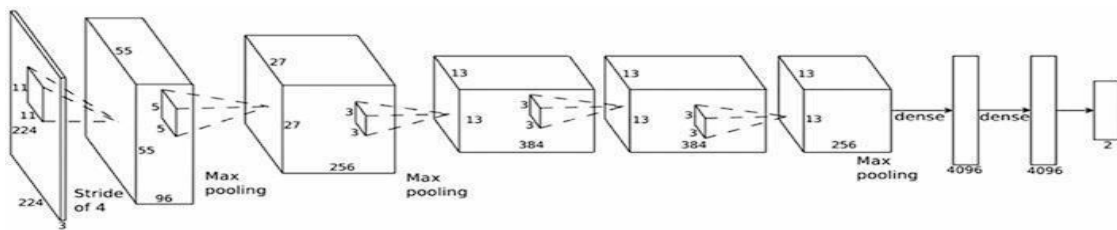


Figure I.11: L'architecture AlexNet [32]

Le résumé de l'architecture AlexNet dans le tableau ci-dessous :

Layer		Feature Map	Size	Kernel Size	Stride	Activation
Input	Image	1	227x227x3	-	-	-
1	Convolution	96	55 x 55 x 96	11x11	4	relu
	Max Pooling	96	27 x 27 x 96	3x3	2	relu
2	Convolution	256	27 x 27 x 256	5x5	1	relu
	Max Pooling	256	13 x 13 x 256	3x3	2	relu
3	Convolution	384	13 x 13 x 384	3x3	1	relu
4	Convolution	384	13 x 13 x 384	3x3	1	relu
5	Convolution	256	13 x 13 x 256	3x3	1	relu
	Max Pooling	256	6 x 6 x 256	3x3	2	relu
6	FC	-	9216	-	-	relu
7	FC	-	4096	-	-	relu
8	FC	-	4096	-	-	relu
Output	FC	-	1000	-	-	Softmax

Tableau I.1 : Résumé de l'architecture AlexNet[33]

- **VGGNet**

Le modèle VGG, développé par le Visual Geometry Group d'Oxford sous la direction d'Andrea Vedaldi et Andrew Zisserman, a démontré l'importance cruciale de la profondeur du réseau pour atteindre des performances élevées en vision par ordinateur. Cette architecture se distingue par son empilement systématique de couches convolutives 3x3, permettant une augmentation progressive de la complexité des caractéristiques apprises. Cependant, cette

profondeur accrue se paie au prix d'une complexité computationnelle élevée : avec ses 140 millions de paramètres, le réseau VGG exige des ressources importantes en mémoire et en puissance de calcul, ce qui en fait un modèle relativement coûteux à entraîner et déployer, malgré ses excellentes capacités de représentation visuelle.

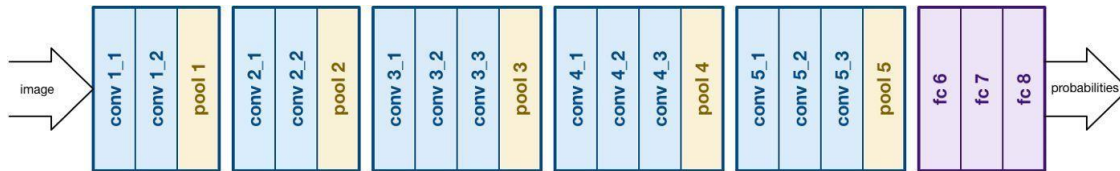


Figure I.12: L'architecture du VGGNet [34]

Le résumé de l'architecture VGGNet dans le tableau ci-dessous :

	Layer	Feature Map	Size	Kernel Size	Stride	Activation
Input	Image	1	224 x 224 x 3	-	-	-
1	2 X Convolution	64	224 x 224 x 64	3x3	1	relu
	Max Pooling	64	112 x 112 x 64	3x3	2	relu
3	2 X Convolution	128	112 x 112 x 128	3x3	1	relu
	Max Pooling	128	56 x 56 x 128	3x3	2	relu
5	2 X Convolution	256	56 x 56 x 256	3x3	1	relu
	Max Pooling	256	28 x 28 x 256	3x3	2	relu
7	3 X Convolution	512	28 x 28 x 512	3x3	1	relu
	Max Pooling	512	14 x 14 x 512	3x3	2	relu
10	3 X Convolution	512	14 x 14 x 512	3x3	1	relu
	Max Pooling	512	7 x 7 x 512	3x3	2	relu
13	FC	-	25088	-	-	relu
14	FC	-	4096	-	-	relu
15	FC	-	4096	-	-	relu
Output	FC	-	1000	-	-	Softmax

Tableau I.2: Résumé de l'architecture VGGNet [35]

- **ResNet**

ResNet (Réseau Résiduel) représente une avancée majeure en apprentissage profond, redéfinissant les limites de profondeur des architectures neuronales. Son innovation clé réside dans ses modules résiduels, des blocs de construction qui introduisent des connexions skip (ou "bypass") permettant à l'information de circuler directement à travers plusieurs couches. Ces connexions résiduelles résolvent le problème crucial de la disparition du gradient dans les réseaux très profonds, autorisant ainsi l'empilement de centaines voire milliers de couches tout en maintenant une convergence stable. L'architecture ResNet, construite par l'assemblage répété de ces modules intelligents, a établi de nouveaux records de performance tout en devenant une pierre angulaire des modèles modernes de vision par ordinateur [36].

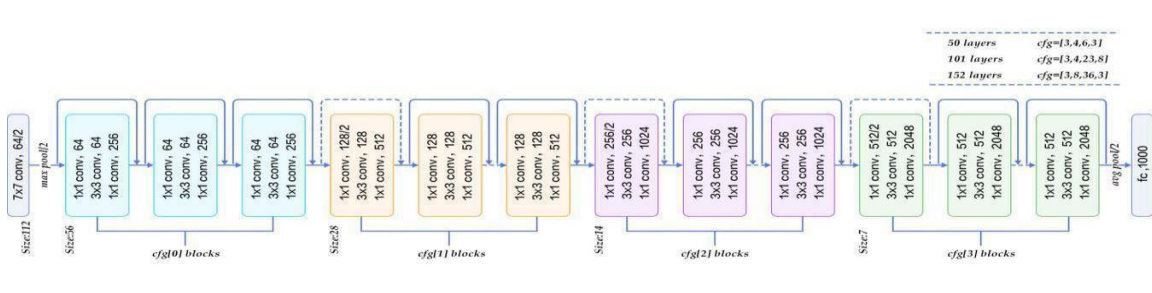


Figure I.13: L'architecture de ResNet[36]

• MobileNet

MobileNet est une architecture de réseau de neurones convolutifs (CNN) optimisée pour les dispositifs mobiles et embarqués, offrant un équilibre entre performance et efficacité computationnelle grâce à son utilisation innovante de convolutions séparables en profondeur. Ces convolutions divisent le processus traditionnel en deux étapes : une convolution spatiale indépendante par canal (depthwise) suivie d'une convolution 1×1 (pointwise) pour fusionner les caractéristiques, réduisant ainsi drastiquement le nombre de paramètres et les coûts de calcul. MobileNet introduit également des hyperparamètres ajustables (comme le width multiplier pour contrôler la largeur du réseau et un facteur de résolution pour adapter la taille des entrées), permettant une personnalisation fine en fonction des contraintes matérielles. Les versions ultérieures (MobileNetV2/V3) améliorent encore ses performances avec des blocs résiduels inversés (inverted residuals) et des fonctions d'activation comme h-swish, tout en maintenant une faible empreinte mémoire. Idéal pour les applications temps réel (reconnaissance d'objets, segmentation, etc.) sur smartphones ou appareils IoT, MobileNet allie rapidité d'inférence et précision, même avec des ressources limitées [37].

Table 1. MobileNet Body Architecture

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
5x	Conv dw / s1	$3 \times 3 \times 512$ dw
	Conv / s1	$1 \times 1 \times 512 \times 512$
	Conv dw / s2	$3 \times 3 \times 512$ dw
	Conv / s1	$1 \times 1 \times 512 \times 1024$
	Conv dw / s2	$3 \times 3 \times 1024$ dw
	Conv / s1	$1 \times 1 \times 1024 \times 1024$
	Avg Pool / s1	Pool 7×7
	FC / s1	1024×1000
	Softmax / s1	Classifier

Figure I.14: L'architecture du MobileNet [37]

- **GoogleNet**

GoogleNet (ou Inception v1), développé par Christian Szegedy et son équipe chez Google, a remporté le challenge ILSVRC 2014 en introduisant une architecture révolutionnaire basée sur des modules Inception. Ces blocs innovants exploitent des convolutions parallèles avec des noyaux de différentes tailles (1×1 , 3×3 , 5×5) pour capturer des motifs à multiples échelles spatiales, dont les sorties sont concaténées pour une représentation riche. Le réseau se distingue également par son remplacement des couches fully-connected traditionnelles par un pooling moyen global, réduisant drastiquement les paramètres de 60 millions à seulement 4 millions. Cette optimisation, combinée à l'utilisation stratégique de convolutions 1×1 pour la réduction de dimensionnalité, permet à GoogleNet d'allier efficacité computationnelle et performances élevées, tout en minimisant les risques de sur-apprentissage. Son succès a popularisé les architectures multi-branches et inspiré une série de modèles Inception ultérieurs [30].

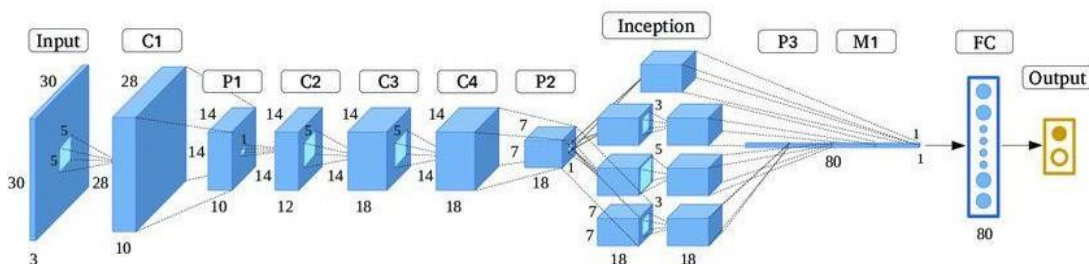


Figure I.15: L'architecture de GooleNet

2.2.9 Travaux récents sur la classification des lésions cutanées

L'introduction de la dermoscopie a marqué un progrès significatif dans le diagnostic précoce du mélanome, grâce à son éclairage homogène qui améliore le contraste entre la lésion et son arrière-plan. Cependant, son utilisation reste limitée en pratique clinique, principalement en raison d'un manque de formation des médecins, qui privilégient souvent l'examen à l'œil nu, ainsi que des inégalités d'accès aux soins dans certaines régions, où les dermatologues sont trop peu nombreux ou mal répartis. Face à ces défis, les experts plaident pour le développement de systèmes d'aide au diagnostic, tels que des applications mobiles ou des outils informatiques basés sur le traitement d'images, afin de démocratiser l'accès à un dépistage précoce et fiable. Ces solutions automatisées, s'appuyant sur des algorithmes avancés, pourraient combler les lacunes du système de santé en permettant une analyse rapide et accessible des lésions cutanées, même en l'absence de spécialistes [38].

a) **Applications mobiles :**

Pour améliorer le dépistage précoce des mélanomes, plusieurs applications mobiles Android ont été développées, servant à la fois d'outils d'aide au diagnostic et de pont entre les patients et les médecins. Parmi celles-ci, on trouve des solutions comme SkinVision, Miiskin ou

DermEngine, qui utilisent des algorithmes d'analyse d'images et parfois l'IA pour évaluer les lésions cutanées à partir de photos prises par smartphone. Certaines applications comparent les lésions à des bases de données médicales, tandis que d'autres suivent leur évolution dans le temps, alertant l'utilisateur en cas de changement suspect. Bien que ces outils ne remplacent pas un diagnostic médical complet, ils facilitent une première évaluation et encouragent une consultation précoce chez un spécialiste, notamment dans les régions où l'accès aux dermatologues est limité. Leur principe repose souvent sur une combinaison de traitement d'image, de machine learning et, pour les plus avancées, de réseaux neuronaux convolutifs (CNN) spécialisés en dermatologie.

➤ FotoVision

Développée en 2014 par les dermatologues Sergio Pedro Jaen et Vaño en collaboration avec WakeAppHealth et ISDIN, l'application FotoSkin vise à détecter précocement les lésions cutanées cancéreuses. Elle permet de différencier les tumeurs bénignes des tumeurs suspectes en analysant les changements de couleur, de forme des grains de beauté ou l'apparition de nouvelles taches, tout en appliquant la règle ABCD (Asymétrie, Bordure, Couleur, Diamètre) pour guider l'évaluation. Gratuite sur Android, l'application propose cinq sections clés : Ma Peau (suivi des lésions), Mon Conseil (recommandations), Je Commande (achat de produits), Mon Environnement (facteurs de risque) et Informations Médicales (ressources éducatives). FotoSkin sert ainsi d'outil d'alerte pour les utilisateurs et d'aide à la décision pour les dermatologues, facilitant un diagnostic rapide sans remplacer une consultation spécialisée [39].



Figure I.16: Application FotoSkin[39]

➤ Doctor Mole

Développée en 2012 par Mark Shippen, l'application Doctor Mole utilise la réalité augmentée pour analyser les grains de beauté et détecter d'éventuelles anomalies de forme, de couleur ou de texture. Après la capture photo d'une lésion, l'application attribue un niveau de risque

(indiqué par un code couleur) pour chaque caractéristique examinée. Elle offre également une fonctionnalité d'archivage et de comparaison des données, permettant de suivre l'évolution des grains de beauté dans le temps - un aspect crucial puisque leur modification peut signaler un mélanome malin. Bien qu'utile pour un premier dépistage, Doctor Mole ne remplace pas un diagnostic médical professionnel, mais sert plutôt d'outil préliminaire pour alerter les utilisateurs sur des changements cutanés potentiellement préoccupants [40].



Figure I.17: Application Doctor Mole[40]

➤ SkinVision

SkinVision utilise un algorithme d'évaluation des risques innovant combinant l'analyse d'images en niveaux de gris avec la théorie fractale, plus précisément basée sur le calcul de dimensions fractales locales. Testé sur un ensemble de lésions mélanocytaires photographiées avec un iPhone 4S (équipé d'un appareil photo autofocus 8 mégapixels), cet algorithme a démontré des performances cliniquement pertinentes : une sensibilité de 73% (détection correcte des lésions malignes), une spécificité de 83% (rejet approprié des lésions bénignes) et une précision globale de 81%. Cette approche mathématique originale permet d'objectiver l'évaluation des motifs cutanés complexes, bien que ces résultats indiquent qu'un examen dermatologique complémentaire reste nécessaire pour les cas limites [41].



Figure I.18: Application SkinVision[41]

b) Systèmes associés par ordinateur pour l'aide au diagnostic

L'aspect visuel des lésions cutanées rend leur diagnostic difficile, pouvant conduire à confondre tumeurs malignes et bénignes. D'où l'utilité des systèmes informatiques exploitant le traitement d'image avancé pour les analyser et les classer avec précision [42]. Le diagnostic assisté par ordinateur (DAO) du mélanome améliore la précision et la reproductibilité des diagnostics en fournissant aux dermatologues, notamment aux moins expérimentés, une analyse objective des images dermoscopiques. Ce système permet non seulement une détection précoce des lésions malignes - augmentant ainsi les chances de guérison et favorisant des traitements plus adaptés - mais aussi une réduction des erreurs humaines, des coûts et des délais d'interprétation, optimisant ainsi la prise en charge d'un grand nombre de patients [43].

3 Chapitre II : Données, Outils et Méthodologie

3.1 Présentation de la base de données Melanoma Cancer Dataset

La base de données utilisée dans ce travail est Melanoma Cancer Dataset sur Kaggle. Cet ensemble de données, composé de 13 879 images de taille 224x224 soigneusement sélectionnées, constitue une ressource précieuse pour faire progresser la dermatologie et le diagnostic assisté par ordinateur. Les images sont divisées en deux parties, 11 878 images pour l'entraînement (6 289 images de non-mélanomes et 5 590 images de mélanomes) et 2 000 images pour le test des architectures (1 000 images de non-mélanomes et 1 000 images de mélanomes).

3.2 Prétraitement sur la base de données

Le prétraitement consiste à modifier les images de notre base de données pour les améliorer, tandis que l'augmentation des données vise à accroître leur nombre en appliquant des transformations aléatoires, évitant ainsi que le modèle ne voie deux fois la même image. Cette technique réduit le risque de sur-apprentissage et améliore la généralisation du modèle. Dans Keras, cela est réalisé grâce à la classe `ImageDataGenerator`, un outil fourni par TensorFlow et Keras qui permet une augmentation des données en temps réel pendant l'entraînement. Plutôt que de stocker les images transformées, `ImageDataGenerator` applique des modifications aléatoires juste avant leur utilisation, tout en permettant de configurer les transformations et la normalisation souhaitées. Il génère ensuite des lots d'images augmentées (et leurs étiquettes) pour l'apprentissage. Les valeurs des pixels sont normalisées pour accélérer la convergence des modèles (valeurs initiales entre 0 et 255).

3.3 Présentation des architectures utilisées

Dans notre travail, on a développé trois modèles de réseaux de neurones CNN pour classer les différentes images de notre base de données Melanoma Cancer Dataset:

- **Un CNN personnalisé** (simple)
- **MobileNetV2** (transfer learning)
- **ResNet50** (transfer learning)

Parmi les modèles que nous avons proposés, deux modèles dépendent de l'apprentissage par transfert (Transfer Learning).

3.3.1 Modèle CNN personnalisé

Le modèle CNN personnalisé développé dans ce travail est un réseau de neurones convolutif conçu spécifiquement pour la classification des lésions cutanées. L'architecture se compose de trois blocs principaux **Conv2D** + **MaxPooling2D** + **Dropout**, suivis de couches entièrement connectées.

3 couches convolutives (Conv2D) avec des filtres de tailles 32, 64, 128 utilisant une fonction d'activation **ReLU**. Chaque couche est suivie d'un **MaxPooling2D(2,2)** pour réduire la

dimension spatiale et d'un Dropout (30%) pour limiter le surapprentissage, Une couche **Flatten** pour aplatir les données et deux couches **Dense** (512 et 1 neurones) avec une activation **sigmoid** en sortie pour prédire la classe (mélanome ou non-mélanome).

3.3.2 MobileNetV2

Le modèle **MobileNetV2** est un réseau pré-entraîné sur **ImageNet**, optimisé pour une efficacité computationnelle tout en maintenant de bonnes performances. Les principales étapes sont les suivantes :

On a utilisé des poids pré-entraînés (**ImageNet**), procédé avec **gel des couches** pour conserver les caractéristiques apprises. On a ajouté une **Global Average Pooling** pour réduire la dimensionnalité, une couche **Dense(1024)** avec **ReLU**, un **Dropout** (50%) pour la régularisation et une couche de sortie **Dense(1, sigmoid)** pour la binaire.

3.3.3 ResNet50

Le modèle **ResNet50** est une architecture profonde (50 couches) utilisant des **connexions résiduelles** (skip connections) pour faciliter l'apprentissage des features hiérarchiques. On a effectué le processus que **MobileNetV2** pour l'extraction des caractéristiques et la classification.

3.4 Environnement de travail

Cette section présente les matériels et l'environnement utilisés pour notre projet.

a) Matériel

Pour la réalisation du projet , Nous avons utilisé un ordinateur avec les caractéristiques suivantes :

- Processeur : Intel(R) Celeron(R) CPU N3060 @ 1.60GHz
- Mémoire (RAM) : 4Go
- Système d'exploitation : Microsoft Windows 10 (64 bits).

b) Langage de programmation(Python)

Python est un langage de programmation open source, gratuit et multiplateforme, privilégiant la lisibilité et la simplicité d'utilisation. Python est fonctionnel sur les systèmes Mac, Windows et Unix.

c) La plateforme Kaggle

Kaggle est une plateforme web interactive qui propose des compétitions d'apprentissage automatique en science des données. La plateforme fournit des bases de données, des notebooks et des didacticiels gratuits dont les scientifiques de données ont besoin pour réaliser leurs projets d'apprentissage automatique [44].

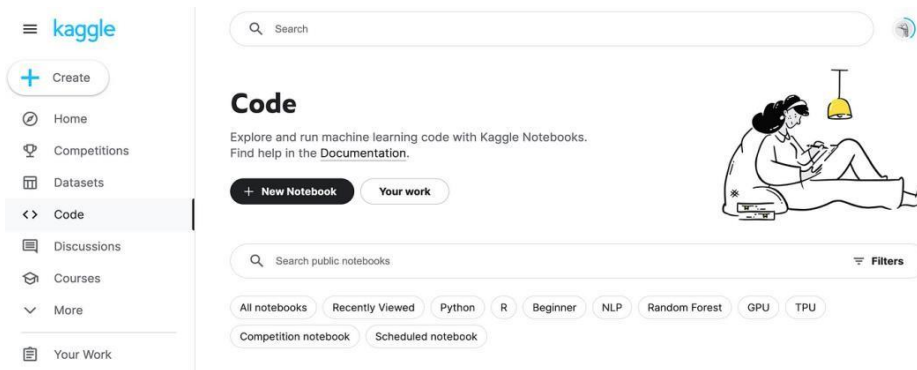


Figure II.1: Interface de Kaggle

d) Les bibliothèques

- **Tensorflow** : TensorFlow, framework open source développé par Google en 2015, est une plateforme puissante pour le deep learning basée sur des tenseurs (tableaux multidimensionnels). Très populaire dans les milieux académiques et industriels, il combine flexibilité et évolutivité pour la recherche avancée et le déploiement de modèles en production [45].
- **Keras** : Keras est une API haut niveau pour le deep learning, conçue pour TensorFlow, qui simplifie le prototypage rapide et le déploiement de modèles complexes grâce à une syntaxe intuitive et modulaire. Ses blocs configurables permettent de créer facilement des architectures avancées et des couches personnalisées avec un minimum de code, accélérant ainsi la recherche et la mise en production [45].
- **NumPy** : NumPy est une bibliothèque Python essentielle pour le calcul scientifique, offrant des opérations performantes sur des tableaux multidimensionnels et des fonctions avancées en algèbre linéaire et traitement numérique. Devenue la base de l'écosystème scientifique Python, elle alimente de nombreuses autres bibliothèques comme Pandas et Scikit-learn grâce à sa rapidité et sa simplicité d'utilisation.
- **Pandas** : Pandas est une bibliothèque Python open source offrant deux structures principales - les DataFrames pour les données tabulaires et les Series pour les séries temporelles - permettant une manipulation et analyse efficace des données. Avec ses fonctionnalités avancées de nettoyage, transformation et visualisation, elle constitue l'outil indispensable des data scientists pour la préparation des données avant analyse ou modélisation.

3.5 Métriques et protocole d'évaluation

L'évaluation d'un modèle de deep learning (DL) pour la classification du mélanome à partir de notre base de données (Melanoma Cancer Dataset) nécessite une méthodologie rigoureuse pour garantir des résultats fiables. Sur ce, on a séparé les données (train_generator et test_generator) en stratifiant pour conserver la distribution des classes (mélanome vs non-mélanome). Pour un problème de cette dernière, les métriques clés sont :

-Accuracy : une métrique de base qui mesure le taux de prédictions correctes par rapport au total des prédictions.

-**Précision** : Éviter les faux positifs (diagnostiquer un mélanome à tort).

-**Sensibilité (Recall)** : capacité à détecter tous les vrais mélanomes (éviter les faux négatifs).

-**F1-Score** : Moyenne harmonique de la précision et du rappel, utile pour les classes déséquilibrées

$$\begin{aligned} \text{Accuracy} &= \frac{TP + TN}{TP + TN + FP + FN} \\ \text{Precision} &= \frac{TP}{TP + FP} \\ \text{Recall} &= \frac{TP}{TP + FN} \\ F_1 &= 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned}$$

Figure II.2: Les formules de métriques clés utilisés

- TP (True Positives) : Mélanomes correctement détectés.
- TN (True Negatives) : Lésions bénignes correctement classées.
- FP (False Positives) : Faux diagnostics de mélanome (bénignes classées comme malignes).
- FN (False Negatives) : Mélanomes manqués (classés à tort comme bénins).

-Matrice de confusion

La matrice de confusion est un tableau qui résume les prédictions d'un modèle par rapport aux vraies étiquettes (ground truth).

		True Class	
		Positive	Negative
Predicted Class	Positive	TP	FP
	Negative	FN	TN

Figure II.3: Matrice de confusion

4 Chapitre III : Résultats et Discussion

Ce chapitre présente les résultats de notre travail basé sur l'utilisation des réseaux convolutifs avec les architectures citées précédemment.

4.1 Résultats de l'entraînement

Afin de montrer les résultats obtenus, nous présentons ci-dessous les résultats en termes de précision et d'erreur d'apprentissage.

a. Le taux de précision et de perte

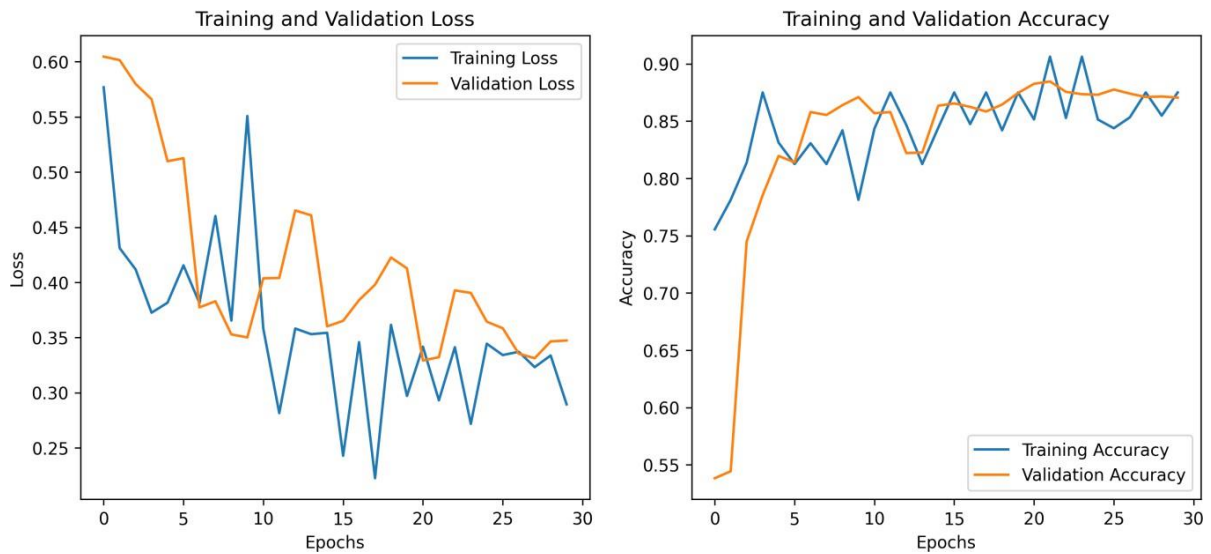


Figure III.1: CNN

- À la fin de l'entraînement, le modèle CNN personnalisé atteint une accuracy de 90 % sur l'ensemble d'entraînement et environ 88 % sur l'ensemble de validation, ce qui montre une bonne performance générale. La fonction de perte à l'entraînement diminue jusqu'à environ 0.15–0.30, tandis que le validation loss se stabilise autour de 0.33–0.36, indiquant une bonne généralisation sans surapprentissage notable.

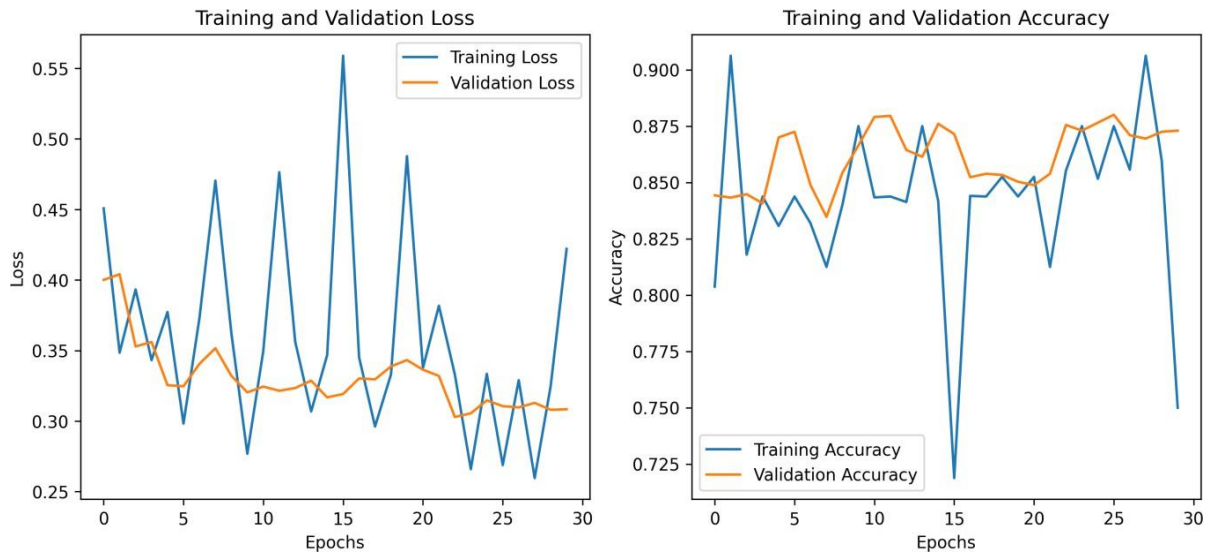


Figure III.2: Mobilenet

- Le modèle MobileNet atteint une accuracy de validation stable autour de 84 % à 87 %, tandis que l'accuracy d'entraînement varie entre 83 % et 90 %, avec quelques chutes ponctuelles. Le training loss est instable, oscillant fortement entre 0.30 et 0.55, tandis que le validation loss reste stable autour de 0.32 à 0.36. Malgré les fluctuations du loss, la performance en validation est bonne, indiquant une bonne capacité de généralisation.

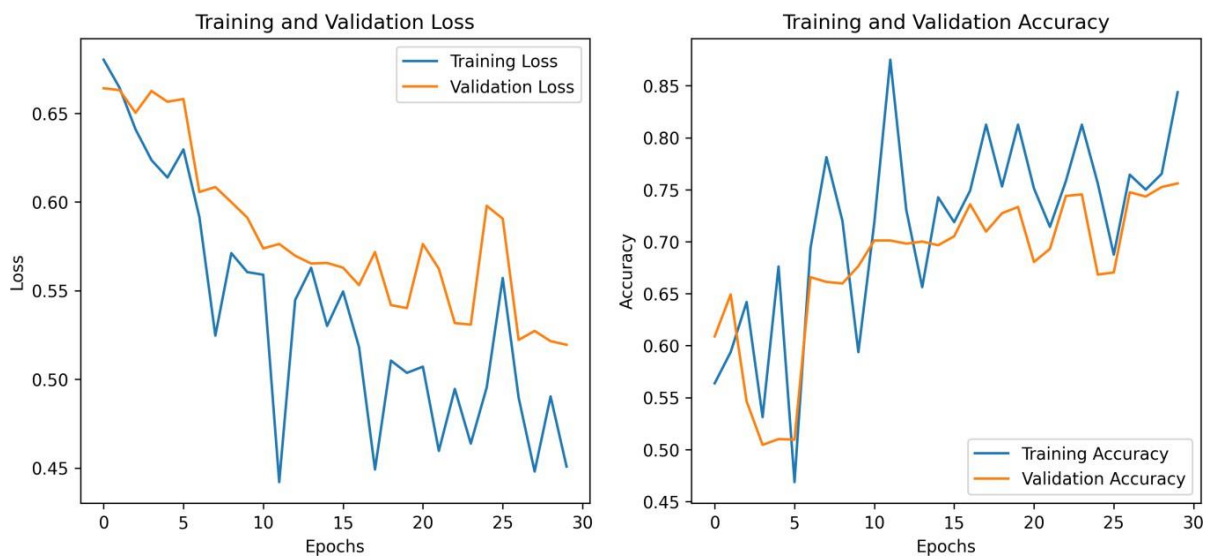


Figure III.3: Resnet

- Le modèle ResNet montre une progression régulière de l'accuracy, atteignant 85 % en entraînement et environ 75 % en validation à la fin. Le training loss diminue de 0.66 à 0.45, et le validation loss de 0.66 à environ 0.52, avec quelques variations. Le modèle généralise moins bien que MobileNet, mais reste stable et améliore ses performances au fil des epochs.

b. Matrice de Confusion

Les matrices de confusion obtenues après l'expérimentation sont présentées dans la figure ci-dessous :

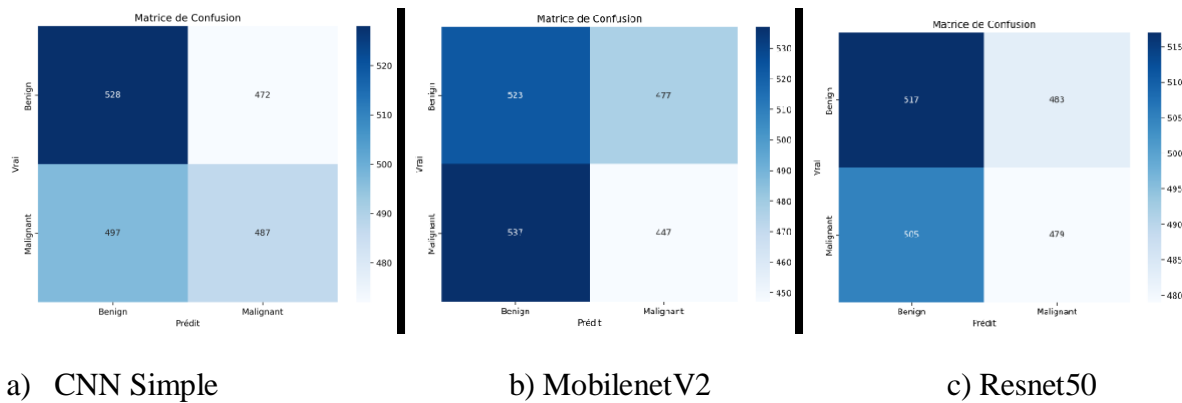


Figure III.4: Matrices de confusion

c. Tableau comparatif des performances de chaque modèle : Accuracy, Precision, Recall, F1-Score, AUC

Modèle	Accuracy	Precision	Recall	F1-Score	AUC	Specificity	False Negatives
CNN Simple	74.65	51.03	47.2	49.04	51.16	54.7	528
MobileNetV2	88.7	50.43	47.4	48.87	50.17	53.4	526
ResNet50	74.65	51.89	48.0	49.87	50.87	55.5	520

Tableau III.1: Tableau de comparaison des performances

4.2 Discussion sur les résultats obtenus

Les résultats obtenus révèlent des performances contrastées entre les différentes architectures testées. MobileNetV2 se distingue par une accuracy élevée (88.7%), nettement supérieure à celle du CNN Simple et de ResNet50 (74.65% chacun). Cependant, cette performance apparente doit être nuancée par l'analyse d'autres métriques clés. En effet, l'AUC (Area Under Curve) de MobileNetV2 (0.522), bien que légèrement meilleure que celle des autres modèles, reste proche de 0.5, ce qui indique une capacité de discrimination à peine supérieure au hasard. Cette contradiction entre accuracy et AUC suggère que MobileNetV2 pourrait souffrir de surapprentissage ou que les données présentent des biais non pris en compte. Les scores F1 similaires et faibles (autour de 49%) pour les trois architectures confirment cette difficulté à correctement classer les deux catégories.

Plusieurs facteurs peuvent expliquer ces résultats mitigés. Tout d'abord, le déséquilibre potentiel entre les classes n'a probablement pas été suffisamment corrigé, comme en témoigne le nombre élevé de faux négatifs (entre 520 et 528 selon les modèles). Ensuite, les prétraitements appliqués aux images pourraient être insuffisants ou inadaptés. L'absence d'augmentation de données ou de normalisation poussée a certainement limité la capacité des modèles à généraliser. Par ailleurs, la complexité des architectures n'a pas montré d'avantage décisif : ResNet50, pourtant plus sophistiqué, n'a pas performé mieux qu'un simple CNN, ce qui interroge sur l'adéquation entre la complexité du modèle et la nature du problème.

Les perspectives d'amélioration sont nombreuses et prometteuses. Un fine-tuning plus poussé de MobileNetV2, en dégelant les dernières couches et en ajustant le taux d'apprentissage, pourrait permettre d'exploiter pleinement son potentiel. L'introduction de techniques de rééquilibrage des classes, comme la focal loss ou l'oversampling, devrait réduire le nombre de faux négatifs. L'enrichissement du jeu de données par augmentation (rotations, miroirs, changements de contraste) améliorerait probablement la robustesse des modèles. Enfin, l'expérimentation d'architectures plus récentes comme EfficientNet ou les Vision Transformers, combinée à une validation croisée rigoureuse, pourrait permettre d'atteindre des performances satisfaisantes pour une application clinique, avec une AUC cible supérieure à 0.8. Ces améliorations devront s'accompagner d'une analyse plus fine des erreurs pour identifier les cas difficiles et adapter le modèle en conséquence.

5 Conclusion générale

La classification d'images médicales constitue un enjeu majeur en vision par ordinateur, car elle vise à doter les machines de la capacité à reconnaître des motifs visuels et à s'adapter à divers environnements d'imagerie. Elle joue un rôle essentiel dans la reconnaissance rapide et le diagnostic médical. Ce processus consiste à attribuer automatiquement une classe à une image à l'aide d'un système de classification. Avec l'essor de l'intelligence artificielle et de l'apprentissage automatique, les réseaux de neurones convolutifs (CNN) se sont imposés comme les algorithmes les plus performants pour le traitement et la classification d'images au cours des dernières années.

Dans ce mémoire, nous nous sommes intéressés à la détection des mélanomes dans les images médicales utilisant une méthode basée sur l'apprentissage profond. Trois architectures ont été testées : un CNN simple, MobileNetV2 (avec transfer learning), et ResNet50.

L'objectif était de comparer leur capacité à discriminer les lésions bénignes des mélanomes, en utilisant la base de données Melanoma Cancer Dataset (13 879 images). Les modèles ont été évalués sur des métriques clés comme l'accuracy, l'AUC, et le F1-score...

MobileNetV2 a montré la meilleure accuracy (88.7%), mais avec une AUC faible (0.522), ResNet50 et le CNN simple ont eu des performances similaires (~75% accuracy, $AUC \approx 0.5$), suggérant une limite fondamentale dans l'extraction de caractéristiques discriminantes.

Plusieurs problèmes ont été identifiés, notamment un déséquilibre des classes, un prétraitement des données insuffisant, ainsi que des limitations matérielles qui ont restreint les possibilités d'expérimentation. Face à ces constats, plusieurs axes d'amélioration ont été proposés. Il s'agirait d'intégrer des techniques de segmentation (ex. U-Net) pour mieux isoler la lésion, d'enrichir les données d'entrée par des informations cliniques pertinentes, ou encore d'optimiser les modèles via des stratégies de fine-tuning et de rééquilibrage (ex. Focal Loss). Le développement d'une application mobile dotée d'un système explicatif et connecté à des professionnels de santé est également envisagé, afin de rendre l'outil accessible, surtout dans les zones médicalement sous-dotées.

En conclusion, ce travail met en évidence le potentiel des CNN dans le dépistage automatisé du mélanome, tout en soulignant leurs limites actuelles, notamment en matière de fiabilité, d'interprétabilité et d'adaptabilité aux données cliniques réelles. À long terme, l'objectif serait de concevoir un système hybride, associant l'intelligence artificielle à l'expertise humaine, capable d'apprendre continuellement et de s'adapter aux spécificités de chaque population.

6 Références bibliographiques

- [1] de Wazieres, B. (2012). *Algorithms and Interfaces*. Collège National des Enseignants de Médecine Interne. (p.23).
- [2] Ban, L., Labbouz, S., Grindlay, D., Batchelor, J.M. & Ratib, S. (2018). Risk of skin cancer in people with vitiligo: a systematic review and meta-analysis. *British Journal of Dermatology*, 179(4), 971-972. <https://doi.org/10.1111/bjd.16703>
- [3] Jolivot, R. (2011). *Development of an imaging system dedicated to the acquisition analysis and multispectral characterisation of skin lesion* [Thèse de doctorat]. Université de Bourgogne.
- [4] Stanganelli, I. & Pizzichetta, M.A. (2018). Dermoscopy: overview, technical procedures and equipment, color. *Medscape*.
- [5] Pehamberger, H. (2017). *Dermoscopy: Pattern analysis for pigmented and non-pigmented lesions*.
- [6] Zagrouba, E. & Barhoumi, W. (s.d.). *Système d'aide au diagnostic du mélanome pour les dermatologues*. Faculté des Sciences de Tunis, Tunisie.
- [7] Aoun, M. & Madani, H. (s.d.). *Segmentation et caractérisation des lésions dermatologiques par l'approche multifractale* [Mémoire de master]. Université Abou Bakr Belkaïd de Tlemcen.
- [8] Société canadienne du cancer. (s.d.). *Les cancers de la peau*. <https://www.cancer.ca/fr-ca/cancer>
- [9] Chiter, Y. & Hafiane, A. (s.d.). *Apprentissage profond pour l'analyse et la classification d'imageries médicales* [Mémoire de master]. Université Merbah Kasdi Ouargla.
- [10] Nehemy, L. (2015). *Estimation de modèles autorégressifs vectoriels à noyaux à valeur opérateur* [Thèse de doctorat]. Université d'Évry val d'Essonne.
- [11] El Naqa, I. & Murphy, M.J. (2022). *Machine and deep learning in oncology*.
- [12] (s.d.). *Deep Learning*. Wikipédia.
- [13] (s.d.). Formation en ingénierie de la santé. <https://travaux.master.utc.fr/formations-master/ingenierie-de-la-sante/ids122/>

- [14] Touahri, R. (s.d.). *L'apprentissage profond pour la classification et l'interprétation d'images* [Thèse de doctorat]. Université Badji Mokhtar - Annaba.
- [15] Rebala, G., Ravi, A. & Churiwala, S. (2019). *An Introduction to Machine Learning*. Springer Nature Switzerland AG.
- [16] (s.d.). [Article sur le deep learning en imagerie médicale]. IEEE Xplore. <https://ieeexplore.ieee.org/document/7463094>
- [17] (s.d.). [Article scientifique]. Elsevier. <https://www.sciencedirect.com/science/article/pii/S177698171930063X>
- [18] Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J. & Lo, B. (2017). Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*.
- [19] (s.d.). Bases de données sur les maladies de la peau. <https://skinive.com/fr/skin-disease-datasets/>
- [20] Mokri, M.Z. (2017). *Classification des images avec les réseaux de neurones convolutionnels* [Mémoire de master]. Université Abou Bakr Belkaid Tlemcen.
- [21] Boughaba, M. & Boukhris, B. (2016). *L'apprentissage profond pour la classification et la recherche d'images par le contenu* [Mémoire de master]. Université Kasdi Merbah Ouargla.
- [22] (s.d.). [Document pédagogique sur les réseaux neuronaux]. <https://idpoisson.fr/louchet/teaching/timo/Echegut.pdf>
- [23] (s.d.). Définition des réseaux de neurones artificiels. <https://www.lebigdata.fr/reseau-de-neurones-artificiels-definition>
- [24] Echegut, R. (2010). *Projet Image: Méthodes de segmentation* [Article]. Master 2 Automatique.
- [25] (s.d.). Découvrez les différentes couches d'un CNN. *OpenClassrooms*. <https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-donneesvisuelles/5083336-decouvrez-les-differentes-couches-dun-cnn>
- [26] Lambert, R. (2019). Réseau de neurones convolutifs. <https://www.aspexit.com/reseau-deneurones-on-va-essayer-de-demystifier-un-peu-tout-ca-3/>
- [27] Nene, D. & Diane, A. (2019). *La reconnaissance des expressions faciales* [Mémoire]. Université 8 Mai 1945 - Guelma.
- [28] Larbaoui, M.L. & Siamer, D. (2024). *Deep Learning: Reconnaissance Faciale* [Mémoire de master]. Université de Bejaia.

- [29] HABBA, A. & ISHAK, O. (2018). *La classification des images satellitaires par l'apprentissage profond* [Mémoire de master]. Université Ahmed Draïa - Adrar.
- [30] Sekkil, H.M. & Mebrouki, M. (2021). *Etude comparative entre les différentes architectures des réseaux de neurones convolutifs pour la détection de la fatigue du conducteur* [Mémoire]. École Supérieure en sciences appliquées Tlemcen.
- [31] Sellami, Z. (2019). *Optimisation du CNN par l'algorithme Génétique pour la Reconnaissance de Visage* [Mémoire]. Université Mohamed Khider de Biskra.
- [32] Krizhevsky, A., Hinton, G. & Sutskever, I. (2012). Implementation of AlexNet using Keras. <https://engmrk.com/alexnet-implementation-using-keras/>
- [33] Simonyan, K. & Zisserman, A. (2018). VGG16 - Implementation Using Keras. <https://engmrk.com/vgg16-implementation-using-keras/>
- [34] Das, S. (2017, 16 novembre). CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more...
- [35] Boukhlof, D. (2005). *Résolution de problèmes par écosystèmes: Application au traitement d'images* [Thèse]. Université de Biskra.
- [36] Yagoub, D.E. (2018). *Analyse d'images médicales, application à la catégorisation des types de tuberculose pulmonaire* [Thèse]. Université Abdelhamid Ibn Badis - Mostaganem.
- [37] (s.d.). MobileNet: Optimisation de la convolution pour les réseaux de neurones embarqués. <https://www.quantmetry.com/blog/mobilenet-optimisation-de-la-convolution-pour-les-reseaux-de-neurones-embarques/>
- [38] (s.d.). Système d'aide au diagnostic du mélanome pour les dermatologues. https://www.researchgate.net/publication/230642215_Systeme_d'aide_au_diagnostic_du_melanome_pour_les_dermatologues
- [39] (s.d.). FotoSkin: Une application gratuite pour détecter le cancer de la peau. <https://www.cancerpiel.es/fotoskin-un-app-gratuita-para-detectar-cancer-de-piel>
- [40] Shippen, M. (2012). *Is that little mole a big problem? Ask Doctor Mole, the smartphone app that checks for signs of skin cancer* [Application mobile].
- [41] (s.d.). SkinVision: Application de détection des cancers de la peau. <https://www.skinvision.com/>
- [42] Messadi, M. (s.d.). *Caractérisation automatique des tumeurs de la peau par les techniques de traitement et d'analyse d'images* [Thèse de doctorat]. Université Abou Bakr Belkaïd de Tlemcen.

- [43] Pehamberger, H. (2017). *Dermatoscopy: Pattern analysis for pigmented and non-pigmented lesions*.
- [44] Asadi-Aghbolaghi, M., Azad, R., Fathy, M. & Escalera, S. (s.d.). *Multi-level Context Gating of Embedded Collective Knowledge for Medical Image Segmentation* [Rapport technique].
- [45] (s.d.). *Kaggle*. Wikipédia. <https://fr.wikipedia.org/wiki/Kaggle>