

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement supérieur et de la recherche scientifique

Université Mohamed Khider - Biskra

Faculté des Sciences et de la technologie

Département: Génie Electrique

Ref :

جامعة محمد خيضر بسكرة

كلية العلوم والتكنولوجيا

قسم: الهندسة الكهربائية

المرجع :

جامعة محمد خيضر بسكرة



Université Mohamed Khider Biskra

Thèse présentée en vue de l'obtention
Du diplôme de Docteur en 3^{ème} Cycle D-LMD

Option: Biométrie et Télé surveillance

**Apprentissage profond pour le diagnostic
des soins de santé**

Présentée par : DJOUIMA HOSSENA

Soutenue publiquement le : 18/01/2026

Devant le jury composé de :

BAARIR Zine-eddine	Professeur, Université de Biskra	Président
ZITOUNI Athmane	Professeur, Université de Biskra	Rapporteur
MEGHERBI Ahmed Chaouki	Professeur, Université de Biskra	Co-Rapporteur
TOUMI Abida	Professeur, Université de Biskra	Examineur
BENLAMOUDI Azeddine	MCA, Université de Ouargla	Examineur
TALEB - AHMED Abdelmalik	Professeur, Univ de Valenciennes, France	Invité

REMERCIEMENTS

Ce travail a été préparé au sein du Laboratoire de Vision et des Systèmes de Communication (VSC Lab) sous la direction de Monsieur Athmane Zitouni, A.C. Megherbi, et S. Sbaa, Professeurs au département de Génie Electrique à l'Université de Biskra.

Je leur exprime ma profonde gratitude pour leur contribution au développement et à la présentation des idées abordées dans ce travail. Leurs qualités humaines, leur soutien constant, leur disponibilité et leurs précieux conseils scientifiques ont été essentiels à l'aboutissement de cette thèse.

Je remercie chaleureusement Monsieur BAARIR Zine-eddine , Professeur à l'Université de Biskra, pour m'avoir fait l'honneur de présider le jury de ma thèse.

Mes remerciements vont de même aux autres membres de jury examinateurs pour l'intérêt qu'ils ont porté à mon travail en participant au jury de ma thèse. Il s'agit, en l'occurrence de :

Madame TOUMI ABIDA, Professeur à l'Université de Biskra,

Monsieur BENLAMOUDI AZZEDINE, Maître de Conférences A à l'Université de Ouargla,

Monsieur TALEB - AHMED Abdelmalik, Professeur à l'Université de Valenciennes, France.

Je tiens à remercier vivement Monsieur D. BENOUDJIT, Professeur à l'Université de Batna 2, pour ses conseils et ses nombreux encouragements.

Que Monsieur A. DJOUIMA, Maître de Conférences à l'Université de Batna 1, et M.DJOUIMA, Maître de Conférences et Docteur spécialisé en diabétologie, soient aussi bien remercié pour leur appui moral et leur soutien.

Je tiens aussi à remercier Madame L. DJOUIMA Maître de Conférences, à l'Université de Constantine 3 et Madame S. DJOUIMA Professeur à l'Université de Batna 1, pour leurs encouragements incessants.

J'exprime toute ma reconnaissance à Messieurs S. DRID, Professeur, à l'Ecole nationale des énergies renouvelables hns-re2sd, et N. BENOUDJIT pour leur disponibilité sans faille.

Hossena Djouima - VSC Lab-

Je dédie ce travail :

- ***À la mémoire de mon père, qu'il repose en paix.***
- ***À ma mère, que Dieu la préserve, avec l'espoir d'avoir accompli son rêve.***

Résumé

La détection précoce du cancer du sein est essentielle pour améliorer les chances de survie, mais le diagnostic repose encore largement sur des méthodes manuelles, comme la biopsie histologique. Ce processus, exigeant en expertise et prenant beaucoup de temps, ralentit le diagnostic et peut compromettre la qualité des soins. Pour surmonter ces limitations, l'introduction des systèmes de conception assistée par ordinateur (CAO) et des modèles de deep learning a considérablement amélioré la détection d'anomalies, notamment dans le cas du cancer du sein.

Dans cette thèse, l'objectif est de concevoir un modèle de réseaux de neurones profonds capables de classer les tumeurs du sein en bénignes ou malignes, en exploitant la base de données histopathologiques BreakHis (Breast Cancer Histopathological Database), caractérisée par un déséquilibre marqué en faveur des tumeurs malignes. Pour atténuer l'effet du déséquilibre des classes, deux approches sont proposées : la génération d'images synthétiques à l'aide de réseaux génératifs adverses convolutifs (DCGAN), et la transformation en ondelettes. Dans les deux cas, des méthodes classiques d'augmentation de données sont également appliquées. Ces méthodes, intégrées dans une architecture DenseNet201, ont permis une amélioration significative de la précision dans la classification des tumeurs bénignes et malignes, contribuant ainsi au développement d'outils d'aide au diagnostic plus performants.

Mots-clés : *Classification, augmentation de données, apprentissage profond, déséquilibre, DCGAN, ondelettes.*

Abstract

Early detection of breast cancer is essential to improve the chances of survival, but diagnosis still relies heavily on manual methods such as histological biopsy. This process, which requires a great deal of expertise and is time-consuming, slows down diagnosis and can compromise the quality of care. To overcome these limitations, the introduction of computer-aided design (CAD) systems and deep learning models has significantly improved anomaly detection, particularly in breast cancer. In this thesis, the aim is to design deep neural network model capable of classifying breast tumours as benign or malignant, by exploiting the BreakHis (Breast Cancer Histopathological Database) histopathological database, characterised by a marked imbalance in favour of malignant tumours. To mitigate the effect of class imbalance, two approaches are proposed: the generation of synthetic images using Deep Convolutional Generative Adversarial Networks (DCGAN), and wavelet transformation. In both cases, standard data augmentation techniques are also applied. These methods, integrated into a DenseNet201 architecture, have led to a significant improvement in the accuracy of classification of benign and malignant tumours, contributing to the development of more effective diagnostic tools.

Keywords: *Classification, data augmentation, deep learning, imbalance, DCGAN, wavelets.*

ملخص الأطروحة

يعد الكشف المبكر عن سرطان الثدي أمراً حيوياً لتعزيز فرص النجاة، ومع ذلك، لا يزال التشخيص يعتمد بشكل كبير على الطرق اليدوية، مثل الخزعة النسيجية. إن هذه العملية، التي تتطلب خبرة عالية وتستغرق وقتاً طويلاً، تؤدي إلى إبطاء التشخيص وقد تؤثر سلباً على جودة الرعاية الصحية. ولتجاوز هذه العقبات، ساهم إدخال أنظمة التشخيص بمساعدة الحاسوب (CAD) ونماذج التعلم العميق (Deep Learning) في تحسين كشف الاختلالات بشكل ملحوظ، لا سيما في حالات سرطان الثدي.

تهدف هذه الأطروحة إلى تصميم نموذج للشبكات العصبية العميقة قادر على تصنيف أورام الثدي إلى حميدة أو خبيثة، وذلك بالاعتماد على قاعدة البيانات النسيجية المرضية BreakHis (Breast Cancer Histopathological Database)، والتي تتميز بعدم توازن ملحوظ لصالح الأورام الخبيثة. وللتخفيف من حدة هذا التفاوت بين الفئات، تم اقتراح منهجين: توليد صور اصطناعية باستخدام الشبكات التنافسية التوليدية الانتفاقية العميقة (DCGAN)، والتحويل بالموجات (Wavelet Transformation) وفي كلتا الحالتين، تم تطبيق أساليب كلاسيكية لتعزيز البيانات (Data Augmentation). أدت هذه الأساليب، المدمجة في هندسة DenseNet201، إلى تحسين جوهري في دقة تصنيف الأورام الحميدة والخبيثة، مما يساهم في تطوير أدوات أكثر كفاءة للمساعدة في التشخيص الطبي.

الكلمات المفتاحية: التصنيف، تعزيز البيانات، التعلم العميق، عدم توازن البيانات، الشبكات التنافسية التوليدية الانتفاقية العميقة، التحويل بالموجات.

Table des matières

Introduction Générale	1
1. Problématique	1
2. Objectifs.....	2
3. Structure de la thèse.....	2

Chapitre I

L'IA dédiée au diagnostic médical

I.1. Introduction	1
I.2. Histoire de l'IA	1
I.3. Histoire de l'IA dans la santé	3
I.4. Application de IA dans le diagnostic médical.....	5
I.5. Approches de classification des images du cancer du sein	10
I.6. Conclusion	12

Chapitre II

Concepts de base pour l'apprentissage profond

II.1. Introduction	13
II.2. Définition de L'apprentissage profond	14
II.3. Evolution historique	14
II.4. Raisons de l'essor du Deep Learning.....	15
II.5. Domaines d'Application	15
II.6. Principe de Base	15
II.6.1.Neurone artificiel	16
II.6.2.Poids et Biais.....	17
II.6.2.a. Poids (Weights)	17
II.6.2.b. Biais (Biases) :.....	17
II.6.3.Fonctions d'activation.....	17
II.6.3.a. Fonction Sigmoidale	17
II.6.3.b. Fonction ReLU (Rectified Linear Unit)	18
II.6.3.c. Fonction Tanh.....	19
II.6.4.Propagation Avant (Forward Propagation).....	20

II.6.5. Rétropropagation (Backpropagation)	20
II.6.6. Fonction de Coût (Loss Function).....	20
II.6.6.a. Erreur MSE (Erreur Quadratique Moyenne)	21
II.6.6.b. Fonction Cross-Entropy	21
II.6.6.c. Fonction de perte Hinge	22
II.6.6.d. Fonction de perte Huber	23
II.6.7. Optimiseurs.....	24
II.6.7.a. Approche SGD (La descente de gradient stochastique)	24
II.6.7.b. Approche Adam.....	25
II.6.7.c. Approche RMSprop	26
II.6.8. Epoche et Lot (Epoch and Batch)	26
II.6.8.a. Epoche (Epoch)	26
II.6.8.b. Lot (Batch)	26
II.7. Limites de l'apprentissage profond.....	26
II.7.1. Overfitting & underfitting.....	27
II.7.2. Manque de données	28
II.7.3. Données déséquilibrées	29
II.8. Algorithmes d'apprentissage	29
II.8.1. Apprentissage supervisé	29
II.8.2. Apprentissage non supervisé (unsupervised learning)	30
II.8.3. Apprentissage par renforcement (RL)	31
II.9. Conclusion	31

Chapitre III

CNN pour la classification des images

III.1. Introduction	32
III.2. Réseaux de neurones convolutifs (CNN)	32
III.3. Applications des CNN	33
III.4. Caractéristiques principales des CNN	33
III.5. Convolution.....	34
III.6. Activation	36
III.7. Sous-échantillonnage Pooling	37

III.8. Couche d'aplatissement (Flatten)	38
III.9. Couche entièrement connectée (FC)	39
III.10. Fonction de perte.....	39
III.11. Optimiseur	39
III.12. Modèles classiques du CNN	40
III.12.1. LeNet Network.....	40
III.12.2. AlexNet Network.....	40
III.12.3. VGGNet.....	41
III.12.4. Google Net / InceptionV1 à V3.....	42
III.12.5. Modules Inception	43
III.12.6. ResNet	45
III.12.7. DenseNet.....	46
III.13. Apprentissage par transfert (Le transfer learning).....	47
III.13.1. Mise au point des techniques d'apprentissage par transfert	47
III.14. Conclusion	49

Chapitre IV

Data Augmentation pour les Images Histopathologiques

IV.1. Introduction	50
IV.2 Origine et création de la base de données BreakHis	52
IV.2.1. Contenu et structure de la base de données BreakHis	52
IV.2.2. Organisation et catégorisation des images	55
IV.2.3. Illustration du déséquilibre des classes selon le grossissement	56
IV.3. Techniques de gestion des données déséquilibrées.....	57
IV.3.1. Rééchantillonnage : Suréchantillonnage et Sous-échantillonnage	57
IV.3.2. Techniques d'Augmentation de Données.....	58
IV.3.2.a. Augmentation géométrique	58
IV.3.2.b. Modifications photométriques.....	59
IV.3.3. Méthodes Basées sur les Algorithmes	60
IV.4. Traitement des données	61
IV.4.1. Nettoyage des données	62
IV.4.2. Encodage des données	62
IV.4.3. Réduction de la dimensionnalité.....	62

IV.4.4. Traitement spécifique selon le type de données	62
IV.4.5. Batching et Shuffling.....	62
IV.5.Conclusion	62

Chapitre V

Etude Expérimentale

V.1. Introduction	64
V.1.1.Méthode 1 : DCGAN et DenseNet201	64
V.1.2.Méthode 2 : WDT et DenseNet201	64
V.2. Première partie	66
V.2.1.Résolution du déséquilibre des classes avec DCGAN	66
V.2.1.a. Les Deep Convolutional Generative Adversarial Networks (DCGAN)....	66
V.2.1.b. Architecture du DCGAN proposé	67
V.2.2.Augmentation des données	72
V.2.3.Extraction de caractéristiques par apprentissage profond	73
V.2.3.a. La structure DenseNet.....	74
V.2.4.Méthodologie	74
V.2.5.Paramètres d'entraînement du modèle	76
V.2.6.Métriques de performance.....	77
V.2.7.Discussion des résultats	78
V.3. Deuxième partie	79
V.3.1.Résolution du déséquilibre des classes avec les ondelettes	79
V.3.2.Transformée en ondelettes discrète (DWT)	80
V.3.3.Méthodologie	82
V.3.4.Paramètres d'entraînement du modèle	87
V.3.5.Analyse des résultats	88
V.4. Conclusion	93
Conclusion Générale	95
Références Bibliographiques	98

Introduction Générale

Au cours des dernières décennies, le domaine de la santé a connu des avancées technologiques sans précédent, transformant profondément les méthodes de diagnostic, de traitement et de gestion des maladies. Parmi ces avancées, l'intelligence artificielle (IA) et, plus particulièrement, l'apprentissage profond (Deep Learning), ont émergé comme des outils puissants capables de relever les défis complexes à partir de données volumineuses et variées.

Grâce à ces capacités, il surpasse les méthodes traditionnelles dans des domaines tels que la reconnaissance d'images, la classification, la prédiction et l'analyse des données. En médecine, ces technologies se sont avérées particulièrement prometteuses pour l'analyse des images médicales, permettant d'identifier des pathologies avec une précision souvent comparable, voire supérieure, à celle des experts humains.

Les recherches et les travaux récents s'intéressent à appliquer l'apprentissage profond dans le diagnostic du cancer du sein, l'une des principales causes de mortalité chez les femmes dans le monde, avec 2,3 millions de cas féminins et 670 000 décès recensés en 2022 selon l'OMS. Malgré les progrès médicaux, les erreurs de diagnostic et les retards de détection restent des défis majeurs, compromettant les chances de survie des patientes.

1. Problématique

Le diagnostic du cancer du sein repose principalement sur l'analyse d'images histopathologiques, une tâche qui nécessite l'expertise de pathologistes pour identifier les anomalies cellulaires. Cependant, cette évaluation reste subjective, dépend fortement du niveau d'expérience du spécialiste et peut être influencée par la variabilité des échantillons. Avec l'essor du deep learning, des modèles de classification automatique des images histopathologiques ont été développés pour améliorer la précision du diagnostic. Néanmoins, plusieurs défis persistent :

- **Le déséquilibre des classes dans les bases de données :** Les bases de données médicales, comme BreakHis, sont souvent déséquilibrées, avec une prédominance d'échantillons appartenant à certaines classes (ex. :

tissus malins), ce qui peut biaiser les performances des modèles d'apprentissage profond.

- **Complexité et variabilité des images** : Les structures tissulaires présentent une diversité importante, avec des variations entre malades et des différences subtiles entre les tissus bénins et malins.
- **L'extraction optimale des caractéristiques pertinentes.**

Face à ces défis, une question clé émerge : Comment développer une approche basée sur l'apprentissage profond capable d'améliorer la classification des images histopathologiques du cancer du sein, tout en prenant en compte les défis liés à la qualité des données, à l'extraction des caractéristiques et à la robustesse du modèle ?

2. Objectifs

Cette thèse s'articule autour des objectifs suivants :

- Développer une méthode innovante pour surmonter le déséquilibre des classes dans les ensembles de données d'images histopathologiques, en explorant des techniques de prétraitement, telles que la transformation par ondelettes.
- Exploiter les capacités des modèles pré-entraînés, tels que DenseNet201, pour améliorer l'extraction des caractéristiques discriminatives et capturer des informations pertinentes à différents niveaux d'abstraction.

Cette thèse ambitionne d'apporter une contribution significative à la recherche en apprentissage profond appliqué au diagnostic médical, en proposant des solutions concrètes et performantes pour l'analyse des images histopathologiques. Ces travaux visent à renforcer le potentiel des technologies d'IA dans les environnements cliniques, notamment pour la détection précoce et la prise en charge du cancer.

3. Structure de la thèse

La présente thèse est constituée de cinq chapitres présentés comme suit :

Le premier chapitre sera consacré à l'évolution de l'intelligence artificielle (IA) et à son exploitation croissante dans le diagnostic médical.

Dans le deuxième chapitre, nous aborderons les concepts fondamentaux de l'apprentissage profond, en mettant l'accent sur la structure des réseaux neuronaux, les approches de régularisation et les méthodes d'optimisation.

Dans le troisième chapitre, nous approfondirons l'étude des réseaux de neurones convolutifs (CNN) , qui constitue un pilier majeur de la vision par ordinateur. Nous y examinerons leur architecture en détail, en mettant l'accent sur la manière dont les CNN sont conçus pour extraire des caractéristiques visuelles. De plus, nous aborderons le concept de « transfert. Learning », une technique qui permet d'adapter un modèle pré-entraîné à de nouvelles tâches, facilitant ainsi l'application des CNN dans des domaines spécifiques, tels que la classification d'images médicales.

Dans le quatrième chapitre, nous présenterons la base de données BreakHis utilisée pour cette étude, qui présente un déséquilibre de classe et un nombre d'échantillons réduits. Nous explorons les défis associés à ces problèmes, notamment les biais potentiels dans les résultats du modèle. Ce chapitre passe en revue les différentes solutions proposées dans la littérature pour contourner ces problèmes, telles que les techniques de rééchantillonnage, et l'augmentation de données.

Nous présenterons dans le cinquième chapitre, les résultats expérimentaux obtenus pour la classification des images histopathologiques du cancer du sein, en mettant l'accent sur les méthodes utilisées pour améliorer l'efficacité de notre modèle proposé.

Enfin, cette thèse se terminera par une conclusion générale, accompagnée de perspectives éventuelles.

Chapitre I

L'IA dédiée au diagnostic médical

I.1.Introduction

Le diagnostic représente l'une des tâches primordiales des cliniciens, car les erreurs dans ce domaine ont une incidence significative sur les taux de maladie et les pertes de vies humaines, affectant environ 10 à 15 % des cas en raison de diagnostics incorrects ou tardifs [1]. Améliorer la précision des diagnostics demeure un défi capital pour les systèmes de santé à l'échelle mondiale [2]. Initialement, la recherche se concentrait sur des méthodes statistiques, mettant l'accent sur les données quantitatives mais limitant la prise en considération des facteurs contextuels et qualitatifs. Les orientations de recherche ont évolué par la suite vers des approches prenant en compte des aspects cliniques, comportementaux, sociaux et psychologiques. De nos jours, les diagnostics médicaux bénéficient de l'intégration de diverses sources d'informations telles que l'imagerie médicale et les avancées technologiques, ce qui permet une prise en charge plus précise des patients [3]. Cette transformation est en grande partie liée à l'émergence de l'intelligence artificielle (IA), une branche de l'informatique dédiée à la conception de systèmes aptes à simuler certaines capacités humaines telles que la perception, la compréhension du langage, le raisonnement, et la prise de décision. L'IA soutient les praticiens dans des tâches administratives, la documentation clinique, la sensibilisation des patients, ainsi que dans des applications spécifiques comme l'analyse des images médicales, le suivi des dispositifs médicaux et des patients. La figure I.1 met en évidence neuf applications essentielles de l'intelligence artificielle dans le domaine médical [4]. Ce chapitre offre un aperçu général sur l'histoire et les fondements de l'IA dans le diagnostic médical.

I.2.Histoire de l'IA

Au cours des années 1950 et 1970, les premières explorations de l'IA ont été marquées par des contributions importantes de chercheurs tels qu'Alan Turing et John McCarthy [5][6].

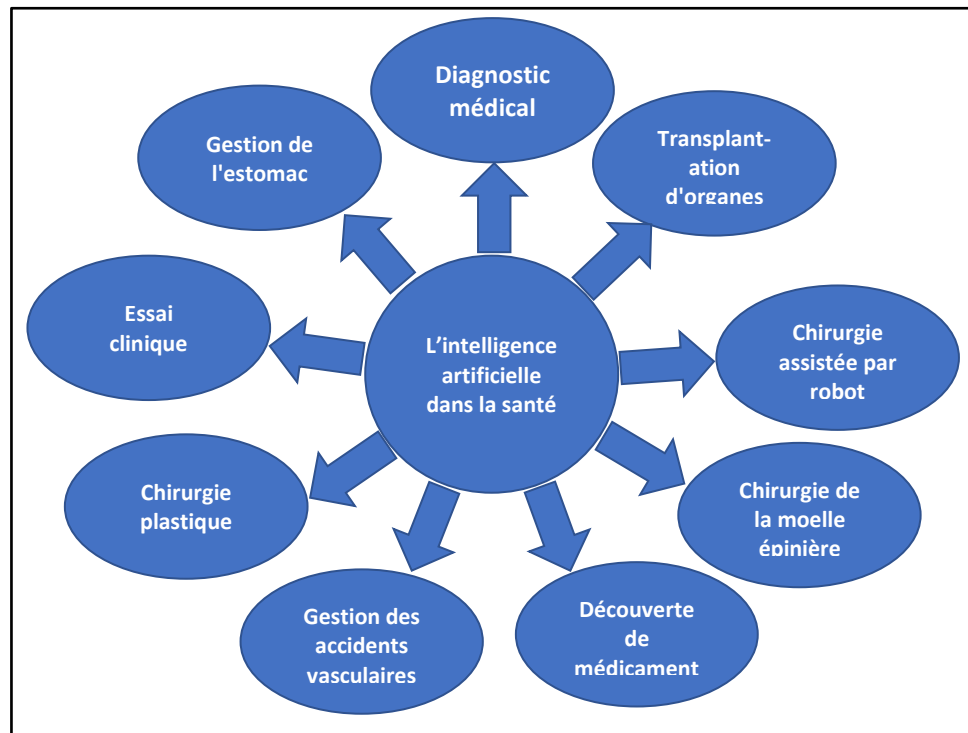


Figure I.1. Contributions de l'intelligence artificielle au domaine médical.

L'article de Turing de 1950 a posé les bases de l'IA en définissant l'objectif de reproduire l'intelligence humaine dans des machines [7][8]. L'intelligence artificielle a été formellement introduite lors de la conférence de Dartmouth en 1956, Un événement marquant le début de l'IA moderne, sous la direction de chercheurs renommés tels que John McCarthy et ses collaborateurs [8][9]. Au cours des années ultérieures, les chercheurs ont développé des programmes permettant aux ordinateurs d'exécuter des tâches faisant appel à des capacités cognitives, notamment les échecs, la détection de formes et l'analyse du langage naturel.

Toutefois, les progrès ont été freinés, en raison des limites des capacités de traitement et des systèmes de stockage des ordinateurs de l'époque [10].

L'émergence de l'apprentissage automatique dans les années 1980 a redonné un vif intérêt à l'IA, facilitant l'apprentissage des ordinateurs à partir de données sans être programmes de façon directe. Ce développement a favorisé des avancées marquantes, en particulier dans des secteurs comme la reconnaissance vocale et la vision par ordinateur [11].

Au cours des dernières décennies, l'IA a enregistré un développement rapide en raison de l'énorme volume de données, l'optimisation des performances des systèmes de traitement (la puissance de calcul) et l'essor des réseaux de neurones profonds. Ces avancées ont permis à l'IA d'accomplir des fonctions complexes telles que la traduction automatique, la conduite autonome et la reconnaissance faciale [12]. Ces dernières années, le domaine de l'IA a enregistré une évolution spectaculaire, dans plusieurs secteurs notamment la santé, les transports, la finance, le marketing, ou encore l'éducation. Parmi les nombreuses applications concrètes, on peut citer les assistants vocaux (Siri et Alexa), les véhicules autonomes, ainsi que les robots collaboratifs, qui illustrent l'intégration croissante de l'IA dans la vie quotidienne. La technologie ne cesse de se développer rapidement, ouvrant la voie à des innovations inédites et à une multiplication des usages potentiels [13]. Une chronologie simplifiée des étapes importantes de l'IA est présentée dans la figure I.2.

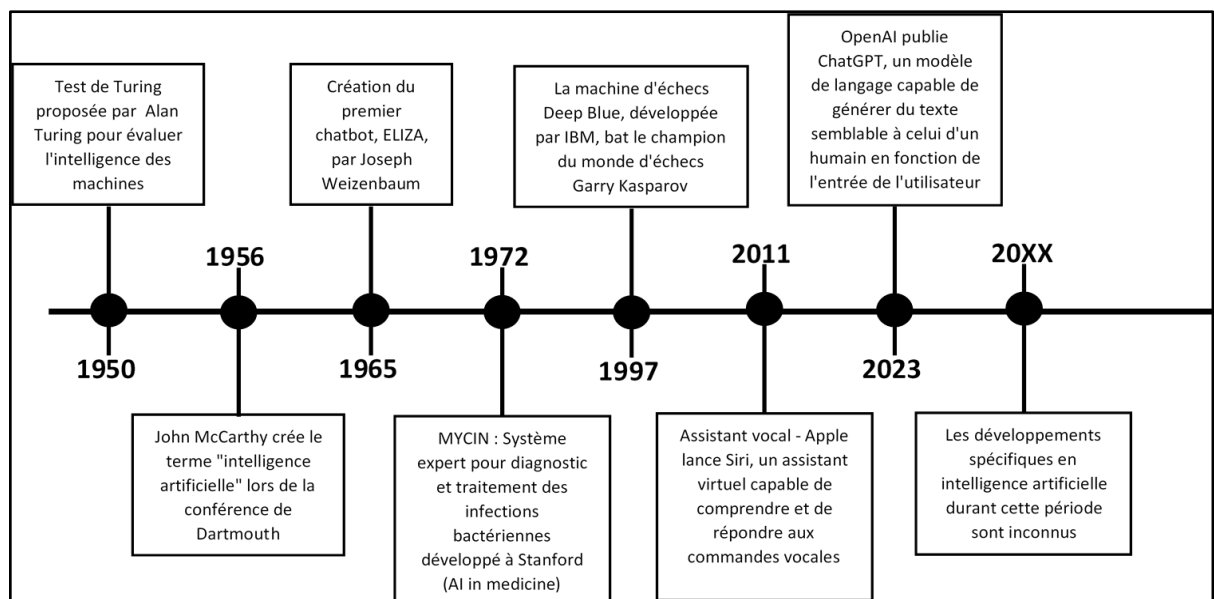


Figure I.2. Brève chronologie de l'histoire de l'IA.

I.3.Histoire de l'IA dans la santé

L'avènement de l'IA dans le domaine médical a pris racine dans les années 1960- 1970, avec le développement de MYCIN par Edward Shortliffe et son équipe à l'Université Stanford marquant ainsi le début la première génération d'IA en médecine [14]. Ce système expert visait à identifier les bactéries responsables d'infections graves (bactériémie, méningite), et à recommander

des antibiotiques. Son nom est dérivé du suffixe “-mycin” utilisé pour de nombreux antibiotiques [8]. MYCIN utilisait une base de connaissances d'environ 600 règles. Il interrogeait le médecin exécutant le programme via une série de questions simples (oui/non ou textuelles). A la fin, il fournissait une liste de bactéries suspectes classées par probabilité de diagnostic, avec diagnostics expliqués et traitements médicamenteux adaptés [1]. Parallèlement, l'utilisation de techniques d'IA, notamment les réseaux neuronaux, pour modéliser des processus biologiques complexes, ainsi que l'émergence des techniques de reconnaissance de formes, ont ouvert la voie aux systèmes de diagnostic assisté par ordinateur (DAO) en radiologie et en pathologie [15].

Les avancées réalisées au cours des années 1980 dans les domaines des arbres de décision, des réseaux neuronaux et des machines à vecteurs de support ont marqué un tournant pour l'intégration de l'IA en médecine, contribuant ainsi à une amélioration significative du diagnostic et du traitement des maladies [16].

Durant les années 1990, l'IA a largement contribué à l'analyse d'images médicales, utilisant des réseaux de neurones artificiels et d'autres techniques d'apprentissage automatique (machine Learning) pour détecter et classer les anomalies, améliorant ainsi la précision des diagnostics radiologiques [17][18]. Parallèlement, elle a été employée pour évaluer les risques de maladies chez les patients en analysant leurs antécédents médicaux et d'autres données cliniques, permettant aux médecins de détecter les cas à risque élevé et d'instaurer des actions préventives adaptées [19][20]. Dans le cadre des études cliniques, l'IA a accéléré l'analyse de vastes ensembles de données, identifiant des schémas et des tendances essentiels pour la découverte de médicaments et la compréhension des maladies, contribuant ainsi à l'avancement des connaissances médicales [21][22].

Les avancées des années 1990 se sont poursuivies durant les années 2000, avec l'introduction de robots dotés d'algorithmes d'IA. Ces robots ont été utilisés dans le domaine médical pour réaliser des interventions chirurgicales telles que le traitement de la tumeur pulmonaire, de la membrane interne de l'utérus, des organes prostatiques et coliques [4].

Depuis la dernière décennie (2010), les techniques de Deep Learning (DL), en particulier les CNN (réseaux de neurones convolutifs), ont connu une adoption croissante dans de nombreux secteurs de la santé. En radiologie, en pathologie et en dermatologie, les CNN ont été massivement intégrés pour des tâches de reconnaissance d'images [23]. De plus, dans le cadre du traitement du langage naturel (Natural Language Processing), les CNN sont utilisés pour analyser les dossiers électroniques de santé, ouvrant ainsi de nouvelles perspectives pour l'exploitation des données textuelles en santé [24].

Dans les années 2020, l'intelligence artificielle (IA) continue de se développer dans les domaines de la médecine de précision, de la découverte de médicaments et de la génomique. On assiste également au développement de robots et d'assistant virtuels alimentés par l'IA pour prendre en charge les patients et faciliter les consultations à distance (télé-médecine) [25].

I.4. Application de IA dans le diagnostic médical

L'exploitation de l'IA dans les systèmes de diagnostic médical, présente un immense potentiel pour améliorer la précision, la rapidité et l'efficacité des diagnostics, tout en réduisant les erreurs humaines. En analysant des volumes importants de données (images médicales, tests sanguins et des antécédents), elle fournit des diagnostics à la fois fiables et précis. Cette revue présente quelques travaux publiés sur les méthodes d'apprentissage profond et examine leurs succès dans le diagnostic.

Diwakaran et al ont développé un modèle, de diagnostic du cancer du sein basé sur la BCP-TL (Breast Cancer Prognosis Based Transfer Learning) à l'aide d'architectures CNN telles que Xception et Channel Boosted CNN sur la base de données MIAS (Mammographic Image Analysis Society). Ce modèle a démontré d'excellentes performances, avec une précision globale de 98,96% grâce à la combinaison de Xception et de Channel Boosted CNN [26].

Izadkhah et al, ont proposé un modèle de deep learning pour la classification du cancer du sein, atteignant une précision élevée sur des images échographiques et histopathologiques. Inspiré des architectures GoogLeNet et ResNet, ce modèle intègre des connexions courtes, des convolutions 1×1 pour réduire les coûts de calcul et améliorer les performances, ainsi que

l'informatique granulaire pour une extraction fine des caractéristiques. Les précisions obtenues s'élèvent respectivement à 93 % et 95 % [27].

Khan et Elfouly ont développé un modèle de diagnostic des maladies cardiovasculaires basé sur un réseau neuronal convolutionnel unidimensionnel (1D CNN), utilisant des paramètres cliniques issus du NHANES (National Health and Nutrition Examination Survey). Pour pallier le déséquilibre des données, ils ont appliqué des techniques de sous-échantillonnage et de pondération des classes. Les résultats indiquent que le modèle 1D CNN surpasse les algorithmes d'apprentissage automatique traditionnels en termes de précision ainsi que de réduction des erreurs positives et négatives [28].

Hassan et al ont développé ont conçu un modèle de diagnostic du COVID-19 spécifiquement destiné aux patients cardiaques, en s'appuyant sur l'apprentissage profond et l'analyse d'images ECG. Ce modèle combine des techniques d'apprentissage par transfert et ensembliste, intégrant les architectures pré-entraînées VGG-19, AlexNet et ResNet-101. L'utilisation de techniques d'augmentation des données a permis d'améliorer ses performances, atteignant une précision globale de 99,1 %, une sensibilité de 99 % et une précision de 100 %. Cette étude met en avant l'importance de modèles adaptés à des populations spécifiques pour optimiser le diagnostic malgré des données limitées [29].

L'article de Hameed et al présente une méthode basée sur un ensemble de modèles d'apprentissage profond pour classifier les images histopathologiques du cancer du sein en carcinome (malin) et non-carcinome (bénin ou normal). En utilisant des variantes des architectures VGG16 et VGG19, les auteurs ont combiné les prédictions des modèles ajustés via une stratégie d'ensemble. Cette méthode a atteint une précision globale de 95,29 % et une sensibilité de 97,73 % pour le carcinome, démontrant son efficacité [30].

L'article de Wang et al présente une méthode de classification de la sclérose en plaques basée sur DenseNet-201 et une stratégie de facteur d'apprentissage composite (CLF). En exploitant le transfert d'apprentissage, cette approche optimise l'entraînement des couches et utilise un pré-calcul pour réduire les besoins en stockage. Avec une sensibilité de 98,27 %, une spécificité de 98,35 %, et une précision de 98,31 %, DenseNet-201 (réglage D)

surpasse les méthodes classiques et avancées. Les auteurs soulignent son efficacité et proposent d'explorer des modèles combinés et d'autres modalités [31].

Suganya et al, ont proposé une approche de diagnostic automatique et interprétable de la maladie de Parkinson (PD), en combinant l'imagerie par résonance magnétique quantitative et les images pondérées T1. Leur pipeline se compose de deux étapes : (1) une segmentation des régions d'intérêt cérébrales (noyaux gris profonds) à l'aide d'un CNN (réseau de neurones convolutifs), et (2) un modèle de classification basé sur un CNN avec un mécanisme d'attention anatomique pour différencier les patients atteints de PD des sujets sains. Les résultats obtenus montrent une précision de 92,0 % et une aire sous la courbe (AUC) de 0,901 pour le diagnostic de la PD. L'utilisation du mécanisme d'attention anatomique renforce l'interprétabilité du modèle en identifiant les régions cérébrales les plus pertinentes pour le diagnostic, ce qui en fait un outil réellement applicable en clinique pour la détection précoce et le suivi de la maladie [32].

Yoo et al ont proposé un algorithme basé sur l'apprentissage profond (DL) pour diagnostiquer la myopathie et la neuropathie à partir des résultats d'électromyographie (EMG) à l'aiguille. Les chercheurs ont développé un réseau neuronal convolutif unidimensionnel et ont comparé les performances de leur modèle avec celles de six médecins expérimentés. Les résultats ont révélé que l'algorithme basé sur le DL a obtenu une Accuracy de 0,875, un Recall de 0,820 et une spécificité de 0,904, tandis que les médecins ont obtenu une précision moyenne de 0,694, une sensibilité de 0,537 et une spécificité de 0,773. L'étude suggère que l'apprentissage profond (DL) pourrait jouer un rôle clé dans le diagnostic des maladies neuromusculaires, offrant une méthode plus précise et efficace pour interpréter les données EMG [33].

Podder et coll. ont une approche développée basée sur le deep learning pour détecter la COVID-19, la pneumonie et distinguer les patients normaux à partir d'images radiographiques pulmonaires. En utilisant DenseNet-169 et DenseNet-201 optimisés avec transfert d'apprentissage, prétraitement et augmentation des données, ils ont obtenu des précisions de 91,95 % et 92,35 %, surpassant les modèles existants. Cette étude met en évidence l'impact de

l'optimisation des architectures et des techniques d'augmentation sur le diagnostic précoce des maladies pulmonaires.[34].

Sfayyih et coll. ont réalisé une revue complète des architectures de réseaux de neurones profonds basées sur l'analyse acoustique pour le diagnostic des maladies pulmonaires. Leur travail aborde le prétraitement des signaux acoustiques, incluant la détection des bruits de fond et la normalisation, ainsi que l'extraction de caractéristiques telles que les Coefficients cepstraux en fréquence de Mel (MFCC) et les transformés en ondelette. Ils comparent ensuite diverses architectures de réseau, comme VGG, ResNet, LeNet, Inception Net et AlexNet, en se référant à des bases de données telles que RSD ICBHI 2017, HF_Lung_V1 et Respiratory-Database@TR. Les auteurs soulignent également les défis tels que le déséquilibre et la qualité des données, ainsi que les perspectives pour des applications cliniques. Leur revue inclut des exemples de performances des modèles et discute des méthodes d'augmentation des données, tout en mettant en lumière les avantages et inconvénients des DNN par rapport à d'autres méthodes de classification. [35]. L'article de Podder propose un nouveau cadre de deep learning (DL) pour l'analyse des maladies pulmonaires infectieuses, y compris le COVID-19 et la pneumonie, à partir de scanners thoraciques et d'images radiographiques (CXR). Le modèle proposé, appelé LDDNet, est basé sur DenseNet201 optimisé avec des couches supplémentaires de pooling global 2D, des couches denses et de dropout, et la normalisation par lot. LDDNet a été évalué sur trois ensembles de données multiclasse de maladies pulmonaires provenant de sources open access. Les résultats montrent que LDDNet offre une meilleure performance que les modèles ResNet152V2 et XceptionNet existants pour la détection de COVID-19 et la pneumonie. LDDNet peut être considéré comme un modèle potentiel pour la classification multiclasse des images de scanners thoraciques et radiographiques pour le diagnostic des maladies pulmonaires infectieuses [36].

Solano et al ont présenté une comparaison de différentes architectures de réseaux de neurones pour la segmentation des vaisseaux sanguins rétiniens dans le cadre du diagnostic de la rétinopathie diabétique. Les architectures étudiées comprennent U-Net (CNN), UNETR et Swin-UNET et ConvMixer).

Les résultats montrent que ConvMixer, avec moins de paramètres, atteint les performances les plus remarquables sur les plans de la précision et de la robustesse. Les auteurs concluent que les architectures hybrides comme ConvMixer peuvent offrir un bon compromis entre complexité et performance pour les tâches de segmentation d'images médicales [37].

Aslan et al ont développé une approche originale fondée sur l'apprentissage profond permettant la prédiction du diabète en convertissant les données cliniques numériques en images. L'étude utilise l'algorithme de sélection de caractéristiques Relief pour déterminer l'importance des caractéristiques et les convertir en images. Des techniques d'augmentation de données sont appliquées sur les images pour améliorer la diversité et la robustesse de l'ensemble de données. Les modèles CNN pré-entraînés ResNet18 et ResNet50 sont utilisés pour extraire des paramètres profonds à partir des images, qui sont ensuite fusionnées et classées avec un algorithme SVM pour prédire le statut du diabète. Les performances de différentes architectures CNN et de différentes fonctions noyau SVM sont comparées pour évaluer la méthode proposée. Les résultats révèlent que la méthode atteint une précision de classification de 92,19 %, dépassant les résultats de la plupart des études antérieures sur le même ensemble de données [38].

Xu et al ont présenté un modèle de prédiction de diabète innovant qui combine l'apprentissage profond et l'intégration de représentations de connaissances médicales. Il construit un graphe de connaissances en utilisant les plages normales des indices de bilan de santé et les valeurs de détection, puis utilise la représentation de connaissances pour encoder ces données en vecteurs. Ces vecteurs sont ensuite intégrés dans un modèle d'apprentissage profond utilisant un mécanisme d'auto-attention et des réseaux de neurones convolutifs (CNN) pour extraire les paramètres pertinents. Les résultats montrent une amélioration significative en termes d'exactitude et de rappel par rapport aux méthodes de machine Learning et d'apprentissage profond existantes. L'étude de l'impact des changements de la dimension des vecteurs de représentation des connaissances montre que la performance est optimale avec une dimension de 256. Enfin, la visualisation des poids de l'auto-attention illustre l'interprétabilité du modèle et sa capacité à capturer les relations pertinentes entre les indices de bilan de santé et les valeurs de

détection. Cette approche novatrice offre des performances améliorées et une interprétabilité accrue pour la prédiction du diabète, avec des implications importantes pour les applications médicales et la recherche en intelligence artificielle [39].

Aptekarev et al ont présenté une méthode innovante de diagnostic de l'asthme bronchial basée sur l'analyse de sons respiratoires en utilisant des techniques d'apprentissage profond. Les chercheurs ont constitué une base de données anonyme comprenant des enregistrements de sons respiratoires de patients atteints de diverses maladies respiratoires et de volontaires en bonne santé. Les enregistrements ont été capturés à quatre points clés : dans la cavité buccale, au-dessus de la trachée, sur la poitrine et sur le dos. Le logiciel développé fournit des classifications binaires, notamment "malade/sain" et "patient asthmatique/non-patient asthmatique et sain", en utilisant des modèles d'apprentissage profond entraînés sur la base de données. Les métriques de sensibilité, de spécificité et d'exactitude pour les deux classifieurs ont été évaluées sur des échantillons de test allant de 50 enregistrements dans le groupe de contrôle (personnes en bonne santé) à 50 enregistrements dans le groupe de comparaison (patients atteints de maladies respiratoires), avec des résultats allant de 82% à 93% [40].

I.5.Approches de classification des images du cancer du sein

Dans le domaine de la classification des images histopathologiques du cancer du sein, plusieurs études ont été réalisées en se concentrant sur des modèles d'apprentissage profond . Ces travaux utilisent des approches d'apprentissage profond pour relever des défis complexes, tels que le déséquilibre des classes et la variabilité des données.

Afin d'établir une comparaison pertinente avec les contributions de cette thèse, un tableau récapitulatif est présenté ci-dessous. Il met en évidence Le prétraitement des données, les modèles appliqués et les performances obtenues.

Reference	Prétraitement des données	Méthode/Modèle	Performances	Conclusion
Das et al. [41].	<ul style="list-style-type: none"> - Décomposition des images avec EWT et VMD. - Utilisation de trois CNN pour classificateurs de base. 	<ul style="list-style-type: none"> - Ensemble avec MLP combinant les sorties des trois CNN. 	<ul style="list-style-type: none"> - Précision : 98.08% sur la base de données de cancer du sein histopathologique. 	<ul style="list-style-type: none"> - Méthode d'apprentissage profond en ensemble efficace pour la détection du cancer du sein, améliorant les performances avec EWT et VMD.
Khan et al. (2022) [42]	<ul style="list-style-type: none"> - Plusieurs techniques d'augmentation des données sont appliquées - le nombre d'images dans le dataset BreakHis est passé de 7909 à 54403. 	<ul style="list-style-type: none"> - Utilisation de plusieurs modèles CNN pré-entraînés suivie d'une fusion de ces caractéristiques pour la classification. 	<ul style="list-style-type: none"> - 99% de précision pour la classification binaire. - Aire sous la courbe ROC (AUC) : 0.993. 	<ul style="list-style-type: none"> - Le cadre MultiNet a montré une Précision élevée dans la classification des images,
Liew et al. (2021) [43]	<ul style="list-style-type: none"> - Le rééchantillonnage corrige le déséquilibre des classes. - La normalisation des couleurs atténue les variations. - Les images sont redimensionnées et augmentées 	<ul style="list-style-type: none"> - Amélioration de l'apprentissage en utilisant DenseNet . - XGBoost remplace la couche entièrement connectée d'origine dans DenseNet201. 	<ul style="list-style-type: none"> - La méthode combine l'apprentissage en profondeur et XGBoost pour une classification efficace, avec une précision de 97 %. 	<ul style="list-style-type: none"> - Une précision et une généralisation améliorées. - Le rééchantillonnage et l'augmentation des données permettent de corriger efficacement les biais et le surajustement. - La robustesse des tâches binaires est démontrée.
Hiren et al. (2020) [44]	<ul style="list-style-type: none"> - Les images ont été divisées en patches de 512x512 pixels avec un chevauchement de 50 %. - Des transformations telles que la rotation, mirroring vertical et le décalage ont été appliquées. 	<ul style="list-style-type: none"> - Une structure CNN qui intègre des caractéristiques spatiales et spectrales obtenues à partir d'une transformation en ondelettes - La normalisation par lots est appliquée après chaque couche pour améliorer la convergence. 	<ul style="list-style-type: none"> - La méthode Proposée a atteint Une précision Moyenne de 97,58 % 	<ul style="list-style-type: none"> - L'intégration des Caractéristiques Spectrales et spatiales améliore considérablement la précision de la classification.
Saini and Susan (2020) [45]	<p>Redimensionnées à une taille uniforme de $224 \times 224 \times 3$ (RVB) à l'aide d'une interpolation bilinéaire</p> <p>Méthode DCGAN pour générer des données synthétiques</p>	<ul style="list-style-type: none"> - Combinaison de DCGAN pour l'augmentation des données de la classe minoritaire et du transfert d'apprentissage avec le modèle pré-entraîné VGG16. 	<ul style="list-style-type: none"> - Une précision de 96,5 % 	<ul style="list-style-type: none"> - L'approche de transfert d'apprentissage associée à l'augmentation des données avec DCGAN est efficace pour traiter les déséquilibres dans la classification des images histopathologiques.

I.6. Conclusion

Ce chapitre a exploré l'application de l'intelligence artificielle (IA) au diagnostic médical, soulignant son potentiel à améliorer la précision et la rapidité des diagnostics, tout en dépassant les limites des approches traditionnelles. En intégrant des données variées comme l'imagerie médicale, l'IA contribue à l'optimisation des systèmes de santé, de l'analyse d'images au suivi des patients. Parmi ses avancées, le Deep Learning s'impose grâce à sa capacité à traiter de grandes quantités de données et à extraire automatiquement des paramètres pertinents. Les prochains chapitres approfondiront ce domaine et son impact croissant sur le diagnostic médical.

Chapitre II

Concepts de base pour l'apprentissage profond

II.1.Introduction

Le Deep Learning également connu sous le nom d'apprentissage profond, est une avancée dans le domaine de l'apprentissage automatique (ML), qui fait participer tous les chercheurs grâce à son efficacité dans le traitement et la résolution de problèmes complexes à travers plusieurs domaines, notamment le traitement d'images, la reconnaissance vocale et le traitement du langage naturel [46][47]. Le succès de cette technologie repose principalement sur deux piliers : la disponibilité massive de données ("big data") et L'expansion significative des capacités de traitement des unités graphiques (GPU). Ces GPU permettent aujourd'hui l'entraînement de réseaux neuronaux profonds, composés de multiples couches cachées, qui peuvent capturer des représentations hiérarchiques et abstraites des données. En effet, cette capacité à traiter des volumes de données énormes tout en apprenant des caractéristiques complexes fait du Deep Learning l'une des technologies les plus adaptées pour relever les défis modernes de l'intelligence artificielle [48].

La figure II.1 illustre trois facteurs qui ont contribué à l'essor du Deep Learning : la disponibilité des données d'apprentissage en grandes quantités, la puissance des GPU, et l'accès aux frameworks open-source tels que TensorFlow et PyTorch, offerts par des géants technologiques comme Google, Facebook, et Microsoft. Ces outils ont démocratisé l'usage de l'apprentissage profond, ouvrant la voie à des innovations dans des domaines diversifiés [49].

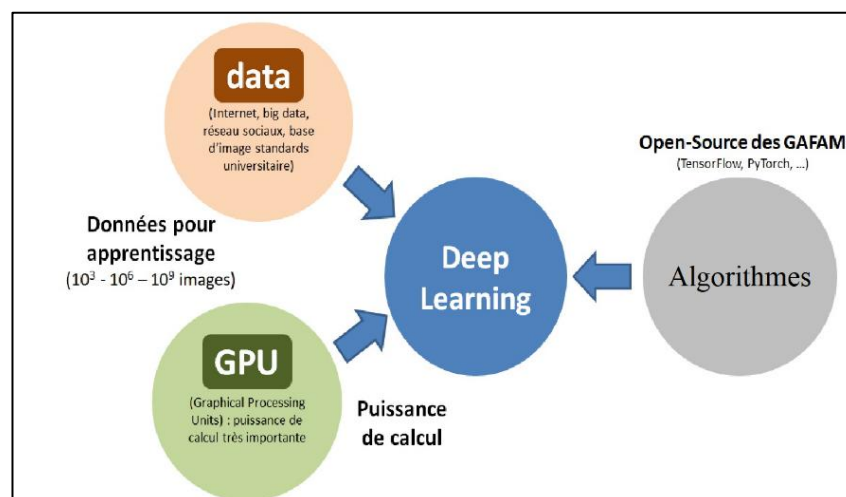


Figure II.1. Facteurs clés pour l'essor du Deep Learning[48].

Ce chapitre, examine en détail les composants essentiels du Deep Learning et leur rôle dans la résolution de problèmes complexes. Il explore l'architecture des réseaux neuronaux profonds, en expliquant comment ils sont construits et comment ils fonctionnent.

II.2. Définition de L'apprentissage profond

L'apprentissage profond (AP), est une branche de l'IA (Intelligence artificielle) permettant aux ordinateurs d'acquérir des connaissances et de traiter de manière autonome des problèmes complexes en utilisant de grandes quantités de données, sans programmation explicite [50]. Inspiré par le fonctionnement du cerveau humain, le Deep Learning vise à imiter le réseau de neurones biologiques en utilisant des structures de calcul appelées réseaux de neurones artificiels. Cette approche a permis des progrès notables dans divers domaines notamment la compréhension du langage naturel, la reconnaissance d'images, la traduction automatique, etc. [51].

La figure II.2 illustre la hiérarchie des concepts d'intelligence artificielle, de machine Learning et de Deep Learning.

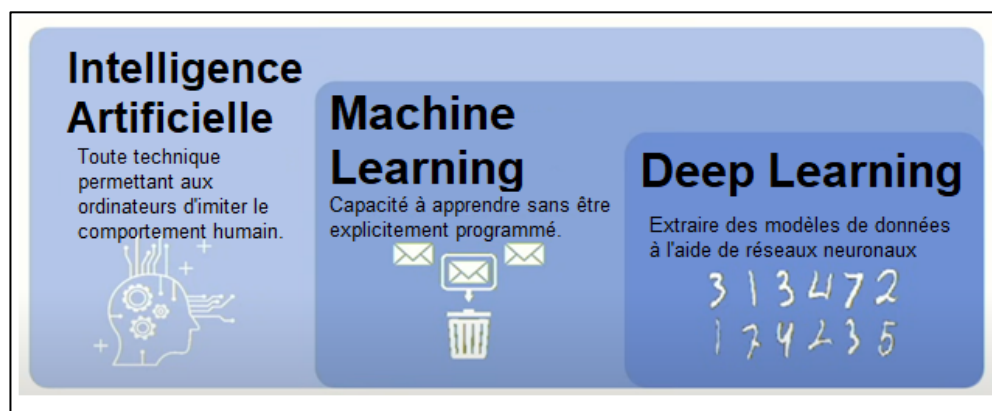


Figure II.2. Relation entre IA, ML et DL.

II.3. Evolution historique

L'évolution historique du Deep Learning est étroitement liée aux progrès de l'intelligence artificielle et des réseaux neuronaux. Ses origines remontent aux années 1940 avec la proposition du perceptron, un modèle simple de neurone artificiel. Cependant, les premières tentatives de réseaux multicouches furent freinées par des limitations informatiques et théoriques, notamment l'absence de méthodes efficaces pour l'entraînement.

Ce n'est qu'à partir de la fin des années 1980, avec l'avènement de l'approche de rétropropagation du gradient, que les réseaux de neurones à couches profondes sont devenus plus pratiques. Les années 2000 ont vu un essor majeur, grâce à l'augmentation des puissances de traitement et à l'accès à de grandes quantités d'échantillons. Des architectures avancées comme les CNN et les RNN ont contribué à des progrès considérables dans des domaines notamment les systèmes intelligents pour analyser les données textuelles et visuelles de manière automatisée. Aujourd'hui, le Deep Learning s'est imposé comme une méthode centrale de l'IA, transformant des secteurs variés, de la médecine à la robotique, tout en continuant d'évoluer avec l'introduction de nouvelles architectures [52].

II.4.Les Raisons de l'essor du Deep Learning

Le Deep Learning a révolutionné le domaine de l'IA pour plusieurs raisons :

- Capacité à Apprendre des Représentations Complexes.
- Performances Supérieures.
- Adaptabilité [53].

II.5.Domains d'Application

Le DL a trouvé des applications dans un large éventail de domaines :

- Vision par Ordinateur : Utilisé à la reconnaissance des visages et d'images, à l'identification d'objets, à la segmentation des images, etc.
- Traitement du Langage Naturel (NLP) : utilisé dans le domaine de la rédaction de texte, de l'analyse de sentiments, de la traduction automatique, la reconnaissance vocale, etc.
- Santé : Détection de maladies à partir d'images médicales, prédiction des résultats cliniques, découverte de médicaments.
- Automobile : Développement de véhicules autonomes, systèmes avancés d'assistance à la conduite.
- Finance : Prévision des marchés financiers, détection de fraudes, gestion des risques [54].

II.6.Principe de Base

Les réseaux neuronaux artificiels (RNA) constituent l'épine dorsale du Deep Learning, reproduisant des mécanismes cognitifs humains au moyen de systèmes mathématiques.

Ces réseaux sont constitués de couches de neurones artificiels, où chaque couche traite les données d'une manière spécifique avant qu'elles ne soient transmises à la couche suivante.

Il convient de noter que le « profond » dans l'apprentissage en profondeur fait référence à la profondeur des couches d'un réseau neuronal. Un réseau composé de plus de trois couches, qui inclurait les entrées et la sortie, peut être considéré comme un algorithme d'apprentissage en profondeur [55]. Ceci est généralement représenté à l'aide du schéma de la figure ci-dessous.

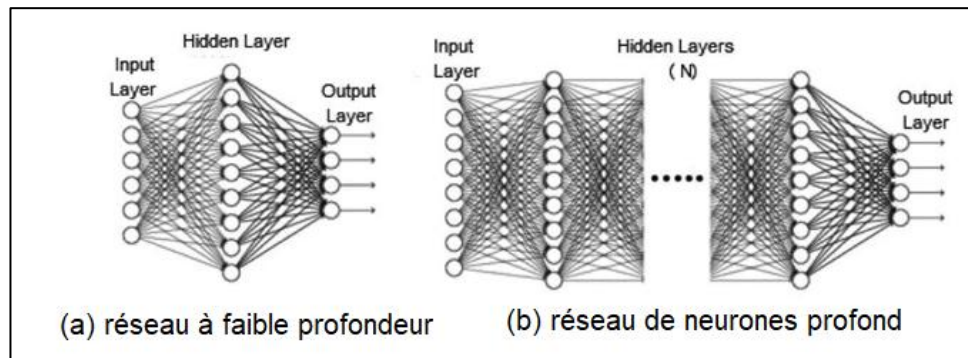


Figure II.3. Architecture d'un Réseau de Neurones Artificiels [56].

- La Couche d'entrée (Input Layer) : Reçoit les données brutes.
- Les Couches cachées (Hidden Layers) : Traitent les données en extrayant des caractéristiques de plus en plus abstraites.
- La Couche de sortie (Output Layer) : Produit le résultat final, tel qu'une étiquette de classe ou une prédiction.

Pour bien saisir le fonctionnement de cette approche puissante, il faut comprendre les composants fondamentaux du Deep Learning.

II.6.1. Neurone artificiel

Le neurone artificiel, également appelé perceptron, constitue l'élément fondamental du réseau de neurones artificiels. Il est conçu pour imiter le fonctionnement d'un neurone biologique (Voir Figure II.4) [57].

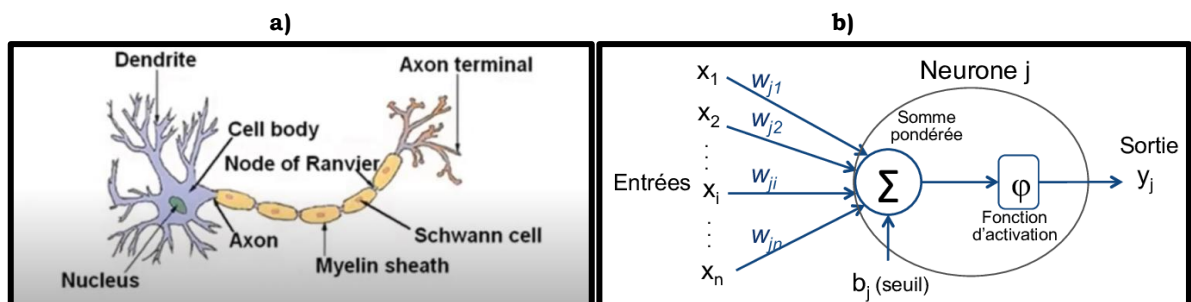


Figure II.4. a) Schéma d'un neurone biologique ; b) Schéma du neurone artificiel [58]

Dans un neurone artificiel, les multiples entrées et la sortie sont analogues aux dendrites et à l'axone d'un neurone biologique. Il reçoit des signaux d'entrée pondérés, effectue une somme pondérée de ces signaux, ajoute un biais, puis applique une fonction d'activation pour générer la sortie [58]. Mathématiquement, cela peut être formulé comme suit :

$$Z = \sum(W_i * X_i) + b \quad (1)$$

$$Y = f(Z) \quad (2)$$

Où Z représente l'activation pondérée ;

W Poids correspondants : $W = [W_1, W_2, \dots, W_n]$;

X Signaux d'entrée : $X = [X_1, X_2, \dots, X_n]$;

b Le biais ;

$f(Z)$ La fonction d'activation non linéaire.

Dans un réseau de neurones, les valeurs Y du neurone sont propagées aux neurones de la couche suivante [59].

II.6.2.Poids et Biais

II.6.2.a.Poids (Weights)

Des coefficients appliqués aux entrées des neurones. Ils sont ajustés durant l'entraînement pour minimiser l'erreur de prédiction [59].

II.6.2.b.Biais (Biases) :

Des valeurs ajoutées au résultat pondéré de chaque neurone avant l'application de la fonction d'activation. Ils permettent de décaler la fonction d'activation, aidant le réseau à mieux modéliser les données [59].

II.6.3.Fonctions d'activation

Introduisent une non-linéarité dans le réseau neuronal, dans le but de modéliser des relations complexes entre les données.

Les fonctions d'activation courantes incluent ReLU (Rectified Linear Unit), Sigmoid et Tanh [60].

II.6.3.a.Fonction Sigmoid

Mathématiquement, la fonction sigmoïde est définie comme suit :

$$f(Z) = \frac{1}{1+e^{-Z}} \quad (3)$$

Elle est fréquemment exploitée dans les couches cachées des réseaux neuronaux, et donne une sortie dans l'intervalle $[0,1]$, ce qui facilite son interprétation probabiliste, notamment dans des tâches de classification binaire [60]. Son graphique est le suivant :

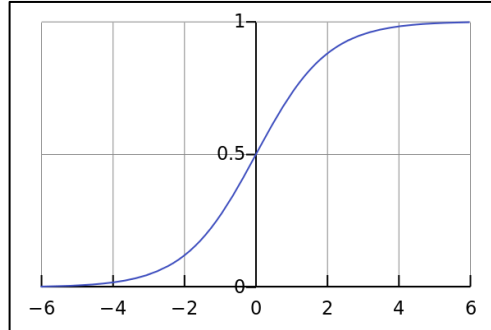


Figure II.5. Fonction Sigmoid

II.6.3.b.Fonction ReLU (Rectified Linear Unit)

Couramment appliquée comme fonction d'activation dans les architectures de Deep Learning [61]. La définition mathématique de la fonction ReLU est la suivante :

$$\text{ReLU}(X) = \max(0, X) \quad (4)$$

La figure II.6 illustre la fonction d'activation ReLU, mettant en évidence deux caractéristiques clés :

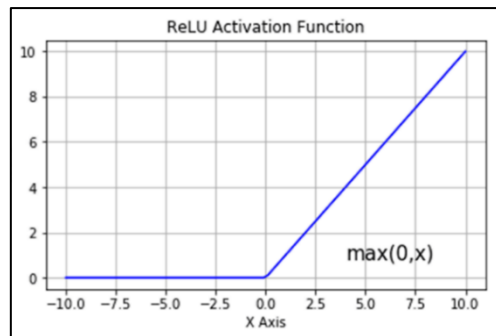


Figure II.6. Fonction ReLU (Rectified Linear Unit).

- **Linéaire pour les Entrées Positives :** Pour des valeurs d'entrée positives, la fonction ReLU renvoie directement l'entrée elle-même, exprimée mathématiquement par $(f(X) = X)$. Cela signifie qu'il y a une relation linéaire directe entre l'entrée et la sortie dans cette région.
- **Nulle pour les Entrées Négatives :** Pour des valeurs d'entrée négatives, la fonction ReLU renvoie toujours 0 $(f(X) = 0)$.

Cette caractéristique rend ReLU une fonction de seuil qui "coupe" toutes les valeurs négatives, produisant ainsi une sortie nulle.

II.6.3.c.Fonction Tanh

Tanh, qui signifie tangente hyperbolique, est une fonction mathématique couramment utilisée en apprentissage automatique.

Elle a le même principe que la fonction sigmoïde, mais elle cartographie les valeurs d'entrée dans une plage entre -1 et 1 au lieu de 0 et 1 [61]. Elle est définie mathématiquement comme suit :

$$\text{Tanh}(\alpha) = \frac{e^{\alpha} - e^{-\alpha}}{e^{\alpha} + e^{-\alpha}} \quad (5)$$

Son graphique est représenté par la figure II.7 :

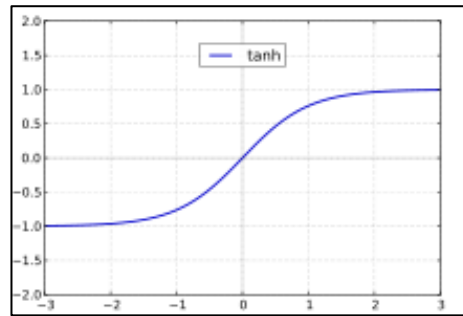


Figure II.7. Fonction tangente hyperbolique (Tanh)

Un exemple concret de la production d'une sortie peut être illustré par le neurone représenté dans la figure II.8.

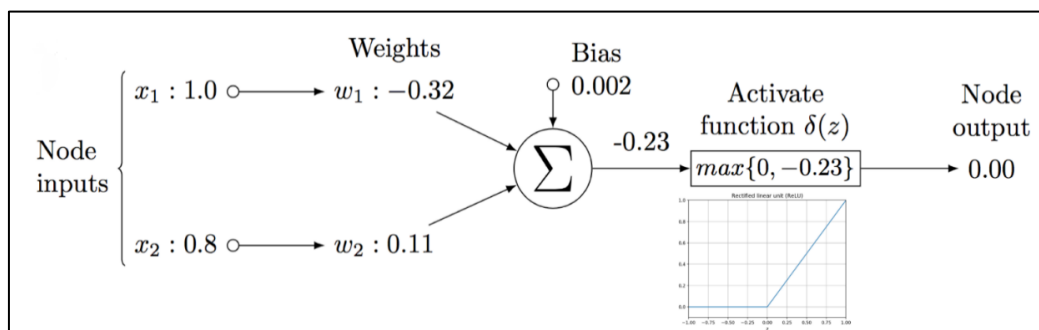


Figure II.8. Exemple de neurone artificiel avec fonction d'activation ReLU [62]

Ce neurone reçoit deux entrées, x_1 et x_2 , auxquelles sont associés des poids de -0,32 et 0,11 respectivement, ainsi qu'un biais de 0,002. La somme pondérée de ces entrées, c'est-à-dire la somme des produits de chaque entrée par son poids, s'élève à -0,23. Cette dernière est ensuite transmise à une fonction d'activation, ici la fonction ReLU (Rectified Linear Unit), renvoie une sortie de 0,00 car ReLU met à zéro toutes les valeurs négatives.

Cette sortie constitue la réponse finale de ce neurone. Cet exemple met en évidence l'influence des poids et du biais sur la sortie du neurone, ainsi que l'impact de la fonction d'activation dans la transformation de la somme pondérée en une sortie finale.

Comprendre ces mécanismes est fondamental pour saisir le fonctionnement d'un réseau de neurones, où chaque neurone interagit avec d'autres pour traiter les informations de manière complexe et produire une sortie globale [62].

II.6.4. Propagation Avant (Forward Propagation)

Flux des données au sein du réseau, depuis la couche d'entrée jusqu'à celle de sortie, pour générer une prédiction [63].

II.6.5. Rétropropagation (Backpropagation)

Méthode d'entraînement utilisée pour ajuster les poids et les biais (Figure II.9). Après chaque prédiction, l'erreur entre la sortie estimée et la valeur cible est calculée, puis propagée en arrière à travers le réseau pour mettre à jour les paramètres [63].

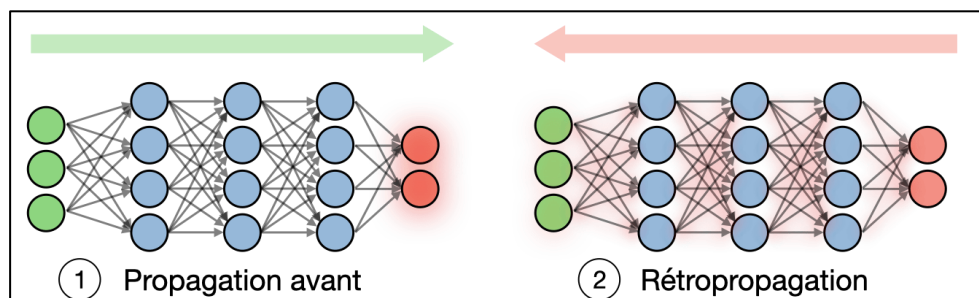


Figure II.9 Propagation Avant et Rétropropagation dans un Réseau Neuronal [64].

II.6.6. Fonction de Coût (Loss Function)

Quantifie la disparité entre les sorties estimées du réseau et les valeurs cibles. Plus cet écart est faible, meilleure est l'efficacité du modèle [65]. La figure II.10 illustre la fonction de coût dans l'apprentissage neuronal.

Divers types de fonctions de perte sont disponibles, chacune étant conçue pour répondre à des problèmes spécifiques, et à des types de données différents. Voici quelques exemples courants :

II.6.6.a. Erreur MSE (Erreur Quadratique Moyenne)

La MSE évalue la moyenne des carrés des différences entre les prévisions du réseau et les valeurs réelles des données d'entraînement. Elle est fréquemment utilisée pour résoudre des problèmes de régression [60].

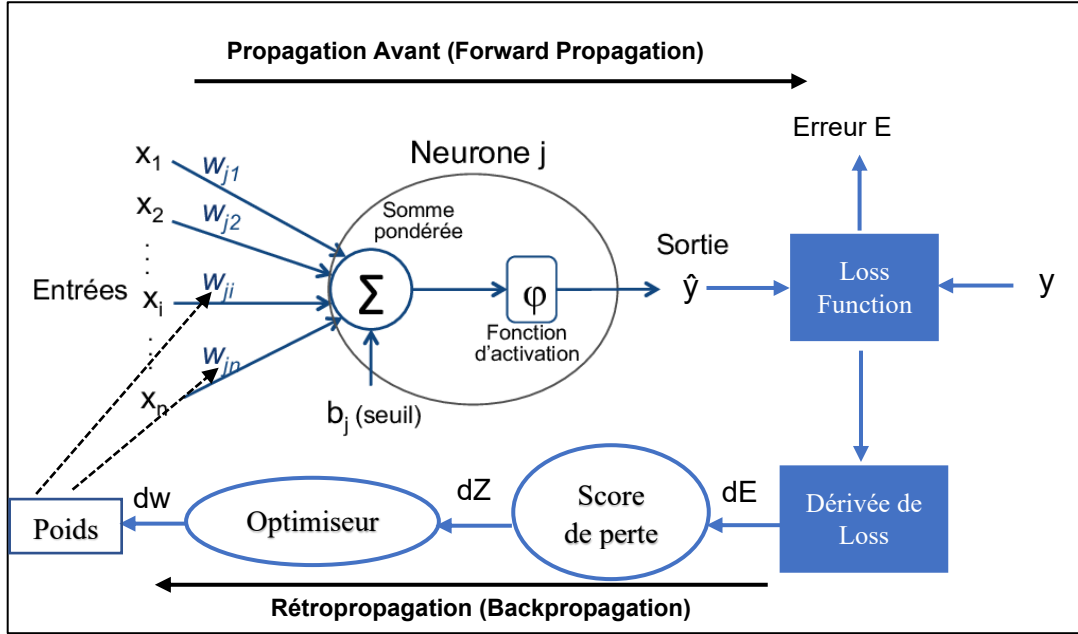


Figure II.10. Illustration de la Fonction de Perte dans l'Apprentissage des Réseaux de Neurones.

- Formule :

La MSE est calculée par le biais de la moyenne des carrés des écarts entre les valeurs prédites (\hat{Y}) et les valeurs réelles (Y) pour chaque instance d'entraînement (i) [60] :

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (6)$$

n : Nombre total d'exemples d'entraînement.

Y_i : Valeur réelle de l'exemple d'entraînement i .

\hat{Y}_i : Valeur prédite par le modèle pour l'exemple d'entraînement i .

II.6.6.b. Cross-Entropy Loss

Aussi connue sous le nom de "log loss", est souvent utilisée pour les problèmes de classification. Elle évalue la différence entre la probabilité prédite par le modèle et la distribution réelle des étiquettes de classe [66].

L'équation de la fonction de perte pour la classification binaire, souvent appelée fonction de perte logistique ou entropie croisée binaire, est la suivante:

$$Loss = - \left[Y * \log(Y_{pred}) + (1 - Y) * \log(1 - (Y_{pred})) \right] \quad (7)$$

- Y : La véritable étiquette de la classe. Y est égal à 1 pour la classe positive et 0 pour la classe négative.
- Y_{pred} : La prédiction du modèle, qui est une probabilité que l'exemple appartienne à la classe positive (compris entre 0 et 1).
- **Pour la classe positive ($Y = 1$) :**
 - La perte devient : $Loss = -\left[\log(Y_{pred})\right]$
 - Si Y_{pred} est proche de 1 (bonne prediction), la perte est faible.
 - Si Y_{pred} est proche de 0 (mauvaise prediction), la perte est élevée.
- **Pour la classe négative ($Y = 0$) :**
 - La perte devient : $Loss = -\log(1 - Y_{pred})$
 - Si Y_{pred} est proche de 0 (bonne prediction), la perte est faible.
 - Si Y_{pred} est proche de 1 (mauvaise prediction), la perte est élevée.

- **Perte totale pour une base de données (dataset)**

Pour une base de données avec N exemples, la perte totale est la moyenne des pertes individuelles [66] :

$$Loss_{total} = \frac{1}{N} \sum_{i=1}^N -[Y_i * \log(Y_{pred,i}) + (1 - Y_i) * \log(1 - (Y_{pred,i}))] \quad (8)$$

La figure II.11, montre deux courbes qui illustrent la fonction de perte logistique (Cross-Entropy) utilisée dans la classification binaire.

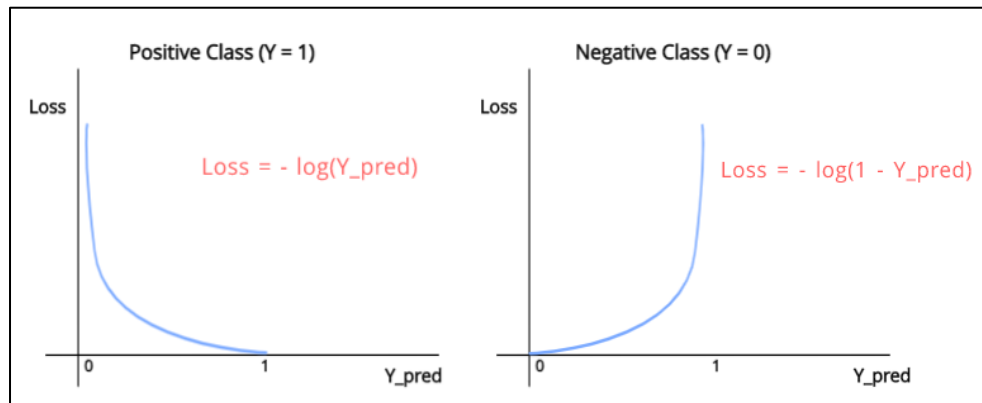


Figure II.11. Courbes de la fonction de perte logistique (Cross-Entropy)

II.6.6.c. Fonction de perte Hinge

Utilisée dans les problèmes de classification, notamment pour les machines à vecteurs de support (SVM), elle mesure la marge entre les prédictions du modèle et les vraies étiquettes [67].

La fonction Hinge Loss est définie comme suit :

$$L(Y, f(X)) = \max(0, 1 - Y \cdot f(X)) \quad (9)$$

Où:

- Y est l'étiquette de classe réelle (1 pour la classe positive, -1 pour la classe négative).
- $f(X)$ est la prédiction du modèle pour l'exemple X
- La fonction max renvoie le maximum entre 0 et $1 - Y \cdot f(X)$.

II.6.6.d. Fonction de perte Huber

Similaire à la MSE, mais moins sensible aux valeurs aberrantes. Elle est souvent utilisée dans les problèmes de régression où les données peuvent contenir du bruit [68].

La fonction de perte de Huber est définie comme suit :

$$L(Y, f(X)) = \begin{cases} \frac{1}{2}(Y - f(X))^2 & \text{si } |Y - f(X)| \leq \delta \\ \delta(|Y - f(X)| - \frac{1}{2}\delta) & \text{sinon} \end{cases} \quad (10)$$

Où :

- Y: est la valeur réelle de la cible,
- $f(X)$: est la prédiction du modèle pour l'exemple X
- δ : est le seuil qui détermine la transition entre la pénalité quadratique et linéaire.

La figure II.12 illustre de manière concrète l'application de la propagation avant (forward propagation) et de la rétropropagation (backpropagation) dans le contexte de l'apprentissage automatique. Au sein d'un réseau neuronal organisé en une couche d'entrée, une couche cachée et une couche de sortie, on peut suivre le cheminement des données d'entrée jusqu'à la génération de la sortie finale. Cette figure met également en évidence, le calcul de la fonction de coût (loss function) et l'utilisation des gradients pour ajuster les poids du réseau, étape importante dans l'optimisation et l'entraînement des réseaux de neurones. Ce processus permet d'améliorer les performances du modèle.

Cet exemple pratique permet de mieux appréhender les principes fondamentaux de l'apprentissage profond et leur mise en œuvre concrète dans des scénarios réels.

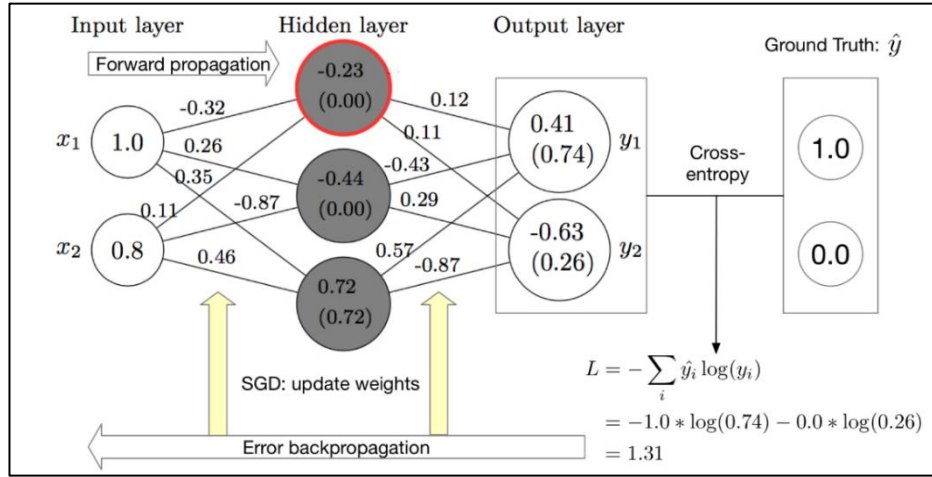


Figure II.12. Exemple concret de propagation avant et rétropropagation dans un réseau neuronal [62].

II.6.7. Optimiseurs

Algorithmes utilisés pour minimiser la fonction de coût en ajustant les poids et les biais. Les optimiseurs courants incluent la descente de gradient stochastique (SGD), Adam et RMSprop.

II.6.7.a. Approche SGD (La descente de gradient stochastique)

Est une méthode d'optimisation utilisée pour minimiser une fonction de coût en ajustant les paramètres du modèle. Elle est particulièrement utile dans l'apprentissage automatique et l'apprentissage profond pour optimiser les réseaux de neurones et autres modèles complexes [69].

▪ Principe de Base

L'objectif de la descente de gradient stochastique est de trouver les paramètres θ (theta) qui minimisent une fonction de coût $J(\theta)$. Plutôt que de calculer le gradient de $J(\theta)$ en utilisant l'ensemble complet des données (comme dans la descente de gradient par lot complet), SGD met à jour les paramètres en utilisant un seul échantillon de données à la fois [70].

▪ Formule de Mise à Jour

La mise à jour des paramètres θ pour un exemple d'entraînement $(X^{(i)}, Y^{(i)})$ est donnée par :

$$\theta = \theta - \eta \nabla J(\theta; X^{(i)}, Y^{(i)}) \quad (11)$$

Où : η : est le taux d'apprentissage,

$\nabla J(\theta; X^{(i)}, Y^{(i)})$ représente le gradient de la fonction de coût J par rapport aux paramètres θ .

Cette mise à jour permet d'ajuster les paramètres du modèle en fonction de l'erreur observée pour chaque exemple d'entraînement [70].

II.6.7.b. Approche Adam

Adam est un algorithme d'optimisation pour l'apprentissage automatique qui combine les avantages de AdaGrad et RMSProp. Il adapte dynamiquement le taux d'apprentissage pour chaque paramètre du modèle en utilisant des estimations exponentiellement pondérées des premiers et deuxièmes moments des gradients [71].

Les étapes clés de l'algorithme Adam sont :

- **Mise à jour des moments** [69] : A chaque étape t , Adam calcule la moyenne mobile des gradients m_t et la moyenne mobile des carrés des gradients v_t .

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (12)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (13)$$

m_t : C'est la moyenne mobile des gradients à l'étape t , qui lisse les gradients sur plusieurs itérations pour réduire le bruit dans les mises à jour des poids [72].

v_t : C'est la moyenne mobile des carrés des gradients à l'étape t , qui estime la variance des gradients pour adapter le taux d'apprentissage selon leur taille [72].

β_1 : Coefficient de décroissance exponentielle pour m_t , contrôlant la contribution des gradients passés, généralement fixé à 0,9 [72].

β_2 : Coefficient de décroissance exponentielle pour v_t , contrôlant la contribution des carrés des gradients passés, généralement fixé à 0,999 [72].

g_t : Gradients calculés à l'étape t pour les paramètres du modèle, représentant la direction et la magnitude du changement nécessaire pour minimiser la fonction de perte [72].

Correction du biais : En raison de l'initialisation à zéro de m_0 et v_0 , les estimations initiales sont biaisées. Adam corrige ce biais en ajustant m_t et v_t [72].

$$\widehat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (14)$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \quad (15)$$

- **Ajustement des paramètres :** Finalement, les paramètres du modèle sont réajustés en utilisant les moments corrigés et un taux d'apprentissage α [72].

$$\theta_{t+1} = \theta_t - \frac{\alpha \hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (16)$$

II.6.7.c. Approche RMSprop

Est un algorithme d'optimisation pour l'apprentissage automatique qui modifie dynamiquement le taux d'apprentissage pour chaque paramètre d'un réseau en fonction des gradients précédents. Il maintient une estimation décadente de la moyenne des carrés des gradients passés pour chaque paramètre et utilise cette estimation pour normaliser le gradient actuel [73].

Les étapes clés de l'algorithme RMSprop sont :

- **Mise à jour de l'estimation des carrés des gradients :** A chaque étape t , RMSprop met à jour une estimation v_t de la moyenne des carrés des gradients passés pour un paramètre θ_t [73].

$$v_t = \beta v_{t-1} + (1 - \beta) g_t^2 \quad (17)$$

- **Mise à jour des paramètres :** Les paramètres sont ensuite mis à jour en utilisant le gradient normalisé par la racine carrée de v_t et un taux d'apprentissage α [73].

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{v_t + \epsilon}} g_t \quad (18)$$

II.6.8. Epoche et Lot (Epoch and Batch)

II.6.8.a. Epoche (Epoch)

Une passe complète à travers l'ensemble de données d'apprentissage [74].

II.6.8.b. Lot (Batch)

Un sous-ensemble des échantillons utilisé pour une itération de réajustement des poids. L'entraînement par lots peut améliorer la stabilité et l'efficacité de l'entraînement [74].

II.7. Limites de l'apprentissage profond

Certaines difficultés sont fréquemment rencontrées lors de l'utilisation de l'apprentissage profond. Nous énumérons certaines d'entre elles.

II.7.1. Surajustement & sous-ajustement

Le surajustement (overfitting) et le sous-ajustement (underfitting) constituent des défis majeurs en apprentissage automatique, affectant à la fois les performances et la capacité de généralisation des modèles. Le surajustement se produit lorsqu'un modèle apprend trop bien les données d'entraînement, y compris le bruit, ce qui conduit à de mauvaises performances sur de nouvelles données [75]. A l'inverse, le sous-ajustement survient lorsque le modèle est trop limité pour capturer les motifs sous-jacents des données, entraînant une faible précision aussi bien sur les ensembles d'entraînement que de test [76]. La Figure II.13 illustre le concept d'underfitting et d'overfitting dans le contexte de l'apprentissage automatique.

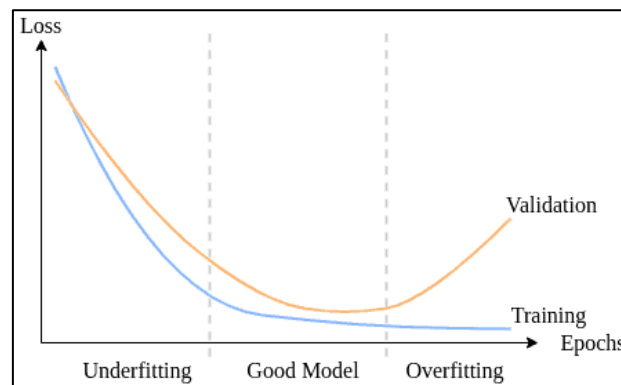


Figure II.13. Illustration de l'overfitting et l'underfitting

- **Courbes :**

- La courbe bleue montre la perte sur l'ensemble d'entraînement, qui diminue au fur et à mesure que le modèle apprend.
- La courbe orange représente la perte sur l'ensemble de validation, qui permet de mesurer la capacité du modèle à généraliser.

- **Zones :**

- **Sous-apprentissage** : Au début, la perte d'entraînement est élevée, signe que le modèle n'a pas assez appris.
- **Bon modèle** : Lorsque les pertes d'entraînement et de validation sont proches, le modèle généralise correctement.
- **Surapprentissage** : Quand la perte de validation augmente alors que celle d'entraînement continue de baisser, le modèle capture trop de bruit et perd en généralisation.

Le point optimal se situe là où la perte de validation est la plus basse, avant l'apparition du surapprentissage, ce qui peut être détecté via des techniques

comme Early stopping. Pour limiter ces problèmes, plusieurs méthodes ont été proposées :

➤ **Early stopping** : est une technique en apprentissage automatique qui vise à prévenir le surapprentissage en arrêtant l'entraînement d'un modèle lorsque sa performance sur un ensemble de validation commence à se dégrader. Le processus consiste à surveiller l'erreur de validation à chaque itération, et lorsque celle-ci augmente, cela signale que le modèle surapprend. L'entraînement est donc interrompu pour préserver le modèle dans son état optimal. En plus de prévenir le surapprentissage, cette approche permet également de réduire les coûts computationnels en évitant des itérations inutiles [77].

➤ **Dropout** : Le dropout est une technique de régularisation utilisée dans les réseaux neuronaux pour prévenir le surajustement en désactivant aléatoirement un pourcentage de neurones pendant l'entraînement, empêchant ainsi les paramètres de s'adapter trop fortement aux données d'entraînement. [77][78].

➤ **Batch Normalization**

Normalise les activations, c'est-à-dire qu'elle les ajuste pour qu'elles aient une moyenne et une variance fixe. Cela rend l'entraînement plus stable et rapide, et aide aussi à améliorer les performances du modèle sur de nouvelles données [79].

➤ **Augmentation de données**

Augmente explicitement la quantité de données d'entraînement pour mieux couvrir la distribution réelle.

II.7.2. Le manque de données

Est un défi majeur dans l'application de l'apprentissage profond, car ces modèles nécessitent beaucoup de données pour apprendre efficacement. Voici trois stratégies pour surmonter ce problème :

➤ **Transfert d'apprentissage (Transfer Learning)** : Utiliser des données provenant de tâches similaires pour aider à apprendre une meilleure fonction de mappage et une meilleure représentation des entrées initiales, sans augmenter directement la quantité de données réelles. On peut également utiliser un modèle pré-entraîné sur une tâche similaire et l'affiner en utilisant les dernières couches et un petit ensemble de données réelles [80].

➤ **Augmentation de données** : Particulièrement utile pour les données d'image (par exemple, rotation, miroir, translation).

II.7.3. Les données déséquilibrées

Les données déséquilibrées surviennent lorsqu'une classe est beaucoup plus représentée que d'autres dans une base de données, ce qui peut entraîner un biais dans l'apprentissage automatique. Cela rend les modèles moins performants sur les classes minoritaires. Pour résoudre ce problème, on utilise des techniques comme le sur-échantillonnage des classes sous-représentées, le sous-échantillonnage des classes majoritaires, ou des algorithmes adaptés aux données déséquilibrées. Ces stratégies visent à améliorer la capacité du modèle à bien prédire toutes les classes [81].

II.8. Algorithmes d'apprentissage

Les algorithmes d'apprentissage servent à ajuster les poids et les biais des connexions d'entrée des neurones pendant l'entraînement du réseau. Ils se répartissent en trois catégories :

- Supervisé (supervised learning) ;
- Non supervisé (unsupervised learning) ;
- Par renforcement (reinforcement learning).

Les différents types d'apprentissage automatique sont présentés dans la figure II.14.

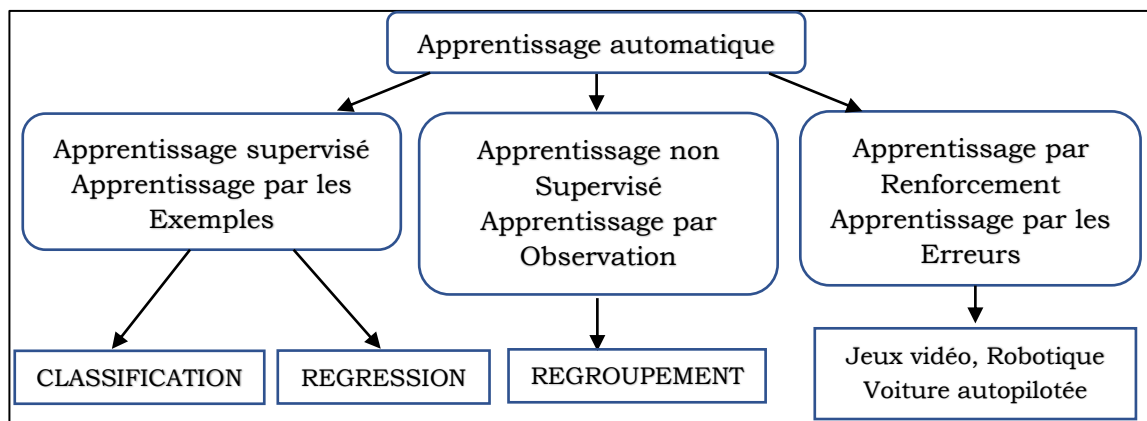


Figure II.14. Taxonomie des méthodes d'apprentissage automatique.

II.8.1. L'apprentissage supervisé

Est une méthode où les modèles de Deep Learning sont entraînés à partir de données étiquetées, c'est-à-dire des données d'entraînement où chaque exemple est associé à une sortie connue.

Les réseaux de neurones profonds sont ajustés en minimisant l'écart entre les prédictions du modèle et les sorties attendues, permettant ainsi au modèle de généraliser et de faire des prédictions précises sur de nouvelles données.

Après l'entraînement, les performances du système sont mesurées sur un ensemble d'exemples différent appelé ensemble de test. Cela permet de tester la capacité de généralisation de la machine, c'est-à-dire sa capacité à produire des réponses. Cette approche est largement utilisée dans des domaines tels que la classification d'images, la détection d'objets et la prédiction de valeurs. Cependant, il nécessite des ensembles de données annotées pour l'apprentissage, ce qui peut être coûteux et laborieux à obtenir [82]. La figure II.15 illustre le principe de l'apprentissage supervisé.

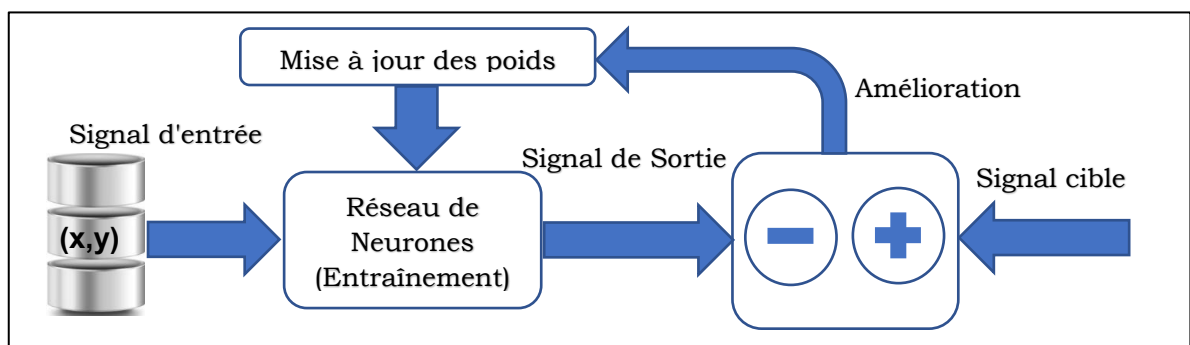


Figure II.15. Principe de l'apprentissage supervisé (SL) [83]

II.8.2. L'apprentissage non supervisé (unsupervised learning)

Est une technique de l'apprentissage automatique qui se concentre sur l'analyse et la compréhension de données brutes, sans l'utilisation de labels ou de réponses prédéfinies. Un ensemble de données de formation est présenté au système au niveau de la couche d'entrée, et les poids de connexion du réseau sont ajustés au fil du temps par une forme de compétition entre les nœuds de la couche de sortie. Le nœud ayant la valeur la plus élevée est retenu comme candidat principal, et ses poids sont ensuite mis à jour pour mieux représenter les données. Ce processus itératif permet aux algorithmes d'identifier des structures internes, des motifs cachés et des relations au sein des données, accomplissant ainsi des tâches telles que la classification non supervisée, la réduction de la dimensionnalité [84].

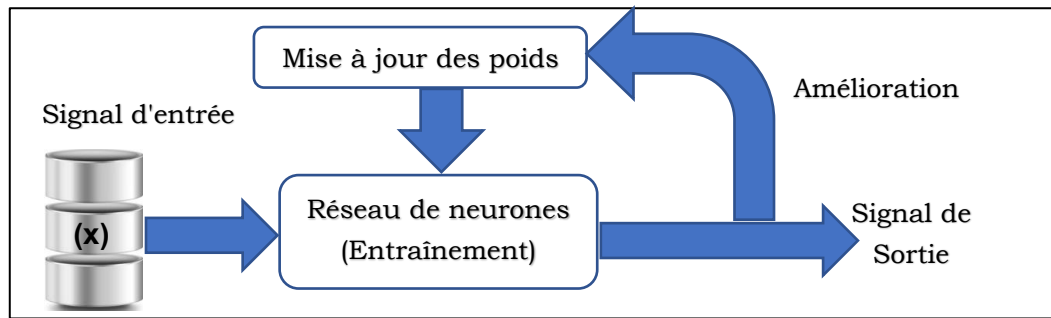


Figure II.16. Apprentissage non supervisé. [83]

II.8.3. Apprentissage par renforcement (RL)

Le RL, également appelé apprentissage graduel, imite d'une certaine manière le comportement d'adaptation des humains qui interagissent avec un environnement physique donné. Les connexions du réseau sont modifiées en fonction des informations de retour fournies au réseau par son environnement. En cas de réponse correcte, les connexions correspondantes menant à cette sortie sont renforcées, sinon elles sont affaiblies.

L'apprentissage par renforcement est appliqué dans divers domaines tels que les jeux, la robotique et l'optimisation de systèmes, où l'agent doit apprendre par essais et erreurs dans un environnement complexe et incertain [85][86].

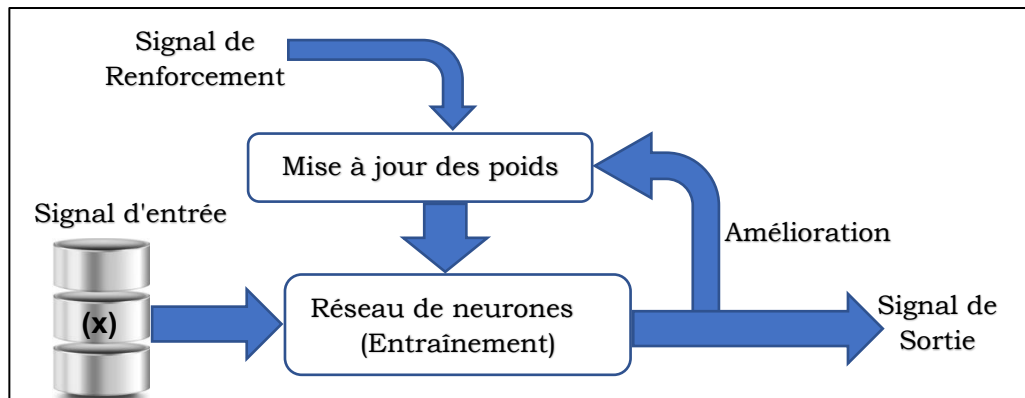


Figure II.16. Apprentissage par renforcement. [83]

II.9. Conclusion

Ce chapitre a établi les bases théoriques de l'apprentissage profond en détaillant les réseaux neuronaux, les fonctions d'activation, les algorithmes d'optimisation, supervisés et non supervisés. Ces notions essentielles forment un socle indispensable pour les contributions des chapitres suivants.

Chapitre III

CNN pour la classification des images

III.1. Introduction

La classification, dans le contexte du Deep Learning, est un domaine dynamique en pleine expansion qui consiste à attribuer des étiquettes ou des catégories aux images en fonction de leur contenu. Ce processus s'appuie sur des algorithmes avancés, notamment les réseaux de neurones convolutifs (CNN), qui se révèlent particulièrement efficaces pour extraire des caractéristiques des images. Cette approche est importante pour le développement de technologies intelligentes et autonomes [87].

Le processus typique de classification comprend les étapes suivantes :

- **Préparation des données :** Collecte et prétraitement des images étiquetées pour l'entraînement.
- **Entraînement :** Le modèle apprend à partir des données d'entraînement, ajustant ses poids pour minimiser l'erreur entre ses prédictions et les étiquettes réelles.
- **Validation :** Evaluation du modèle sur un ensemble de validation pour optimiser les hyperparamètres.
- **Test :** Mesure de la performance du modèle sur un ensemble de test non vu.

La classification est utilisée dans de nombreux domaines, tels que la reconnaissance d'objets, la classification de textes et l'analyse de sentiments. Elle joue un rôle essentiel dans la compréhension des scènes visuelles et l'interprétation des données.

III.2. Réseaux de neurones convolutifs (CNN)

Les réseaux de neurones convolutifs, souvent désignés par l'acronyme CNN (Convolutional Neural Networks), sont des algorithmes populaires dans le domaine de l'apprentissage profond qui permettent de trouver automatiquement les caractéristiques les plus importantes sans intervention humaine [88]. Ils représentent une classe de réseaux de neurones particulièrement reconnue pour son efficacité dans le traitement des données structurées en grille, telles que les images.

Cette fonctionnalité permet une classification efficace dans divers domaines, notamment l'imagerie médicale et la détection d'objets [89]. La figure III.1 représente l'architecture de base d'un réseau de neurones convolutifs.

III.3. Applications des CNN

- **Reconnaissance d'images** : Identification d'objets ou de personnes dans des images.
- **Segmentation d'images** : Attribution de labels à des pixels dans une image.
- **Analyse vidéo** : Détection d'actions ou de mouvements dans des séquences vidéo.
- **Traitement de texte** : Utilisation dans certaines applications de traitement du langage naturel.

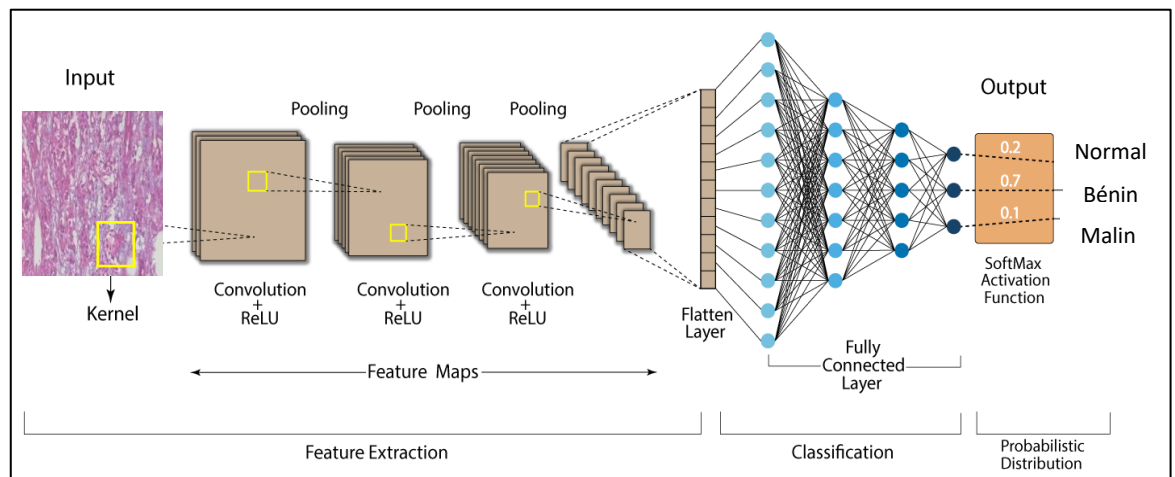


Figure III.1 : Architecture de base d'un réseau de neurones [88].

III.4. Caractéristiques principales des CNN

- **Noyau de Convolution** : Un noyau de convolution est une petite matrice de poids qui glisse sur l'image d'entrée. Chaque noyau est conçu pour détecter des motifs spécifiques, comme des bords, des textures ou d'autres caractéristiques.

Chaque noyau de convolution possède trois dimensions :

- **Longueur (L)** : Correspond à la dimension verticale du noyau.
- **Largeur (W)** : Correspond à la dimension horizontale du noyau.
- **Profondeur (D)** : Représente le nombre de canaux d'entrée. Par exemple, pour une image en couleur, la profondeur serait de 3 (rouge, vert, bleu).

- Taille du Noyau : La taille du noyau est généralement exprimée sous la forme $L \times W$. Les tailles courantes de noyaux de convolution incluent 3×3 ou 5×5 . Le choix de la taille du noyau dépend de la tâche à réaliser et de la complexité des caractéristiques à extraire.

III.5. Convolution

Dans un réseau de neurones convolutifs (CNN), les couches de convolution extraient des caractéristiques de plus en plus abstraites à mesure que l'on monte en profondeur.

Les couches inférieures capturent des éléments simples comme des textures, lignes et bords, tandis que les couches supérieures identifient des caractéristiques plus complexes. Chaque couche contient des noyaux de convolution, qui sont des matrices de poids apprenables, souvent de taille 3×3 , 5×5 ou 7×7 [90].

Le noyau est appliqué à l'image en effectuant une opération de convolution. Cela implique de multiplier les valeurs du noyau par les valeurs des pixels de l'image sous-jacente et d'additionner les résultats pour obtenir une valeur unique.

Ce processus est répété en déplaçant le noyau sur toute l'image, ligne par ligne et colonne par colonne. Chaque position du noyau produit une valeur dans la carte de caractéristiques [91].

La figure III.2 explique le processus de convolution dans un réseau de neurones convolutifs (CNN).

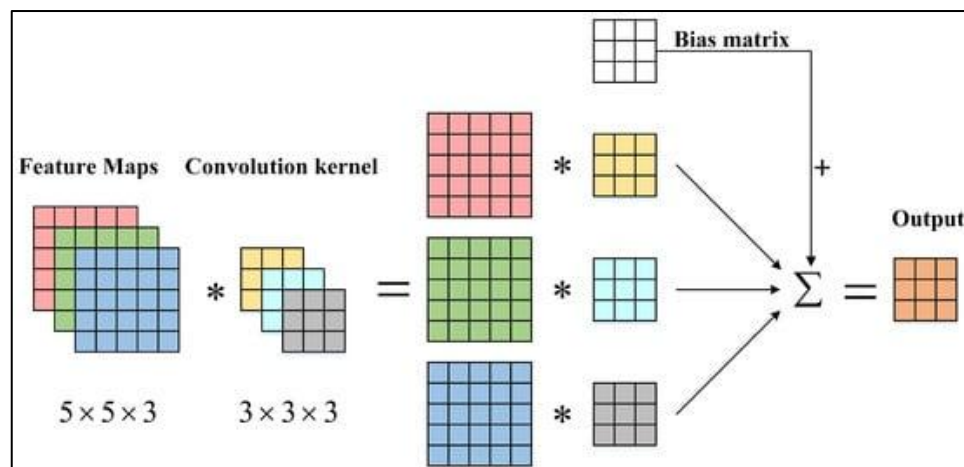


Figure III.2 Processus de convolution [91]

En règle générale, les dimensions des caractéristiques d'entrée sont de $H \times W \times C$ (hauteur H , largeur W et nombre de canaux C).

Chaque noyau de convolution a une taille de $K \times K \times C$, ce qui implique que le nombre de noyaux doit correspondre au nombre de canaux d'entrée. Le traitement des données au sein de la couche de convolution peut être approximativement décrit par l'équation suivante :

$$feature_surfaceout = f(\sum_{i=1}^3 M_i * W_i + M) \quad (19)$$

Où :

- M_i : représente une feature map des feature maps d'entrée.
- W_i : est la matrice des poids du noyau de convolution.
- M : est une matrice de biais.
- $f(.)$: est une fonction d'activation non linéaire.
- $feature_surfaceout$: est la feature map de sortie résultante.

Par exemple, dans la figure III.3, nous avons trois filtres, un pour chaque canal RVB. Pour le filtre rouge en haut à gauche, la valeur de sortie 148 est obtenue en effectuant le calcul suivant sur les valeurs des pixels sous le filtre : $156 * (-1) + 155 * (-1) + 156 * (1) + 153 * (0) + 154 * (1) + 157 * (-1) + 149 * (0) + 151 * (1) + 155 * (1)$. De même, pour le filtre vert, la valeur -8 est calculée en additionnant 167, 164, (-165), (-168), ((160) et (-166) (avec les autres valeurs à 0). Enfin, pour le filtre bleu, la valeur 646 est obtenue par les calculs $162 + 163 + 161 + 156 - 158 + 162$.

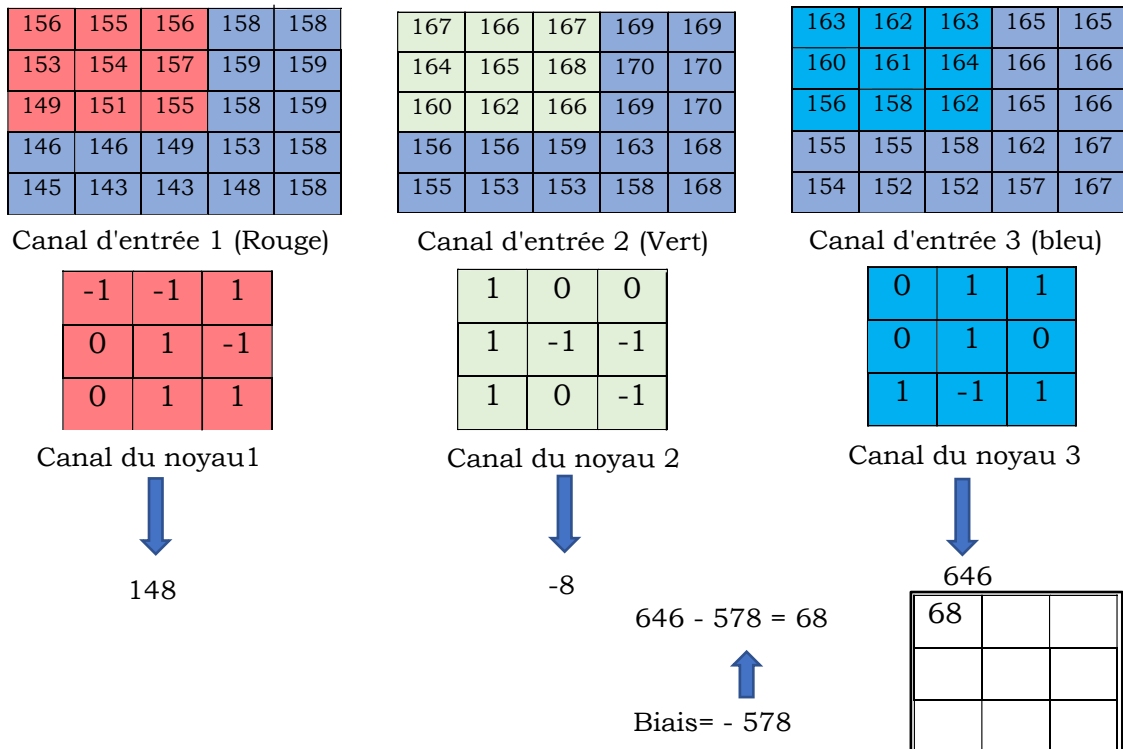


Figure III.3 : opération de convolution sur une matrice d'image 5X5X3 avec un noyau 3X3X3

Remarques

- La valeur du **biais** peut être déterminée par essais et erreurs. Elle peut être ajustée pour optimiser les performances de la transformation de l'image.
- Lorsqu'une convolution est appliquée à une image, un filtre de taille 3x3 transforme les valeurs de 9 pixels en une seule valeur, ce qui réduit la taille de l'image. Pour conserver la taille d'origine de l'image, on peut ajouter des zéros autour de celle-ci, une technique appelée padding. Par exemple, en utilisant une image de 5x5x1 avec un filtre 3x3, la sortie serait une image de 3x3x1. En ajoutant un contour de zéros d'un pixel, l'image devient 6x6x1, et la sortie de convolution reste de 5x5x1. Ainsi, le padding permet de maintenir la taille d'origine de l'image, ce qui est souvent avantageux. Le nombre de zéros ajoutés dépend de la valeur du padding : un padding de 0 signifie aucun ajout, tandis qu'un padding de 1 signifie l'ajout d'un contour de zéros, et ainsi de suite.
- Lors de la convolution, le filtre est déplacé sur l'image, et le décalage effectué à chaque étape est appelé stride. Ce dernier peut être ajusté selon les besoins : avec un stride de 1, le filtre se déplace d'un pixel à la fois, tandis qu'avec un stride de 2, il se déplace de deux pixels à la fois. En général, le choix du stride est fait de manière que la taille de l'image de sortie soit un entier et non un nombre décimal.
- La taille de l'image de sortie d'une convolution peut être calculée à l'aide de la formule suivante, en fonction de la taille de l'image d'entrée (W), de la taille du filtre (F), du padding (P) et du stride (S) [92] :

$$\text{Taille de l'image de sortie d'une convolution} = 1 + \frac{W-F+2P}{S} \quad (20)$$

Prenons un exemple pour illustrer le calcul de la taille de l'image de sortie d'une convolution :

- Soit une image d'entrée de taille 5x5, un filtre de taille 3x3, un padding de 0 et un stride de 1. En appliquant la formule :

$$\text{Taille de l'image de sortie d'une convolution} = 1 + \frac{5 - 3 + 2 * (0)}{1}$$

On obtient donc une image de sortie de taille 3x3.

III.6. Activation

Après une opération de convolution, il est courant d'appliquer une fonction d'activation, pour introduire de la non-linéarité dans le réseau de neurones,

ce qui permet d'établir une relation fonctionnelle entre les entrées et les sorties, ce qui améliore significativement les performances du réseau [93]. La figure III.4 présente des fonctions d'activation telles que Sigmoid et Tanh, appelées non-linéarités saturantes, car elles atteignent des valeurs limites (0 ou 1 pour Sigmoid, -1 ou 1 pour Tanh) lorsque les entrées sont très grandes ou très petites. Pour résoudre ces limitations, des fonctions non-saturantes comme ReLU, Leaky ReLU, PReLU, RReLU et ELU ont été introduites [94].

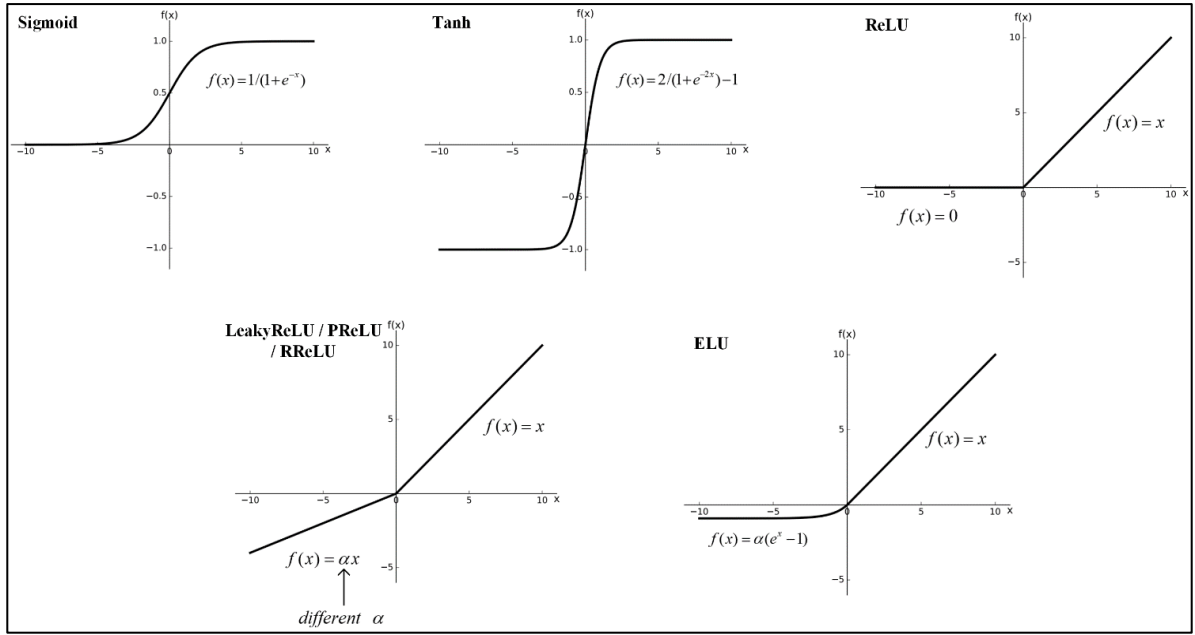


Figure III.4 : Fonctions d'Activation dans les Réseaux de Neurones.

III.7. Sous-échantillonnage (Pooling)

La couche de pooling, généralement placée après la couche de convolution, joue un rôle essentiel dans la réduction de la dimensionnalité des données tout en conservant les caractéristiques importantes. Elle effectue un sous-échantillonnage qui diminue le nombre de connexions dans la couche de convolution, allégeant ainsi la charge de calcul du réseau [93]. Pour comprendre comment les dimensions changent lors des opérations de convolution, on peut utiliser la formule suivante pour calculer la taille de la sortie :

$$O = \left\lceil \frac{(i-k)}{s} + 1 \right\rceil \quad (21)$$

Où O représente la taille de la sortie, i la taille de l'entrée, k la taille du noyau de convolution et s le stride. Par exemple, si l'entrée mesure 32 pixels et que l'on utilise un noyau de 5 pixels avec un stride de 1, la taille de la sortie sera de 28 pixels. Cette réduction de dimension permet d'atteindre une invariance d'échelle, rendant le modèle plus robuste aux variations de taille et de position des objets dans les images. Les méthodes de pooling les plus courantes sont le pooling moyen et le pooling maximal, bien que d'autres variantes, comme le pooling Lp, stochastique ou pyramidal, soient aussi utilisées pour mieux éviter le surajustement des réseaux [95].

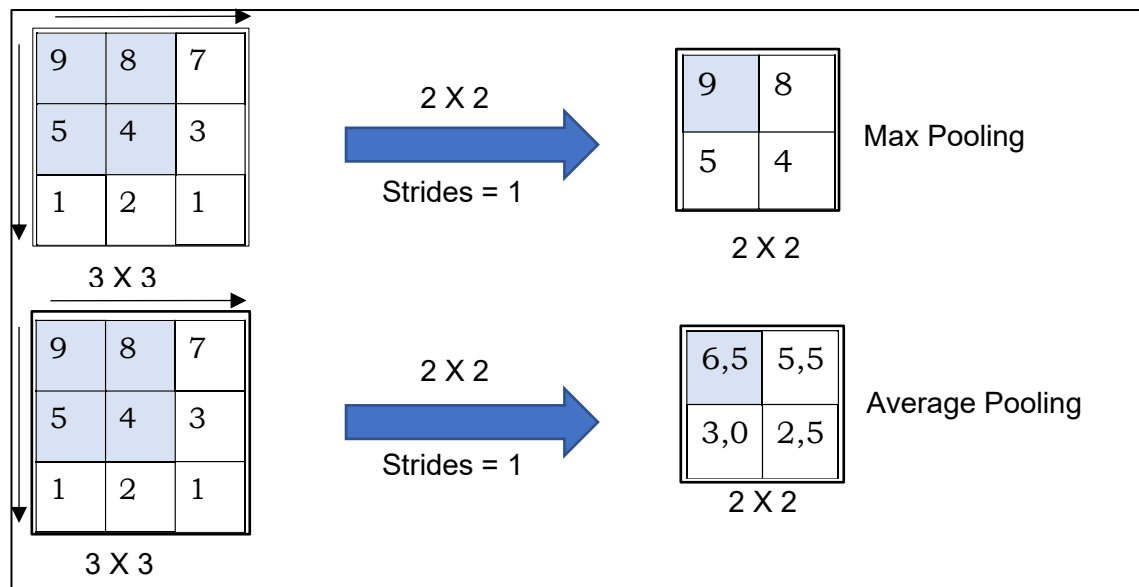


Figure III.5 : Exemples de Max Pooling et Average Pooling dans une couche de réseau de neurones.

III.8. La couche d'aplatissement (Flatten)

La couche Flatten dans un réseau de neurones sert à convertir des données d'entrée multidimensionnelles, comme les images, en un vecteur unidimensionnel. Après plusieurs opérations de convolution et de pooling qui conservent une structure en plusieurs dimensions (par exemple, une image 28x28x1 devient une carte de caractéristiques 28x28), la couche Flatten déplie ces données en un vecteur (28x28 devient 784 éléments). Cette transformation est essentielle pour préparer les données pour les couches denses (entièrement connectées), qui nécessitent des entrées sous forme de vecteurs pour effectuer des tâches de classification ou de régression, tout en conservant toutes les informations extraites par les couches précédentes [96].

III.9. Couche entièrement connectée (FC)

La couche entièrement connectée (FC) se situe généralement après les couches convolutives et de mise en commun dans un réseau de neurones convolutif (CNN). Elle relie tous les neurones entre ses couches et agit comme un classificateur, intégrant les informations locales extraites par la convolution et le pooling pour fournir la catégorie de l'image. Elle contient plusieurs couches cachées qui capturent des caractéristiques complexes. Le nombre de neurones en sortie correspond au nombre de catégories, et la sortie est souvent normalisée via une régression softmax avant la fonction de perte. Pour éviter le surajustement, des techniques comme la régularisation L2 et le dropout sont souvent utilisées [97],[98].

III.10. Fonction de perte

La classification finale dans une architecture de réseau de neurones convolutionnel (CNN) s'effectue à travers la couche de sortie, qui est généralement la dernière couche des couches entièrement connectées (FC). Les performances du CNN dépendent de diverses fonctions de perte, conçues pour des tâches visuelles variées telles que la classification d'images, la reconnaissance faciale et la détection d'objets. La fonction de perte la plus fréquemment utilisée est la combinaison de Softmax et de Cross-Entropy, bien qu'il existe plusieurs versions améliorées, comme center-loss, LSoftmax, A-Softmax, AM-Softmax et PEDCC loss, qui jouent un rôle important dans différentes applications visuelles [99].

III.11. L'optimiseur

Un optimiseur est un algorithme essentiel dans l'entraînement des réseaux de neurones, car il ajuste les paramètres du modèle pour minimiser la fonction de perte. Il met à jour les poids en réponse aux erreurs calculées lors de la rétropropagation. Parmi les optimiseurs courants, on trouve le Stochastic Gradient Descent (SGD) et ses variantes comme Adam, qui adapte les taux d'apprentissage pour une meilleure efficacité. D'autres, comme RMSprop et Adagrad, ajustent également les taux d'apprentissage en fonction des gradients passés, améliorant la stabilité et la vitesse de convergence. Le choix de l'optimiseur influence grandement la performance du modèle [100].

III.12. Les modèles classiques du CNN

III.12.1. LeNet Network

La figure III.6 montre l'architecture du modèle LeNet-5, conçu par Lecun et al en 1998 pour la classification de chiffres manuscrits.

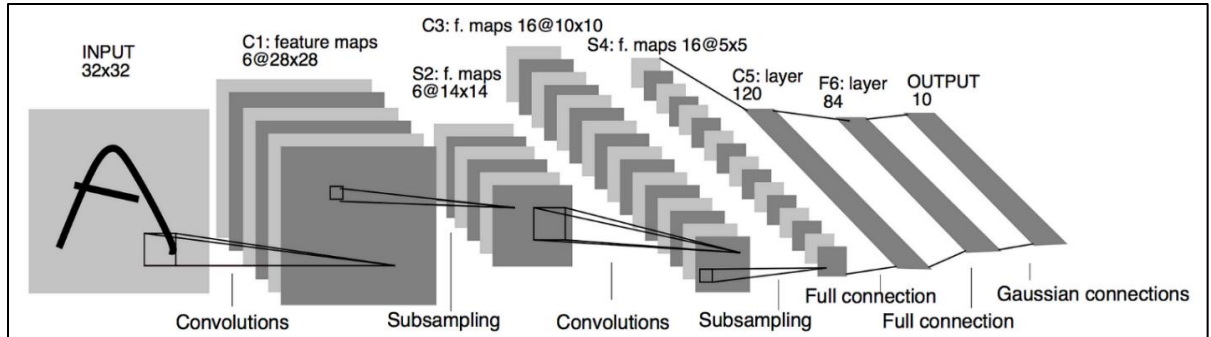


Figure III.6 : Architecture de LeNet-5 telle que publiée dans l'article original

Ce modèle surpassait les méthodes de l'époque et marquait la première utilisation de la rétropropagation dans les réseaux de neurones convolutifs (CNN). LeNet-5 est constitué de 7 couches et environ 60 000 paramètres. Il est divisé en deux parties principales : une zone de convolution et une zone entièrement connectée (fully connected). La zone de convolution, située à gauche, applique des filtres convolutifs suivis de couches de pooling, tandis que la zone entièrement connectée, à droite, intègre et classe les caractéristiques extraites à l'aide de trois couches qui contiennent respectivement 120, 84 et 10 neurones. Le modèle utilise une fonction d'activation sigmoïde et un classificateur softmax en sortie. Bien qu'efficace sur la base de données MNIST, LeNet-5 montrait des limites sur des ensembles de données plus vastes en raison de sa complexité [101].

III.12.2. Réseau AlexNet

En 2012, AlexNet, développé par Krizhevsky et al., a révolutionné la vision par ordinateur en remportant le concours ILSVRC14 (ImageNet LargeScale Visual Recognition Challenge 2014) avec une avance significative, démontrant que les caractéristiques apprises par un modèle surpassent celles conçues manuellement. Ce réseau, entraîné sur un sous-ensemble de la base de données ImageNet comprenant environ 1,2 million d'images pour l'entraînement et 150000 pour les tests, a été le premier CNN à exploiter pleinement le potentiel du deep learning sur des ensembles de données volumineux.

Comme illustré en Figure III.7, l'entrée du réseau est une image de taille $224 \times 224 \times 3$, et la sortie est une fonction softmax pour la classification [102]. L'architecture d'AlexNet se compose de cinq couches de convolution, trois couches de mise en commun maximale, et trois couches entièrement connectées, utilisant des noyaux de tailles variées (11×11 pour la première, 5×5 pour la deuxième, et 3×3 pour les suivantes). Le réseau a introduit des innovations clés telles que l'activation ReLU pour accélérer la convergence, le dropout pour éviter le surajustement, et l'augmentation des données pour améliorer la généralisation. AlexNet a également adopté la mise en commun avec chevauchement pour affiner la précision du modèle. Afin d'accélérer les calculs, il a été formé sur deux GPUs (GTX 580) en traitement parallèle, permettant de surmonter les limitations matérielles. Ces innovations ont établi de nouveaux standards pour les réseaux neuronaux convolutionnels [102].

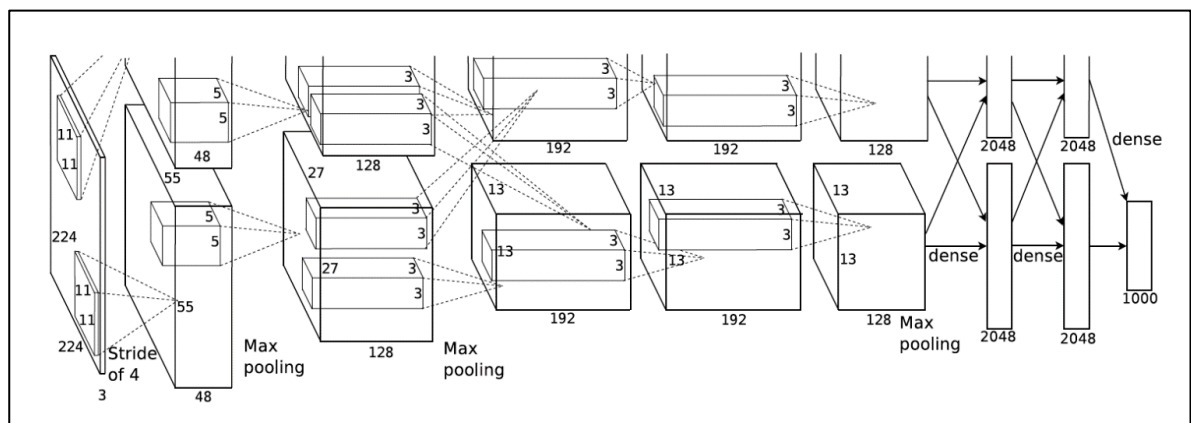


Figure III.7 : Architecture d'AlexNet telle que publiée dans l'article original [102].

III.12.3. VGGNet

En 2014, Simonyan et Zisserman de l'Université d'Oxford ont introduit le modèle VGG, qui a remporté le premier prix pour la localisation et le deuxième pour la classification lors de l'ILSVRC 2014. VGG comprend plusieurs architectures de réseaux convolutifs (figure III.8), dont les plus connus sont VGG16 et VGG19, appliqués à la reconnaissance faciale et à la classification d'images. L'objectif principal de leur recherche était de comprendre l'impact de la profondeur des réseaux convolutifs sur la précision de la classification à grande échelle.

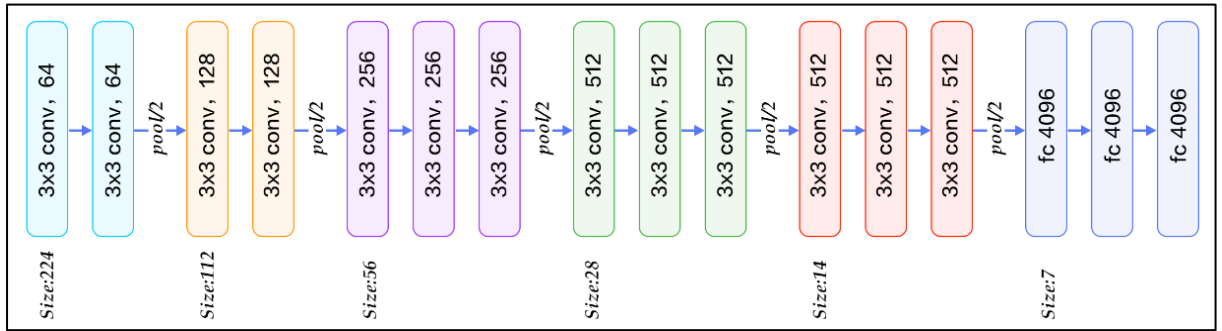


Figure III.8 : Architecture du VGG16.

VGG utilise de petits filtres de 3×3 dans ses couches convolutives, augmentant les performances tout en réduisant le nombre de paramètres. Par exemple Le réseau VGG16 traite des images RGB de 224×224 pixels et se compose de 13 couches convolutives suivies de trois couches entièrement connectées (FC). Les deux premières couches FC ont 4096 neurones chacune, et la dernière utilise une couche softmax pour classifier les images en 1000 catégories [103].

III.12.4. Google Net / InceptionV1 à V3

En 2014, Szegedy et al ont introduit GoogLeNet, un modèle révolutionnaire qui a remporté le premier prix lors de l'ILSVRC 2014 grâce à son architecture innovante. Ce réseau repose sur des modules Inception, combinant plusieurs types de convolutions (1×1 , 3×3 , 5×5) et du pooling, ce qui permet d'extraire simultanément des caractéristiques à différentes échelles. L'utilisation des convolutions 1×1 constitue l'une des avancées majeures de ce modèle, réduisant la dimensionnalité des données et optimisant ainsi le nombre de paramètres, tout en accélérant les calculs. Avec ses 22 couches, GoogLeNet surpasse la profondeur de modèles comme VGG, tout en restant très efficace avec seulement 6 millions de paramètres. Il intègre également des stratégies d'apprentissage telles que le dropout, pour éviter le surapprentissage, et la normalisation des activations afin de stabiliser le processus d'entraînement. Cette architecture se distingue par sa faible consommation en ressources tout en offrant des performances remarquables en classification d'images [104]. La figure III.9 représente la structure de GoogLeNet

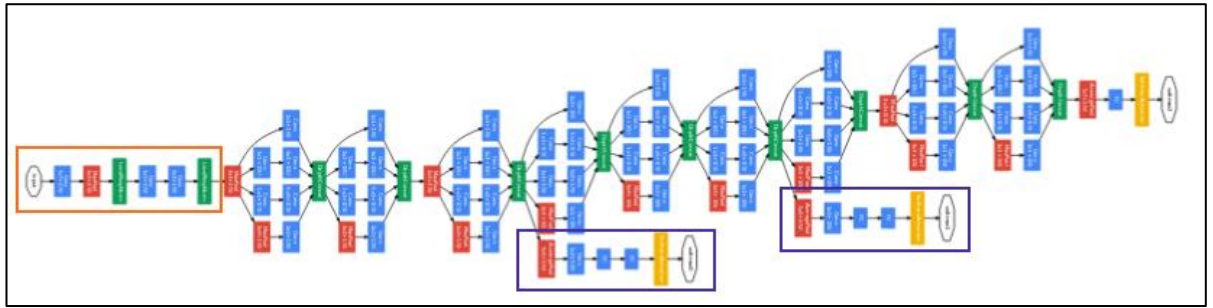


Figure III.9 : structures de GoogLeNet. [104]

III.12.5. Modules Inception

Le module Inception est un composant architectural puissant pour l'apprentissage en profondeur, en particulier pour la classification des images et l'analyse du comportement. Il utilise des filtres convolutifs parallèles de différentes tailles pour capturer des caractéristiques multi-échelles, et une couche de pooling. Les sorties de ces couches sont concaténées et transmises à la couche suivante. Les convolutions 1×1 agissent comme des couches de goulot d'étranglement pour réduire la dimensionnalité des cartes de caractéristiques [105]. Le module Inception, a évolué au cours de plusieurs itérations de la V1 à la V3, améliorant ainsi ses fonctionnalités dans diverses applications. Chaque version apporte des améliorations en termes d'architecture et de performances, ce qui en fait un outil puissant dans des domaines tels que la détection des défauts, le diagnostic médical et la gestion environnementale [105].

- Améliorations architecturales :

- Le module Inception V1 repose sur l'utilisation de filtres de différentes tailles (1x1, 3x3, 5x5) et d'une opération de mise en commun maximale (max pooling) appliqués en parallèle au même niveau. Cela permet au réseau d'être plus large que profond. Les sorties des différentes opérations sont ensuite concaténées et transmises à l'étape suivante. La figure III.10.a montre l'architecture initiale de Inception V1.

Pour optimiser les coûts de calcul, une convolution 1x1 est introduite avant les convolutions 3x3 et 5x5, réduisant ainsi le nombre de canaux d'entrée, ce qui allège la charge computationnelle. La convolution 1x1 est également appliquée après la mise en commun maximale, augmentant l'efficacité globale du module.

La figure III.10.b montre le module Inception V1 avec une réduction de dimension.

L'architecture de réseau neuronal connue sous le nom de GoogLeNet, ou Inception v1, a été construite en utilisant le module d'initiation à dimension réduite [106].

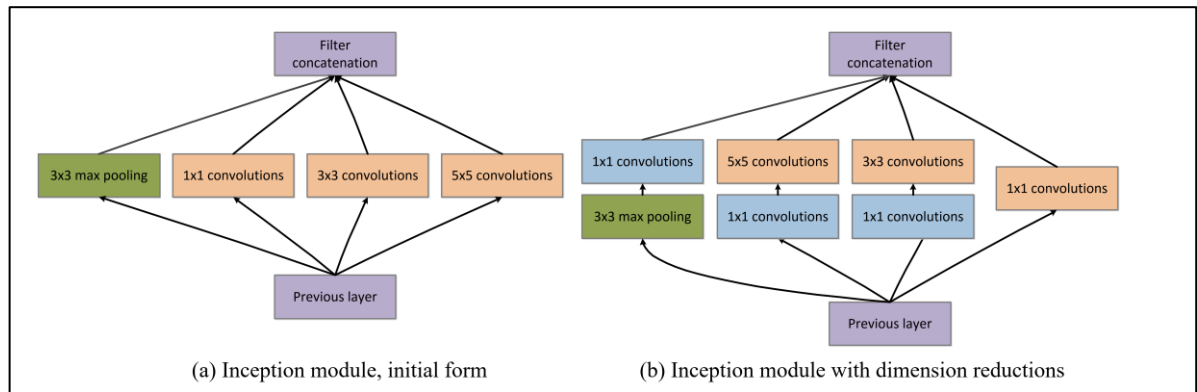


Figure III.10 : Modules Inception. [106]

- Pour réduire la complexité de calcul, Inception V2 (Figure II.11) factorise les convolutions 5x5 en deux convolutions 3x3, ce qui réduit le coût de calcul tout en maintenant la performance. De plus, les convolutions $n \times n$ sont remplacées par une combinaison de convolutions $1 \times n$ et $n \times 1$, réduisant ainsi la complexité de 33 % par rapport aux convolutions classiques [105].

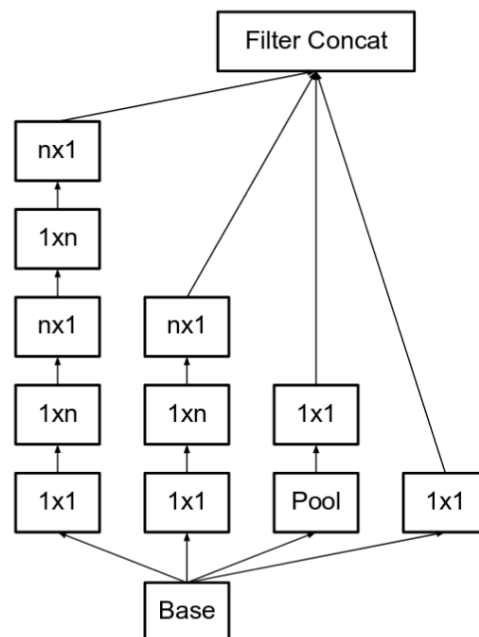


Figure III.11 : Schéma du module Inception V2 avec factorisation des convolutions. [105]

- Inception V3 a encore optimisé le modèle grâce à des techniques telles l'utilisation de l'optimiseur RMSProp, des convolutions factorisées 7x7, l'application de la BatchNorm dans les classificateurs auxiliaires, et le lissage des étiquettes pour éviter le surajustement en régularisant la fonction de perte [105].

III.12.6. ResNet

ResNet (2015), proposé par Kaiming He et al., a remporté l'ILSVRC 2015 grâce à l'introduction de connexions résiduelles, une innovation qui permet de surmonter la "dégradation" observée dans les réseaux profonds au-delà de 20 couches. Ces connexions directes contournent une ou plusieurs couches du réseau, ajoutant l'entrée directement à la sortie. Plutôt que d'apprendre une transformation complète des données, le réseau apprend la différence (ou "résidu") entre l'entrée et la sortie attendue, simplifiant ainsi l'optimisation des réseaux très profonds et améliorant leur performance (voir Figure 12) [107].

Une connexion résiduelle permet de définir la sortie d'un bloc comme :

$$Y = F(X) + X \quad (22)$$

Où : $F(X)$ est la transformation apprise par les couches du bloc, et X est l'entrée du bloc qui est ajoutée à la sortie finale.

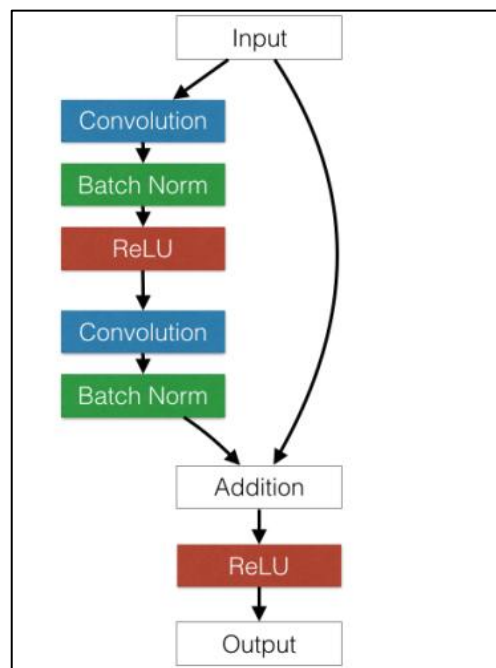


Figure III.12 : Bloc de base d'un ResNet. [108]

L'architecture de ResNet, illustrée dans la figure III.13, se distingue par ses blocs résiduels et ses connexions directes qui permettent de contourner plusieurs couches. Ces connexions facilitent l'optimisation de réseaux très profonds en évitant la dégradation des performances. Les différentes versions de ResNet, telles que ResNet-50, ResNet-101 et ResNet-152, se différencient principalement par le nombre de couches, chacune étant construite autour de ces blocs résiduels.

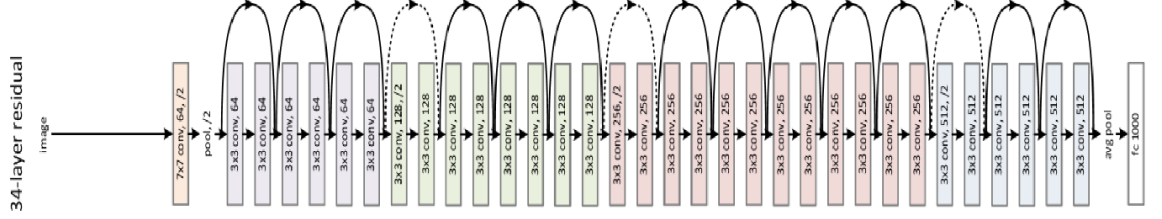


Figure III.13 : Architecture de ResNet [107].

III.12.7. DenseNet

DenseNet, proposé par Huang et al. [109], introduit, une architecture composée de blocs denses, où les cartes de caractéristiques de chaque couche sont connectées à celles de toutes les couches suivantes par concaténation. Cette approche, illustrée à la figure III.14, facilite la propagation des informations et des gradients à travers le réseau, simplifiant ainsi l'apprentissage et améliorant les performances en classification.

Equation:

$$X_l = H_l[X_0, X_1, X_2 \dots X_{l-1}] \quad (23)$$

Où $[X_0, X_1, X_2 \dots X_{l-1}]$ représente la concaténation des sorties des couches antérieures. Cette approche améliore le flux d'informations et des gradients entre les couches, facilitant ainsi l'apprentissage et renforçant les performances en classification. La transformation $H_l[.]$ est une fonction composite comprenant trois opérations : normalisation par lot (Batch Normalization), activation ReLU et convolution 3x3.

DenseNet organise le réseau en blocs denses, entre lesquels des couches de transition sont insérées pour gérer les variations de taille des cartes de caractéristiques. Ces couches effectuent des convolutions et du pooling, permettant de réduire la taille des cartes tout en maintenant la connectivité dense.

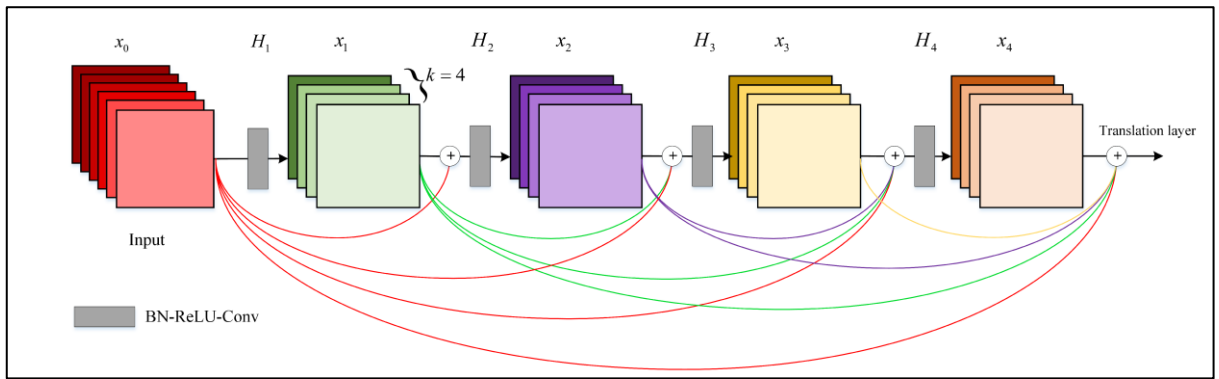


Figure III.14 : DensNet block [110]

III.13. L'apprentissage par transfert (Le transfer learning)

L'apprentissage par transfert est une technique de machine learning permettant d'adapter un modèle préalablement entraîné sur une tâche spécifique à une nouvelle tâche connexe. Cette approche est particulièrement efficace lorsque les deux tâches partagent des similarités et que les données disponibles pour la nouvelle tâche sont limitées. Elle permet de réutiliser les connaissances acquises par le modèle pour accélérer l'entraînement et améliorer les performances sur la nouvelle tâche. Ce concept s'applique notamment dans des domaines comme l'analyse d'images médicales, où il est souvent coûteux et difficile de collecter un nombre suffisant de données d'entraînement. Une approche courante qui consiste à pré-entraîner un réseau de neurones sur une grande base de données, telle qu'ImageNet (qui contient plus de quatorze millions d'images annotées et classées dans plus de vingt mille catégories), puis à ajuster ce modèle pour des applications spécifiques [111].

Le principe du transfer learning, est illustré dans la figure III.15.

III.13.1. Mise au point des techniques d'apprentissage par transfert

En pratique l'apprentissage par transfert se fait comme suit :

➤ Sélection du Domaine Source et Cible

Identifiez un domaine source où vous disposez de données abondantes et d'un modèle pré-entraîné. Par exemple, un modèle de classification d'images pré-entraîné sur ImageNet peut être utilisé pour des tâches spécifiques comme la classification d'images médicales.

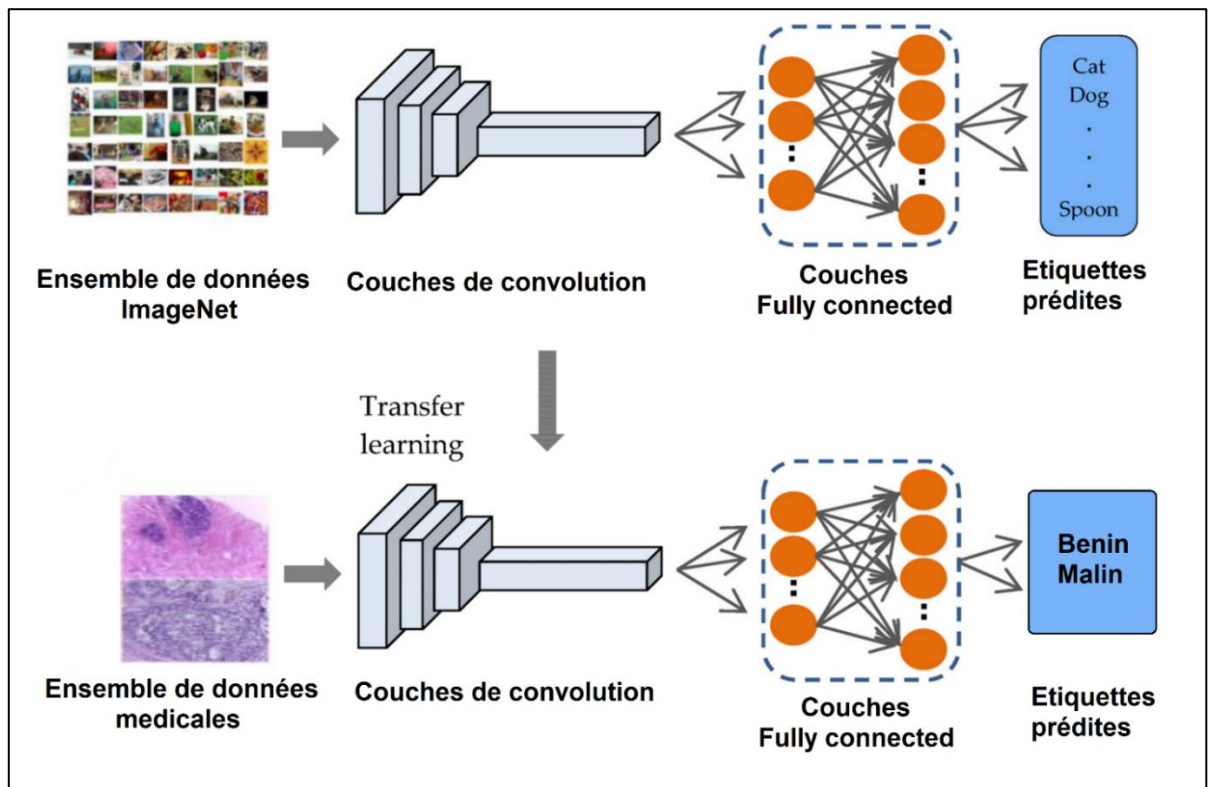


Figure III.15 : transfert d'apprentissage à partir d'ImageNet. [112]

➤ Prétraitement des Données

Les données du domaine cible doivent être prétraitées pour correspondre au format des données du domaine source. Cela peut inclure la normalisation, le redimensionnement des images, ou l'encodage des textes.

➤ Adaptation du Modèle

Le modèle pré-entraîné est souvent ajusté (fine-tuning) pour le domaine cible. Cela peut impliquer :

- Congélation des Couches : Dans cette approche, certaines couches du modèle pré-entraîné sont gelées et les couches restantes sont adaptées à la nouvelle tâche. Généralement, les couches inférieures du modèle pré-entraîné, qui capturent des caractéristiques de bas niveau telles que les bords et les coins, sont gelées, tandis que les couches supérieures, qui capturent des caractéristiques plus complexes, sont affinées.
- Ajout de Couches : Des couches supplémentaires peuvent être ajoutées pour adapter le modèle à la nouvelle tâche.

➤ Entraînement sur le Domaine Cible

Le modèle est ensuite entraîné sur les données du domaine cible.

Cela peut nécessiter moins de données étiquetées que si vous entraînez un modèle à partir de zéro, car le modèle a déjà appris des caractéristiques utiles à partir du domaine source.

➤ **Evaluation et Ajustement**

Après l'entraînement, le modèle est évalué sur un ensemble de validation ou de test. Des ajustements peuvent être nécessaires, comme le réglage des hyperparamètres ou l'ajout de techniques de régularisation pour éviter le surapprentissage.

III.14. Conclusion

Les réseaux de neurones convolutifs (CNN) se sont imposés comme une solution de pointe pour le diagnostic médical, notamment dans le domaine de la vision par ordinateur et de l'analyse d'images médicales, où ils ont atteint des performances à la pointe de la technologie. Grâce à leurs couches de convolution, de pooling et entièrement connectées, les CNN peuvent extraire des informations à différentes échelles et niveaux de granularité. Des modèles populaires tels que ResNet, VGG et DenseNet, ainsi que des techniques comme l'apprentissage par transfert et l'augmentation des données, ont permis d'améliorer considérablement la précision des modèles tout en réduisant le risque de surajustement.

Dans ce chapitre, nous avons présenté plusieurs modèles basés sur les CNN, largement utilisés pour la classification d'images. Ces réseaux sont aujourd'hui au cœur de nombreuses applications complexes, y compris la classification des images médicales, un problème clé abordé dans cette thèse. Le succès indéniable des CNN dans le traitement d'images a motivé leur adoption dans ce travail pour résoudre des problèmes spécifiques en imagerie médicale, en utilisant ce type de réseau comme base des approches proposées. Par conséquent, les CNN continuent d'évoluer, offrant des solutions toujours plus performantes et adaptées aux besoins croissants en analyse d'images.

Chapitre IV

Data Augmentation pour les Images Histopathologiques

IV.1. Introduction

Le cancer du sein est l'un des cancers les plus répandus dans le monde et demeure l'une des principales causes de mortalité chez les femmes. Selon l'Organisation mondiale de la santé (OMS), en 2020, on a recensé **2,3 millions de nouveaux cas** et **685 000 décès** liés à cette maladie, touchant principalement les femmes après la puberté (Figure IV.1). Cette pathologie se caractérise par la multiplication incontrôlée de cellules formant des tumeurs invasives, qui peuvent se propager et endommager d'autres tissus [113].

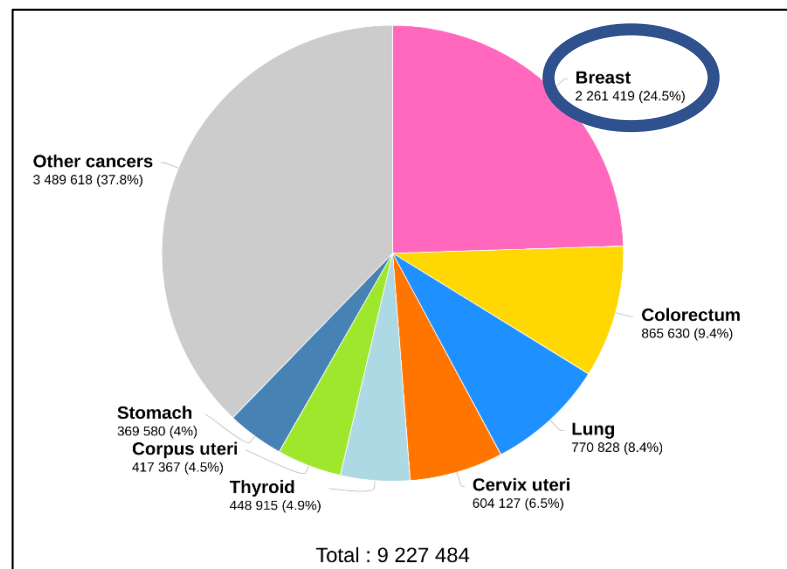


Figure IV.1. : Estimation mondiale des cas en 2020 chez les femmes.

Historiquement, les taux de mortalité du cancer du sein sont restés élevés jusqu'à l'introduction des programmes de détection précoce et de traitements adaptés dans les années 1980 [114]. Ces avancées ont considérablement amélioré les chances de survie des patients et la qualité des soins. La détection précoce du cancer du sein est importante, car elle permet d'initier des traitements adaptés dans des délais réduits, ce qui augmente significativement les chances de guérison [114]. L'analyse d'images histopathologiques est essentielle pour classer les tumeurs mammaires en bénignes ou malignes. Cependant, cette tâche complexe exige une expertise pointue pour détecter des variations

histologiques subtiles. L'interprétation manuelle des biopsies est longue et laborieuse, pouvant retarder le diagnostic jusqu'à deux semaines, avec des risques accrus d'erreurs et de retard de traitement [115]. Dans ce contexte, Les systèmes de diagnostic assistés par ordinateur (CAD), basés sur des algorithmes avancés, offrent une interprétation rapide et précise des images médicales, améliorant significativement les diagnostics en oncologie mammaire [116],[117].

L'apprentissage profond joue un rôle essentiel dans l'automatisation du diagnostic médical, car il renforce l'efficacité et la précision des processus cliniques. Des recherches récentes ont démontré l'efficacité des algorithmes basés sur les réseaux de neurones convolutifs (CNN) dans l'extraction automatique de caractéristiques et la classification des types de tumeurs. Ces modèles nécessitent toutefois une grande quantité de données et un temps de calcul conséquent pour atteindre des performances optimales [118],[119],[120]. Comme illustré dans la figure IV.2, les algorithmes d'apprentissage profond améliorent considérablement leurs performances à mesure que la quantité de données augmente, contrairement aux algorithmes de machine Learning traditionnels, dont les gains de performance sont plus limités [121]. Ainsi, pour maximiser l'efficacité de ces systèmes, il est indispensable de disposer de données variées permettant d'apprendre des caractéristiques complexes et de repérer des motifs subtils dans les images médicales. Cette diversité est essentielle pour former des modèles robustes capables de détecter des anomalies et de fournir des diagnostics fiables [122].

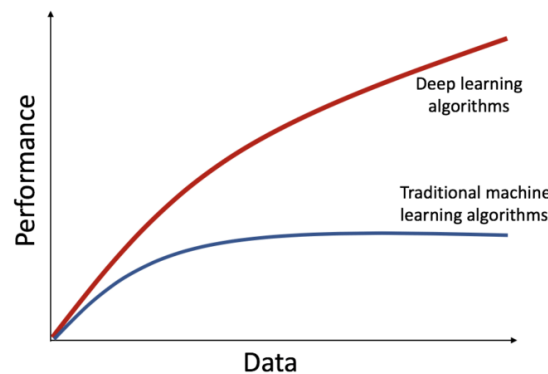


Figure IV.2. Importance de big data dans l'apprentissage profond.

Malheureusement, de nombreux jeux de données en imagerie médicale, comme BreakHis, présentent un déséquilibre des classes et un nombre limité d'images, ce qui complique le développement de modèles robustes [115]. La base de données BreakHis est largement utilisée dans la recherche sur le cancer du sein pour la classification des tumeurs bénignes et malignes à travers des images histopathologiques. Elle se distingue par son organisation en plusieurs niveaux de grossissement (40x, 100x, 200x et 400x), offrant ainsi une représentation diversifiée des échantillons tissulaires. BreakHis constitue une ressource précieuse pour développer et tester des modèles de classification en imagerie médicale, mais son déséquilibre des classes pose des défis que cette recherche s'efforce de surmonter [123].

IV.2 Origine et création de la base de données BreakHis

La base de données BreakHis (Breast Cancer Histopathological Database) a été créée dans le cadre d'une collaboration entre l'Université Fédérale de Paraná (UFPR) au Brésil et des chercheurs spécialisés en imagerie médicale et oncologie, en partenariat avec le Laboratoire P&D d'anatomie pathologique et cytopathologie. Elle a été mise à disposition en 2016 pour fournir à la communauté scientifique un ensemble de données complet et standardisé, destiné au développement de modèles d'intelligence artificielle (IA) pour le diagnostic du cancer du sein. L'objectif principal de cette initiative était de combler les lacunes des jeux de données existants, souvent limités en diversité et en standardisation, ce qui compliquait la comparaison des résultats entre différentes approches de classification. Par ailleurs, l'ensemble de données BreakHis, disponible en accès libre, a été spécialement conçu pour surmonter les défis du diagnostic manuel, en facilitant l'automatisation de l'interprétation des images histopathologiques [124],[125].

IV.2.1. Contenu et structure de la base de données BreakHis

Le diagnostic du cancer du sein repose en grande partie sur l'analyse histopathologique des tissus, qui implique une biopsie suivie d'une observation au microscope (Figure IV.3) [126].

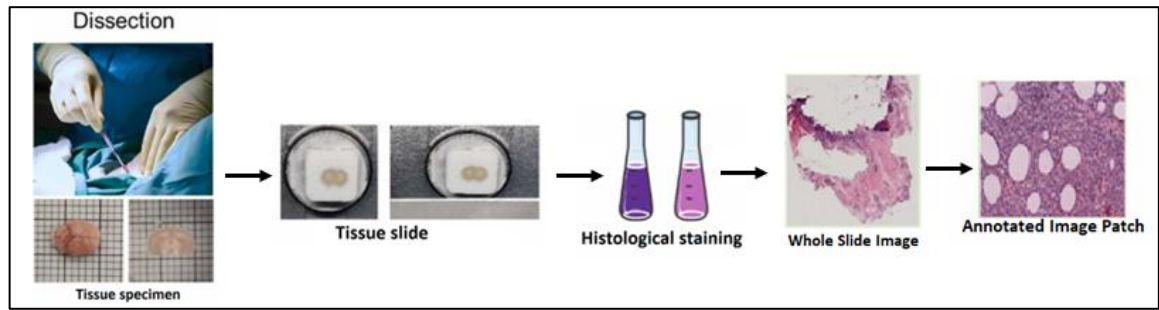


Figure IV.3. Le processus complet d'une biopsie

Ce processus permet d'évaluer la nature bénigne ou maligne des tumeurs en examinant la structure cellulaire et tissulaire. Les étapes principales sont les suivantes :

1. Prélèvement de l'échantillon (Biopsie)

Un échantillon de tissu est prélevé dans la zone suspecte à l'aide d'une aiguille (biopsie à l'aiguille fine) ou d'un punch, ou lors d'une intervention chirurgicale.

2. Fixation du tissu

Le tissu est immergé dans une solution de fixation (comme le formol) pour préserver sa structure cellulaire et éviter sa dégradation pour l'analyse.

3. Inclusion en paraffine

Le tissu est déshydraté et enrobé de paraffine pour former un bloc rigide. Ce bloc permet de réaliser des coupes fines (environ 3 à 5 micromètres) à l'aide d'un microtome.

4. Coupe et montage sur lame

Les coupes de tissu sont posées sur des lames de verre en vue de leur observation au microscope. Cela facilite l'examen des structures internes des cellules.

5. Coloration

Les lames sont colorées avec des agents spécifiques, comme l'hématoxyline-éosine (H&E), pour mettre en évidence les structures cellulaires et faciliter l'analyse. Des colorations supplémentaires ou des techniques d'immunohistochimie peuvent être employées pour détecter des marqueurs associés à des tumeurs spécifiques.

6. Observation au microscope

Le pathologiste examine les coupes sous différents grossissements (40x, 100x, 200x et 400x) pour évaluer les caractéristiques morphologiques des cellules et tissus (Figure IV.4) [127].

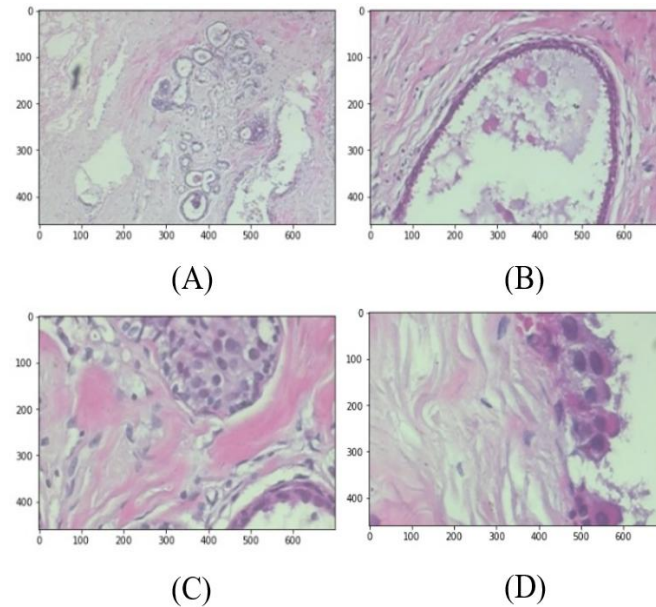


Figure IV.4. Diapositive d'une tumeur bénigne du sein vue sous différents facteurs d'agrandissement : (A) 40x, (B) 100x, (C) 200x, et (D) 400x [123].

- Les images issues de cette analyse sont capturées à travers trois canaux (RGB), permettant une visualisation détaillée des tissus sous divers niveaux de détail.
- L'observation porte sur des aspects tels que la taille, la forme et l'organisation des cellules, permettant de détecter les anomalies caractéristiques des tumeurs bénignes ou malignes.

7. Interprétation et diagnostic

En fonction de l'analyse microscopique, le pathologiste identifie si le tissu présente des lésions bénignes ou malignes. Le stade et le type de tumeur (par exemple, carcinome canalaire ou lobulaire) peuvent également être déterminés.

8. Rapport histopathologique

Le pathologiste rédige un rapport diagnostique, qui détaille les observations et fournit un diagnostic final. Ce rapport est transmis au médecin pour définir le traitement approprié.

La base de données BreakHis comprend un total de 7909 images microscopiques de biopsies mammaires, chacune d'une résolution de 700x460 pixels. Ces images proviennent de 82 patients et se répartissent en 2480 images de tumeurs bénignes et 5429 images de tumeurs malignes [128].

IV.2.2. Organisation et catégorisation des images

Les images de BreakHis sont classées selon deux grandes catégories : tumeurs bénignes et tumeurs malignes. Chaque catégorie comprend plusieurs sous-catégories histopathologiques :

Tumeurs bénignes	Tumeurs malignes
<ul style="list-style-type: none"> - Adenosis (A) - Fibroadenoma (F) - Phyllodes Tumor (PT) - Tubular Adenoma (TA) 	<ul style="list-style-type: none"> - Ductal Carcinoma (DC) - Lobular Carcinoma (LC) - Mucinous Carcinoma (MC) - Papillary Carcinoma (PC)

La répartition des images entre les catégories bénignes et malignes, en fonction du niveau de grossissement, est présentée par la figure IV.5 et le tableau VI.1. Ces données permettent aux chercheurs d'explorer l'influence du déséquilibre des classes sur les performances des modèles et de développer des stratégies pour y remédier.

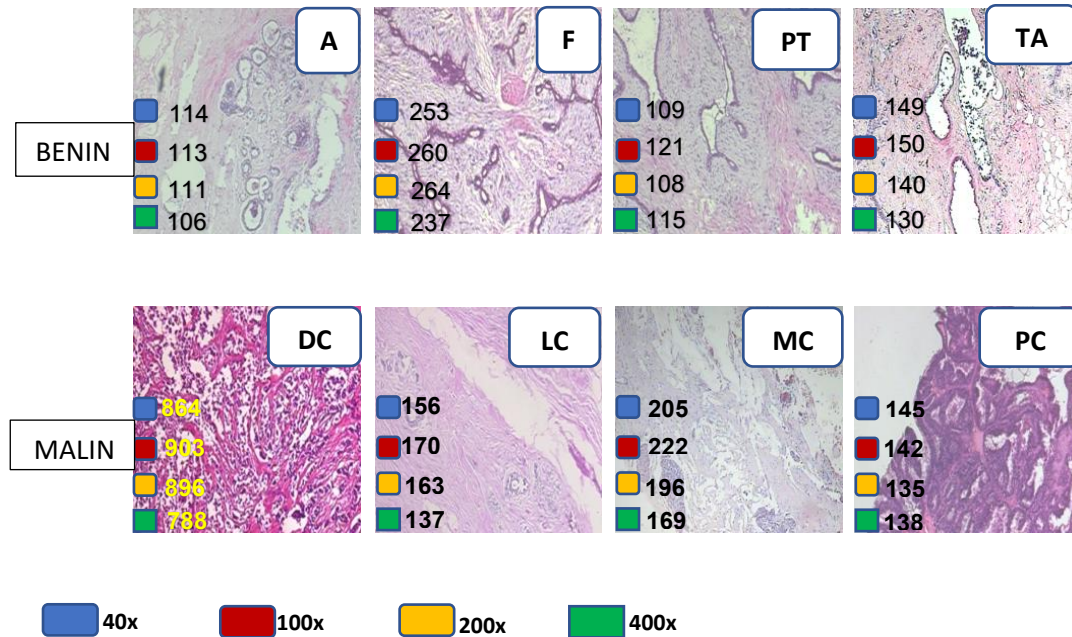


Figure IV.5. Distribution des images bénignes et malignes selon les grossissements.

Catégorie		Bénin					Malin					Total
Sous-catégorie		A	F	PT	TA	Total Bénin	DC	LC	MC	PC	Total Malin	
Nombre d'images Par facteur D'agrandissement	40x	114	253	109	149	625	864	156	205	145	1370	1995
	100x	113	260	121	150	644	903	170	222	142	1437	2081
	200x	111	264	108	140	623	896	163	196	135	1390	2013
	400x	106	237	115	130	588	788	137	169	138	1232	1820

Tableau VI.1. Catégorisation des images bénignes/malignes par grossissement [123].

IV.2.3. Illustration du déséquilibre des classes selon le grossissement

La Figure IV.6 présente la distribution des images entre les tumeurs bénignes et malignes à travers les quatre niveaux de grossissement disponibles dans la base de données BreakHis (40x, 100x, 200x, et 400x).

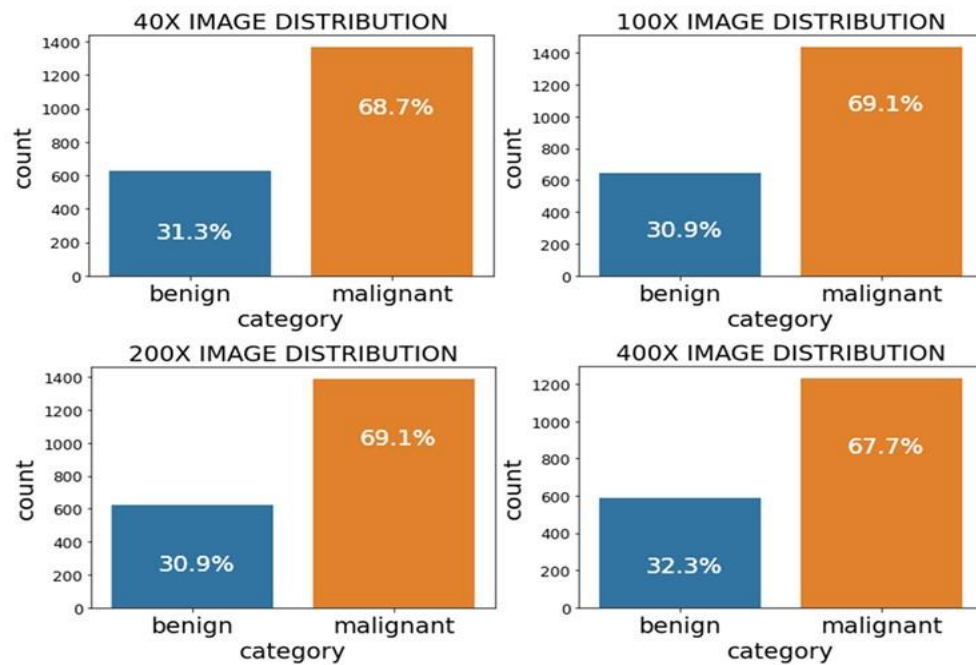


Figure IV.6. Distribution des classes BreakHis [113].

On observe un déséquilibre notable entre les deux catégories : les images de tumeurs malignes sont systématiquement plus nombreuses, représentant environ 67 % à 69 % des échantillons à chaque niveau de grossissement. À l'inverse, les tumeurs bénignes représentent seulement 30 % à 32 % des images [113].

Le manque de données et le déséquilibre des classes limitent la généralisation des modèles, favorisant les classes majoritaires comme les tumeurs malignes. Des techniques comme le rééchantillonnage, l'augmentation de données et des méthodes spécifiques améliorent la généralisation et traitent les classes sous-représentées. Un résumé de ces techniques est présenté ci-dessous.

IV.3. Techniques de gestion des données déséquilibrées

IV.3.1. Rééchantillonnage : Suréchantillonnage et Sous-échantillonnage

Le rééchantillonnage vise à équilibrer le nombre d'échantillons entre les classes. Le suréchantillonnage consiste à augmenter artificiellement le nombre d'exemples de la classe minoritaire, par duplication simple ou par la génération de données synthétiques avec des techniques telles que SMOTE et ADASYN. En parallèle, le sous-échantillonnage réduit le nombre d'exemples de la classe majoritaire, ce qui peut être utile lorsque la classe dominante contient beaucoup d'exemples redondants.

1. Suréchantillonnage (Oversampling)

- Duplication simple : Répéter les échantillons de la classe minoritaire.
- SMOTE (Synthetic Minority Over-sampling Technique) : Générer des échantillons synthétiques en interpolant entre des points proches de la classe minoritaire.
- ADASYN (Adaptive Synthetic Sampling) : Extension de SMOTE, cette technique se concentre sur les échantillons difficiles à classifier en générant plus de points synthétiques autour d'eux.

2. Sous-échantillonnage (Undersampling)

- Echantillonnage aléatoire : Réduire le nombre d'échantillons de la classe majoritaire en supprimant des données.
- Cluster-based Undersampling : Grouper les échantillons majoritaires en clusters et n'en garder qu'un sous-ensemble représentatif.
- NearMiss : Sélectionner les exemples majoritaires les plus proches des exemples minoritaires pour un meilleur équilibre.

IV.3.2. Techniques d'Augmentation de Données

L'augmentation des données est une technique qui consiste à générer de nouveaux ensembles de données à partir de données existantes, en créant des données synthétiques pour augmenter artificiellement la taille de l'ensemble de données. Cette méthode vise à améliorer les performances et la généralisation des modèles d'apprentissage profond, en réduisant les erreurs, et est particulièrement utile pour les tâches de classification. Les principales stratégies de cette technologie sont présentées ci-dessous.

II.3.2.a. Augmentation géométrique

Sont des opérations appliquées aux images qui modifient leur forme ou leur position, tout en préservant les informations visuelles essentielles. Ces transformations sont importantes pour l'augmentation des données et l'amélioration des performances des modèles d'apprentissage profond. Voici quelques opérations courantes :

- **Rotation** : Faire pivoter les images à différents angles. Voir Figure IV.7.

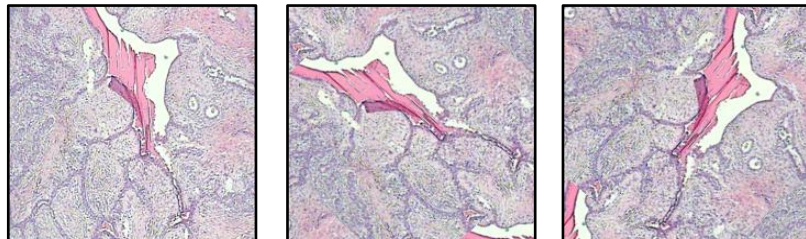


Figure IV.7. Opération de rotation.

- **Miroir** : Cette méthode consiste à faire pivoter l'image le long de l'axe horizontal ou vertical en partant du principe que la taille de la matrice est cohérente (Voir Figure IV.8).

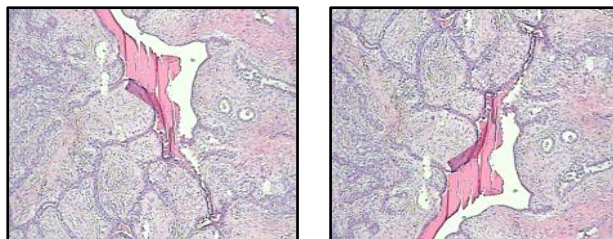


Figure IV.8. Opération de miroir.

- **Zoom** : Appliquer un zoom avant ou arrière (Voir Figure IV.9)

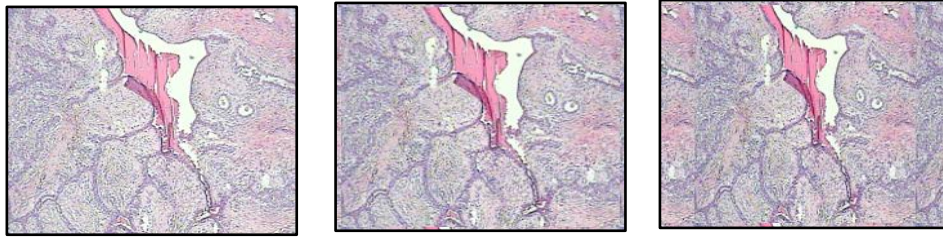


Figure IV.9. Opération de Zoom.

- **Shift** : déplace les pixels d'une image vers une direction spécifique (haut, bas, gauche ou droite). (Voir Figure IV.10).

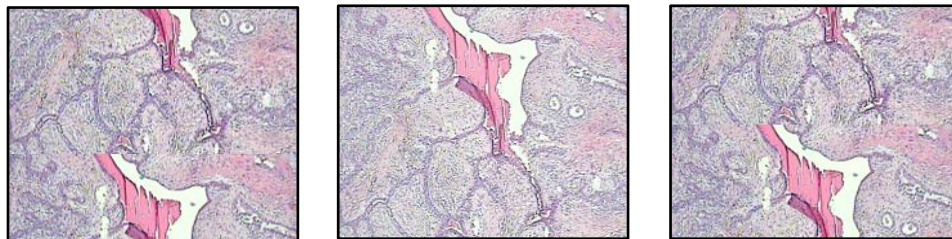


Figure IV.10. Opération de shift.

II.3.2.b. Modifications photométriques

Ajustement de la luminosité, du contraste (Figure IV.11a) ou ajout de bruit (Figure IV.11b).

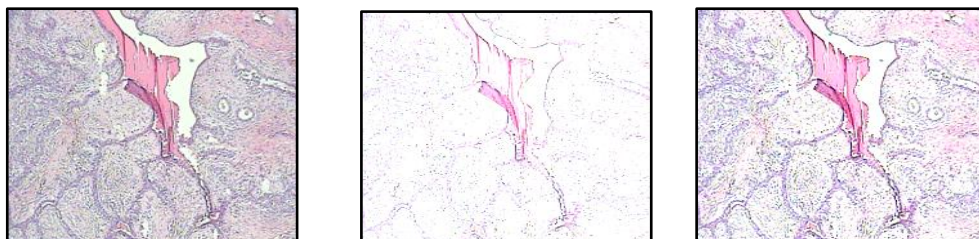


Figure IV.11a. Modifications de luminosité et de contraste.

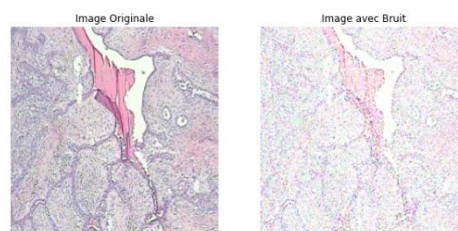


Figure IV.11b. Ajout de bruit

➤ **Mixup**

Est une technique d'augmentation de données qui consiste à créer de nouvelles images (Figure IV. 12) en combinant linéairement deux images existantes et en faisant de même pour leurs étiquettes (labels).

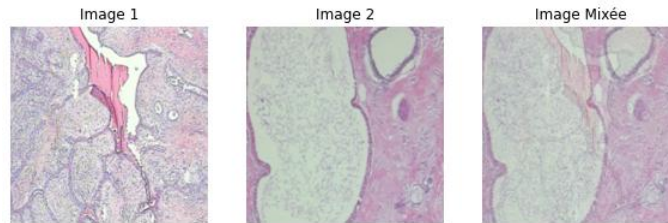


Figure IV.12. Technique de Mixup

➤ **CutMix**

Est une technique d'augmentation de données qui consiste à remplacer une région d'une image par une région d'une autre image (Figure IV. 13). Cette méthode améliore la robustesse des modèles de classification en combinant les informations de deux images différentes.

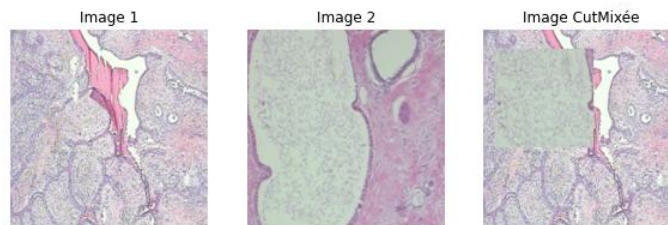


Figure IV.13. Technique de CutMix

IV.3.3. Méthodes Basées sur les Algorithmes

1. Ensembles d'apprentissage (Ensemble Learning)

Utilisation de plusieurs modèles pour améliorer la précision globale, par exemple:

➤ Bagging (Bootstrap Aggregating) : est une technique d'apprentissage automatique utilisée pour améliorer la précision des modèles en combinant les prédictions de plusieurs modèles de base (ex : Random Forest).

➤ Boosting : est une technique d'apprentissage automatique qui vise à améliorer la précision des modèles en combinant plusieurs modèles faibles pour créer un modèle puissant (ex : AdaBoost, XGBoost).

Ces techniques peuvent ajuster les modèles pour donner plus d'importance aux classes minoritaires.

2. Réseaux Adversaires (GANs)

L'utilisation GANs permet de générer des échantillons synthétiques dans des ensembles de données déséquilibrés.

3. Autoencodeurs

Apprennent à générer de nouvelles variations des données à partir de représentations internes, ce qui enrichit l'ensemble d'entraînement avec des exemples synthétiques.

4. Le transfer learning

L'apprentissage par transfert consiste à réutiliser les poids d'un modèle déjà entraîné sur un grand ensemble de données pour une nouvelle tâche. Cette méthode est utile lorsque les données spécifiques sont limitées, comme avec des modèles pré-entraînés tels que GoogLeNet ou ResNet formés sur ImageNet. Elle permet de diminuer le temps d'entraînement et les coûts tout en offrant de bonnes performances.

IV.4. Le traitement des données

Le traitement des données regroupe l'ensemble des étapes et techniques utilisées pour préparer, nettoyer, transformer et analyser des données brutes, afin de les rendre exploitables par les modèles d'apprentissage profond. Les données peuvent exister sous plusieurs formes : tables structurées ou non structurées, images, fichiers audio, vidéos, etc. Une machine ne peut pas interpréter directement du texte libre, des vidéos ou des images sous leur forme brute ; il est nécessaire de convertir ces informations en une forme numérique, principalement en 1 et 0. Par conséquent, les données brutes ne peuvent pas être directement fournies à un modèle d'apprentissage automatique en espérant qu'il soit capable de s'entraîner efficacement. Le prétraitement des données est une étape fondamentale de l'apprentissage profond.

Il consiste à transformer ou encoder les données afin de les rendre compréhensibles et utilisables par la machine. Un prétraitement soigné garantit que les données sont dans un format que la machine peut facilement traiter et analyser. C'est l'une des étapes les plus importantes pour assurer une bonne performance de généralisation du modèle, influençant directement sa capacité à

faire des prédictions précises sur des données nouvelles. Ci-après quelques techniques couramment utilisées pour le traitement des données dans le cadre de l'apprentissage profond.

IV.4.1. Nettoyage des données

- **Suppression du bruit** : Identification et suppression des données incohérentes ou erronées (par exemple, des échantillons mal étiquetés).
- **Normalisation des données** : Conversion des données dans un intervalle commun (par exemple, entre 0 et 1) pour faciliter l'entraînement des modèles.

IV.4.2. Encodage des données

- **Encodage one-hot** : Transformation des données catégorielles en vecteurs binaires.

IV.4.3. Réduction de la dimensionnalité

- **PCA (Analyse en Composantes Principales)** : Réduit le nombre de variables en identifiant les combinaisons les plus significatives.

IV.4.4. Traitement spécifique selon le type de données

- **Images** : Conversion des images en matrices de pixels (par exemple, RGB en vecteurs).
- **Audio** : Transformation en spectrogrammes pour une analyse plus facile.

IV.4.5. Batching et Shuffling

- **Batching** : Division des données en lots (batches) pour un entraînement plus efficace.
- **Shuffling** : Mélange aléatoire des données à chaque itération pour éviter le surapprentissage d'un ordre particulier.

IV.5. Conclusion

Dans ce chapitre nous avons exploré les différentes techniques décrites dans la littérature pour gérer les données déséquilibrées et améliorer les performances des modèles de deep learning appliqués aux données histopathologiques du

cancer du sein. Ces approches incluent des techniques d'augmentation de données (transformations géométriques, photométriques, et méthodes génératives) ainsi que des stratégies de gestion du déséquilibre des classes, telles que le suréchantillonnage et l'utilisation de GAN pour enrichir la classe minoritaire avec des données synthétiques réalistes. Le prochain chapitre sera consacré aux solutions adoptées pour résoudre le problème du déséquilibre des classes, un défi majeur en imagerie médicale.

Chapitre V

Etude Expérimentale

V.1. Introduction

L'objectif de ce chapitre est de classer les images du cancer du sein issues de l'ensemble de données BreakHis en deux catégories : bénin et malin, afin d'améliorer les diagnostics assistés par ordinateur (DAO) dans le domaine de la santé en utilisant l'apprentissage profond. Les progrès des architectures de réseaux de neurones convolutifs (CNN) ont rendu possible l'analyse automatisée des images histopathologiques numériques, malgré la complexité liée à la diversité morphologique des cellules tumorales et au déséquilibre des classes.

Ce déséquilibre, associé à un nombre limité d'échantillons, peut biaiser les modèles et réduire leur capacité à effectuer une classification précise.

Deux éléments clés influencent les performances des modèles : la qualité des données d'apprentissage et le choix du modèle d'extraction des caractéristiques. Ce dernier est particulièrement essentiel, car il conditionne la capacité du modèle à détecter les caractéristiques discriminantes nécessaires à une classification précise.

Relever ces défis est indispensable pour développer des systèmes robustes et fiables, capables de fournir des diagnostics précis dans des contextes cliniques.

Afin de résoudre cette problématique, nous avons proposé deux méthodes.

V.1.1. Méthode 1 : DCGAN et DenseNet201

Cette approche équilibre les classes en générant des images synthétiques avec DCGAN (Deep Convolutional Generative Adversarial), suivie d'une augmentation de données. Le modèle DenseNet201 pré-entraîné est utilisé pour extraire les caractéristiques.

V.1.2. Méthode 2 : WDT et DenseNet201

L'approche repose sur l'utilisation d'images générées à partir de la transformation par ondelettes discrètes pour augmenter la classe bénigne, corrigeant ainsi le déséquilibre des classes dans la base de données BreakHis.

En exploitant le modèle DenseNet201, elle améliore la classification des images histopathologiques du cancer du sein en intégrant les informations spatiales et spectrales, réduisant le risque d'erreurs critiques, comme la classification incorrecte de cas malins en bénins.

La structure générale du modèle proposé est décrite dans la figure V.1.

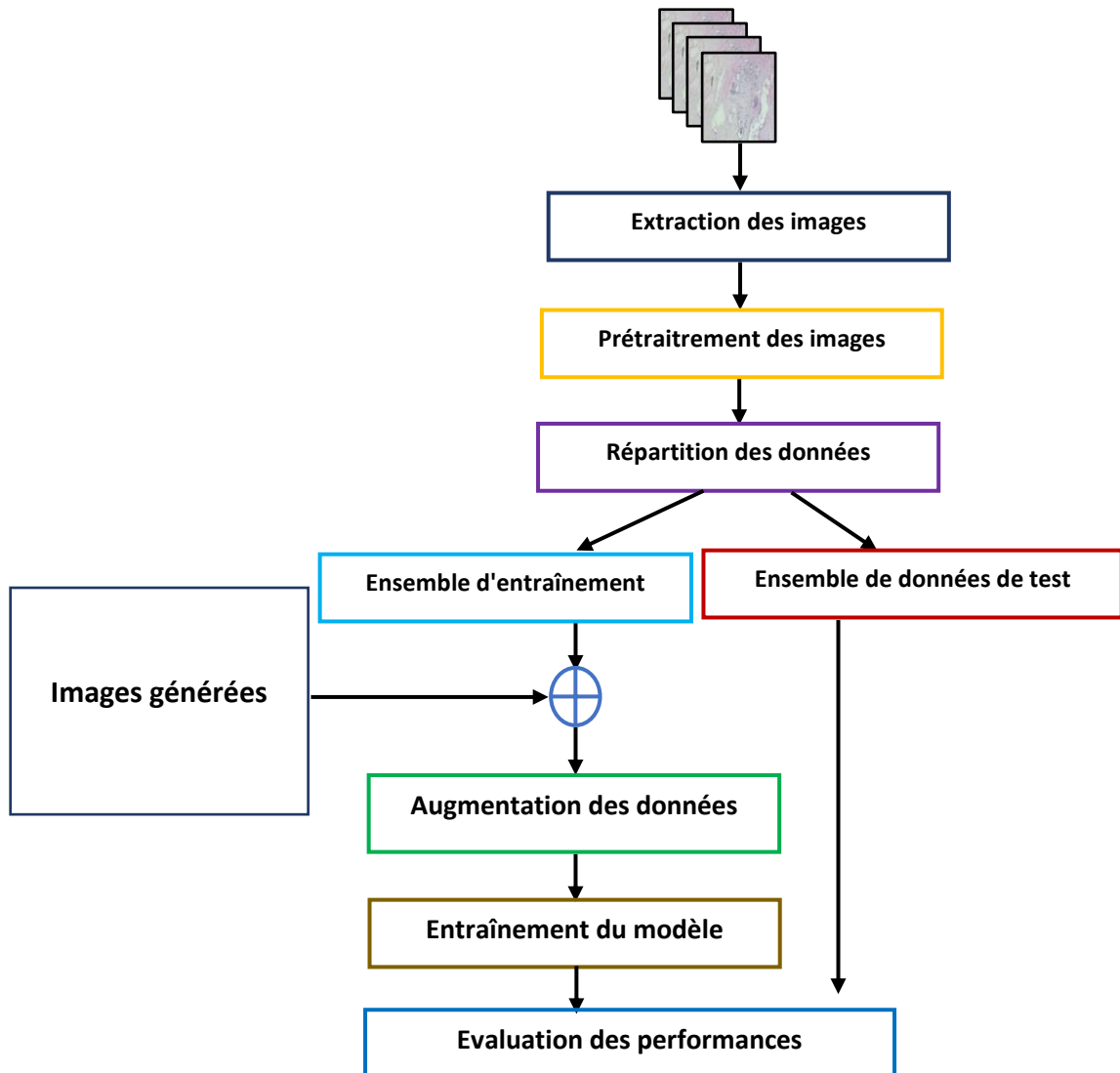


Figure V.1. Structure générale du modèle de classification des images histopathologiques du cancer du sein.

Les images générées par les deux méthodes proposées sont représentées par la figure V.2.

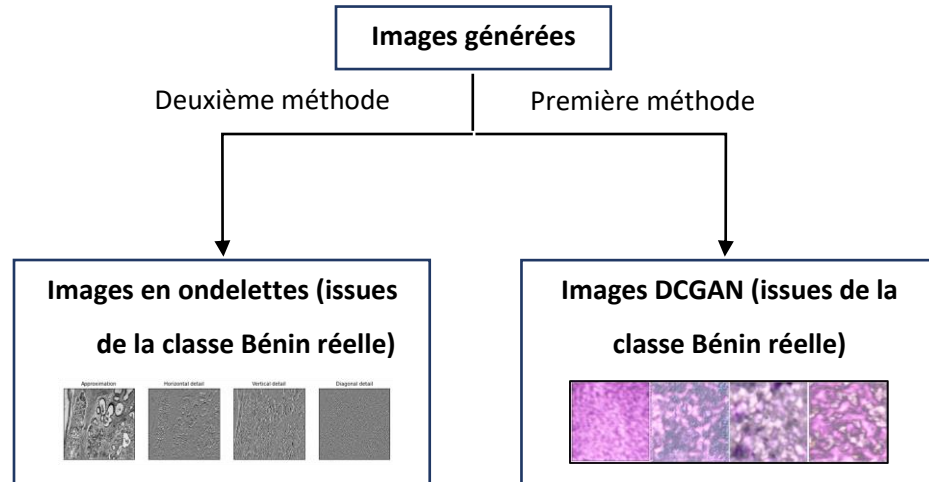


Figure V.2. Méthodes proposées pour la génération des images.

V.2. Première partie

V.2.1. Résolution du déséquilibre des classes avec DCGAN

Dans cette première partie, nous présentons l'utilisation du Deep Convolutional Generative Adversarial Network (DCGAN), pour résoudre le déséquilibre des classes dans la base de données BreakHis. Les DCGAN intègrent des réseaux de neurones convolutifs marquant ainsi une avancée significative dans la capacité des modèles génératifs à produire des images de haute qualité.

V.2.1.a. Deep Convolutional Generative Adversarial Networks (DCGAN)

Les Réseaux Antagonistes Génératifs Convolutionnels (DCGAN) sont une extension des GAN qui exploitent des couches de réseaux de neurones convolutionnels pour améliorer la qualité des images générées. Introduits en 2015, les DCGAN sont particulièrement efficaces pour générer des images synthétiques à haute résolution en exploitant les propriétés de la convolution [113]. Le schéma fonctionnel d'un DCGAN, illustré par la figure V.3, montre le processus de génération d'images synthétiques. Ce principe repose sur deux réseaux en compétition : un générateur (G) et un discriminateur (D).

Le générateur prend en entrée un vecteur aléatoire issu de l'espace latent et utilise des couches de convolution transposée (Conv2DTranspose) pour créer des images synthétiques à partir de cette entrée. En parallèle, le discriminateur, basé sur un réseau de convolution classique, essaie de différencier les images générées par le générateur de celles provenant de l'ensemble de données réelles.

Le processus d'apprentissage repose sur l'amélioration continue des deux réseaux : le générateur apprend à produire des images plus réalistes, tandis que le discriminateur affine sa capacité à distinguer les images générées des images réelles [129].

Les DCGAN sont connus pour leur capacité à générer des images visuellement cohérentes et détaillées, grâce à l'utilisation de mécanismes de Batch Normalization. L'adoption de fonctions d'activation comme LeakyReLU contribue à la convergence rapide du modèle.

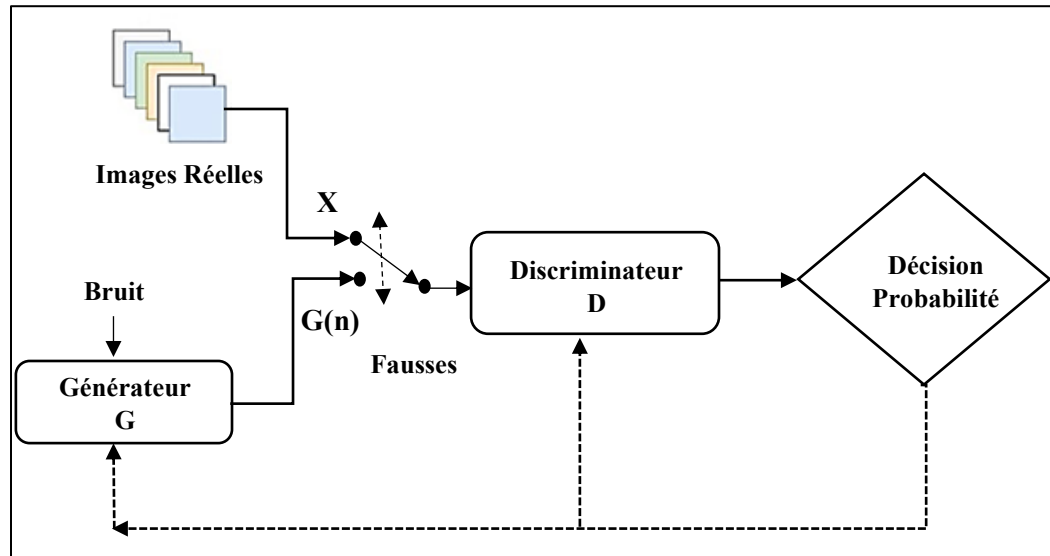


Figure V.3. Schéma fonctionnel d'un DCGAN.

V.2.1.b. Architecture du DCGAN proposé

L'architecture du DCGAN proposé est constituée de :

1. Le Générateur

La figure V.4 illustre l'architecture détaillée du générateur du DCGAN développée dans ce travail. Le réseau commence par une entrée latente (vecteur de bruit aléatoire) transformée progressivement en une image synthétique grâce à une série de couches dense, de convolutions transposées, de normalisation par lot et de fonctions d'activation. L'entrée est un vecteur latent aléatoire z de dimension 100.

•Etapes du réseau :

1. Dense Layer : Une couche dense de taille $224 \times 224 \times 10$.

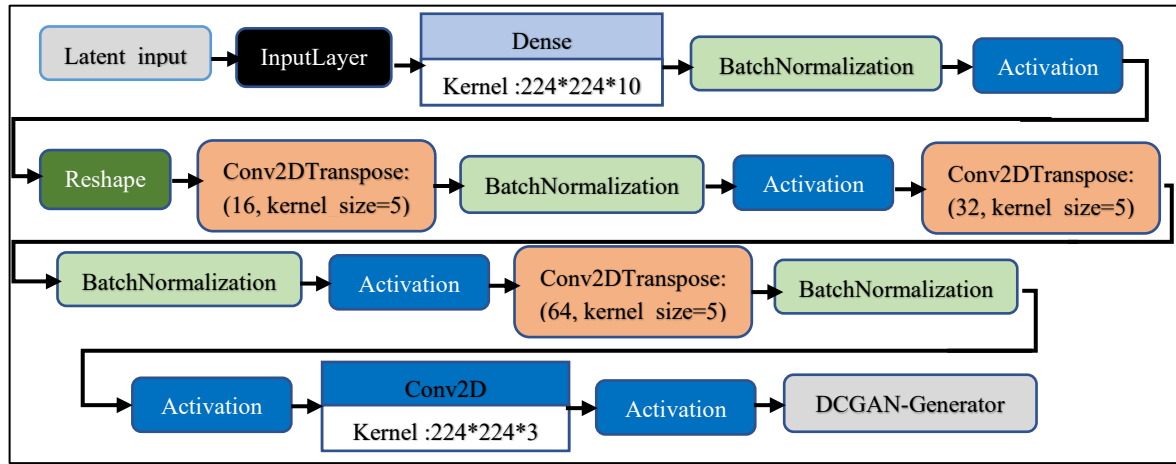


Figure V.4. Architecture détaillée du générateur du DCGAN proposé.

2. Batch Normalization : Utilisée pour stabiliser l'entraînement et normaliser les activations.
3. LeakyReLU : Fonction d'activation
4. Convolution Transposée : Plusieurs couches avec des filtres de tailles différentes :
 - Conv2DTranspose (5x5x16) : Augmente la taille sans perte d'information.
 - Conv2DTranspose (5x5x32) et (5x5x64) : Améliore la structure spatiale de l'image.
5. Dernière couche de sortie : Conv2D avec une activation sigmoïde pour produire une image normalisée avec trois canaux (RGB) dans la plage [0,1].

2. Le discriminateur

Le discriminateur (figure V.5), agit comme un classificateur binaire, prenant en entrée une image (réelle ou générée) et produisant une probabilité indiquant si l'image est réelle ou non. Il utilise plusieurs couches convolutives, suivies de normalisation et de fonctions d'activation.

1. Image_input (224,224,3) : Ce bloc représente les images d'entrée avec une taille de 224×224 pixels et 3 canaux (RVB). Cette taille est cohérente avec les dimensions couramment utilisées dans les architectures de réseaux de convolution modernes, telles que DenseNet, tout en offrant un bon compromis entre résolution et efficacité computationnelle.

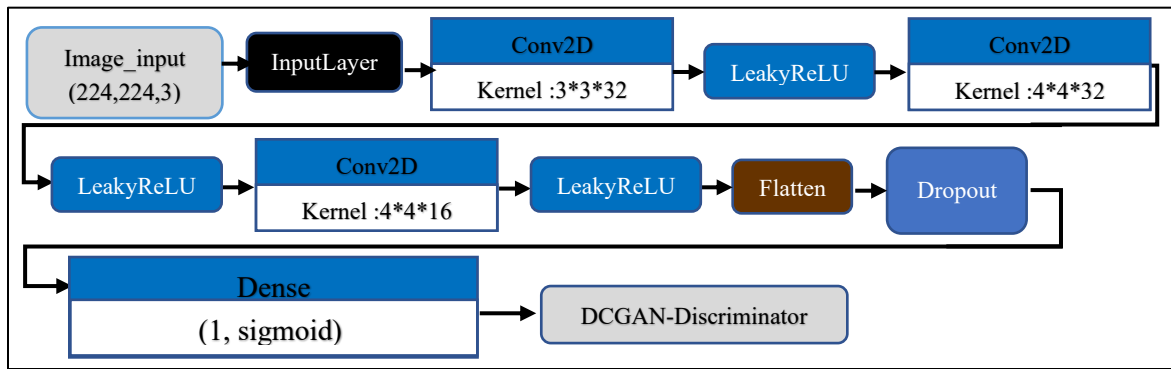


Figure V.5. Architecture détaillée du discriminateur du DCGAN proposé.

2. **InputLayer** : Ce bloc indique le début de l'architecture du réseau, préparant les données d'entrée pour les couches suivantes.
 3. **Première couche (Conv2D (Kernel : 3×3×32))** : Cette couche effectue une convolution 2D sur l'image d'entrée en utilisant 32 filtres de taille 3×3 pour extraire les caractéristiques spatiales locales.
 4. **Fonction d'activation (LeakyReLU)** : Appliqués après chaque couche de convolution pour introduire une non-linéarité. Ceci est pour permettre au réseau de modéliser des relations complexes.
 5. **Conv2D (Kernel : 4×4×32)** : Applique une autre convolution 2D avec 32 filtres de taille 4×4 pour extraire davantage de caractéristiques.
 6. **Conv2D (Kernel : 4×4×16)** : Réduit le nombre de filtres à 16 pour extraire des caractéristiques spécifiques et réduire progressivement la taille des données.
 7. **Flatten** : Transforme les données en une représentation vectorielle unidimensionnelle pour les connecter à des couches denses.
 8. **Dropout** : prévenir le surapprentissage en désactivant aléatoirement 20% des neurones de cette couche pendant l'entraînement.
 9. **Dense** : Connecte toutes les entrées à toutes les sorties via des poids (dense layer), générant un vecteur de taille correspondant au nombre de neurones spécifié.
 10. **La sortie** est une couche Dense finale avec une activation sigmoïde donnant une probabilité entre 0 et 1, indiquant si l'image est réelle ou générée.
- 3. L'entraînement du DCGAN** repose sur une compétition entre les deux réseaux: le générateur et le discriminateur. Le déroulement de l'apprentissage peut être décrit selon les étapes suivantes :

1. Génération d'images

- Passage par le générateur : Des vecteurs latents (bruit aléatoire) sont échantillonnés et transmis au modèle du générateur.
- Transformation : Le générateur, qui est un réseau de neurones, applique plusieurs transformations non linéaires au bruit d'entrée Z . Ces transformations permettent au générateur de capturer des caractéristiques complexes des données d'entraînement et de générer des images artificielles aux mêmes grossissements (40x, 100x, 200x, 400x) que les images réelles.
- Résultat : Le générateur produit un lot d'images générées qui tentent d'imiter les images réelles de la base de données.

2. Calcul des pertes

- Discriminateur

- Le discriminateur reçoit deux types d'images :
 - Des images réelles (classe bénigne) provenant de la base de données BreakHis avec plusieurs niveaux de grossissement (40x, 100x, 200x, 400x).
 - Des images synthétiques, générées par le générateur.
- Le discriminateur tente de classer chaque image en lui attribuant une étiquette:
 - 1 pour une image réelle.
 - 0 pour une image générée.
- La perte du discriminateur est calculée à l'aide d'une fonction de perte (binaire croisée) qui mesure l'écart entre :
 - Les prédictions du discriminateur.
 - Les étiquettes correctes (1 pour les images réelles, 0 pour les images générées).

- Générateur :

- Le générateur est évalué en utilisant les images qu'il a générées et qui ont été transmises au discriminateur.

Le but du générateur est de tromper le discriminateur pour que celui-ci classifie les images générées comme étant réelles.

- La perte du générateur est également calculée avec une fonction de perte, qui évalue dans quelle mesure le discriminateur a mal classé les images générées (le générateur cherche idéalement à obtenir une étiquette de 1 pour ces images).

3. Mise à jour des poids

- Une fois les pertes calculées, les poids des deux réseaux (générateur et discriminateur) sont mis à jour en utilisant un algorithme d'optimisation (Adam).
- Ce processus est répété sur un grand nombre d'itérations (ou époques). L'entraînement se fait en alternant la mise à jour du générateur et du discriminateur. A chaque itération, les deux réseaux s'améliorent progressivement :

- Le discriminateur devient meilleur à identifier les images générées.
- Le générateur produit des images de plus en plus réalistes pour tromper le discriminateur. La Figure V.6 présente le modèle DCGAN proposé.

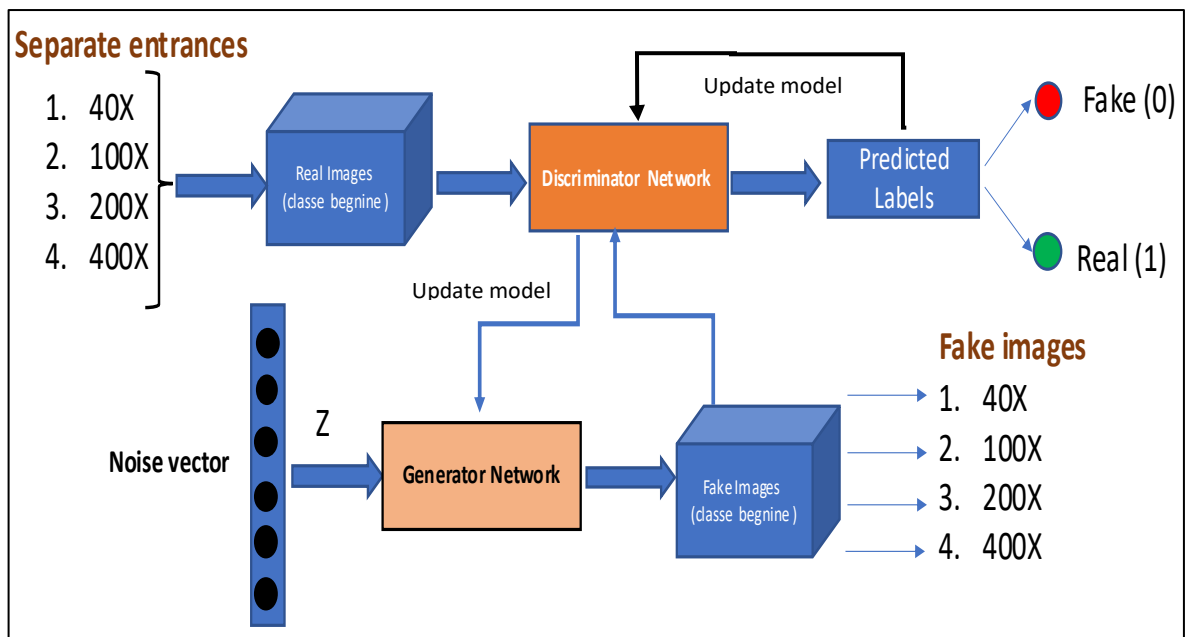


Figure V.6. Modèle DCGAN proposé [113]

La Figure V.7, illustre des échantillons d'images réelles et générées pour la classe minoritaire (bénigne).

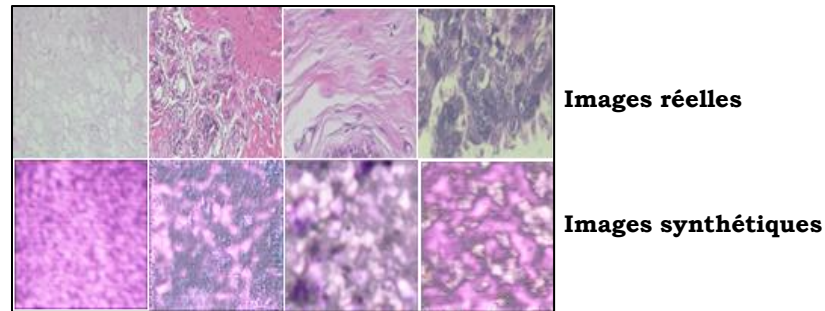


Figure V.7. Images réelles et synthétiques des tumeurs bénignes

Le tableau V.I présente le nombre d'images après l'utilisation de la technique d'augmentation DCGAN, permettant d'équilibrer les échantillons entre les classes « Bénin » et « Malin » à différents niveaux de grossissement.

Niveau de Grossissement	Type d'échantillon	Images Originales	Images réservées	Images restantes	Images générées	Total après Augmentation
40x	Bénin	625	76	549	674	1223
	Malin	1370	147	1223	/	1223
	Total	1995	223	1772	674	2446
100x	Bénin	644	79	565	722	1287
	Malin	1437	150	1287	/	1287
	Total	2081	229	1852	722	2574
200x	Bénin	623	77	546	697	1243
	Malin	1390	147	1243	/	1243
	Total	2013	224	1789	697	2486
400x	Bénin	588	72	516	493	1009
	Malin	1232	223	1009	/	1009
	Total	1820	295	1525	493	2018

Table V.1 Distribution des images avant et après DCGAN

V.2.2. Augmentation des données

Dans cette étude, nous avons appliqué plusieurs techniques d'augmentation de données pour améliorer l'entraînement du modèle, éviter le surapprentissage et optimiser ses performances. Ces techniques incluent le retournement, le recadrage, la rotation, le zoom et la mise à l'échelle. En générant différentes versions des images, ces transformations augmentent la diversité de la base de données BreakHis, renforçant ainsi la robustesse du modèle face à la variabilité

des données. Des exemples d'images augmentées sont présentés dans la figure V.8.

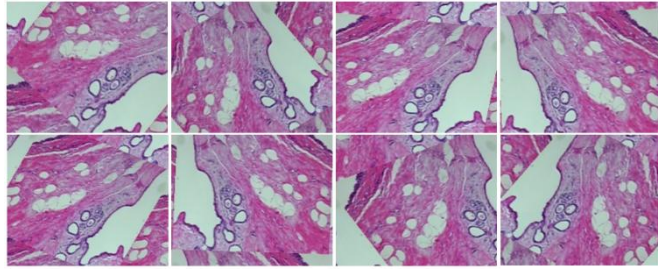


Figure V.8. Exemples d'images augmentées [123]

Le tableau V.2 récapitule les valeurs utilisées pour les transformations géométriques appliquées dans le cadre de l'augmentation de la base de données.

Data augmentation	Value
Rotation range	15
shear_range	0.2
zoom_range	0.2
horizontal_flip	True
fill_mode	'nearest'
width_shift_range	0.1
height_shift_range	0.1

Table V.2. Techniques d'augmentation des données.

V.2.3. Extraction de caractéristiques par apprentissage profond

Les réseaux de neurones convolutifs (CNN) sont utilisés pour extraire automatiquement des informations pertinentes des images brutes, pixel par pixel. Cependant, l'application des CNN nécessite une grande quantité de données, rendant difficile l'obtention des ensembles de formation et de test correspondants. Pour surmonter ce défi, l'apprentissage par transfert est introduit, permettant d'améliorer la précision de classification en utilisant les connaissances de modèles pré-entraînés sur de vastes ensembles de données. Cette approche est particulièrement avantageuse dans des scénarios où les données disponibles pour l'entraînement sont limitées. L'apprentissage par transfert implique souvent un entraînement sur ImageNet, un ensemble de données contenant 14 millions de photographies annotées, réparties sur environ 1000 classes distinctes.

Ce concept a gagné en popularité dans divers domaines de la vision par ordinateur, tels que la détection d'objets, la classification d'images et la segmentation.

V.2.3.a. La structure DenseNet

DenseNet présente une architecture de réseau neuronal qui atténue la perte d'information en permettant aux couches de recevoir des entrées de toutes les couches précédentes et de les transmettre aux couches suivantes. Cette architecture compacte privilégie la concaténation séquentielle des cartes de caractéristiques. DenseNet a démontré des performances de classification supérieures sur des ensembles de données de référence tels que CIFAR-100 et ImageNet. La figure V.9 illustre la précision top-1 en validation, une métrique utilisée en DL pour évaluer les performances de modèles de classification sur un ensemble de validation. La comparaison montre clairement que DenseNet surpasse les autres modèles pré-entraînés, ce qui justifie son choix en tant que modèle de base pour cette étude.

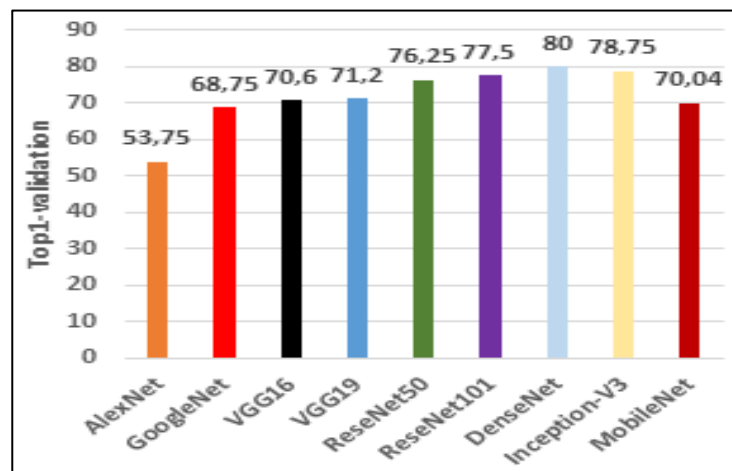


Figure V.9. Meilleures performances obtenues pour ImageNet [\[123\]](#)

V.2.4. Méthodologie

Un sous-ensemble de test a été réservé pour chaque classe avant la correction du déséquilibre. Afin de compenser l'inégalité entre les classes, l'approche DCGAN a été utilisée pour générer des images bénignes synthétiques et équilibrer leur nombre avec celui des images malignes. L'ensemble des images, initialement de taille (700×460×3), a été redimensionné en (224×224×3) afin de réduire la

complexité computationnelle et de respecter les contraintes d'entrée du modèle DenseNet-201. Après équilibrage, les données ont été réparties en deux ensembles : 96 % pour l'entraînement et 4 % pour la validation, et ce pour chaque niveau de grossissement (40x, 100x, 200x et 400x).

Les images, provenant de divers niveaux de grossissement (40x, 100x, 200x et 400x), ont été entraînées et testées séparément selon leur grossissement respectif. L'extraction des caractéristiques a été réalisée à l'aide du modèle pré-entraîné DenseNet201. Contrairement à l'utilisation traditionnelle du réseau, où l'on applique un Flatten, cette approche utilise le Global Average Pooling (GAP) pour réduire la dimensionnalité, évitant ainsi un nombre excessif de paramètres qui pourrait entraîner un surapprentissage et compliquer l'entraînement du modèle. Les caractéristiques ont été extraites des couches pool3_pool du bloc dense 2, pool4_pool du bloc dense 3, et conv5_block32_concat du bloc dense 4. Ces sorties ont été transformées en une représentation vectorielle via le GAP, aboutissant à un vecteur de 3072 dimensions. En outre, une couche de régularisation (DROPOUT) a été intégrée pour atténuer le surapprentissage, et une couche dense finale, avec une fonction SOFTMAX, a permis la classification binaire. La figure V.10 illustre l'architecture proposée.

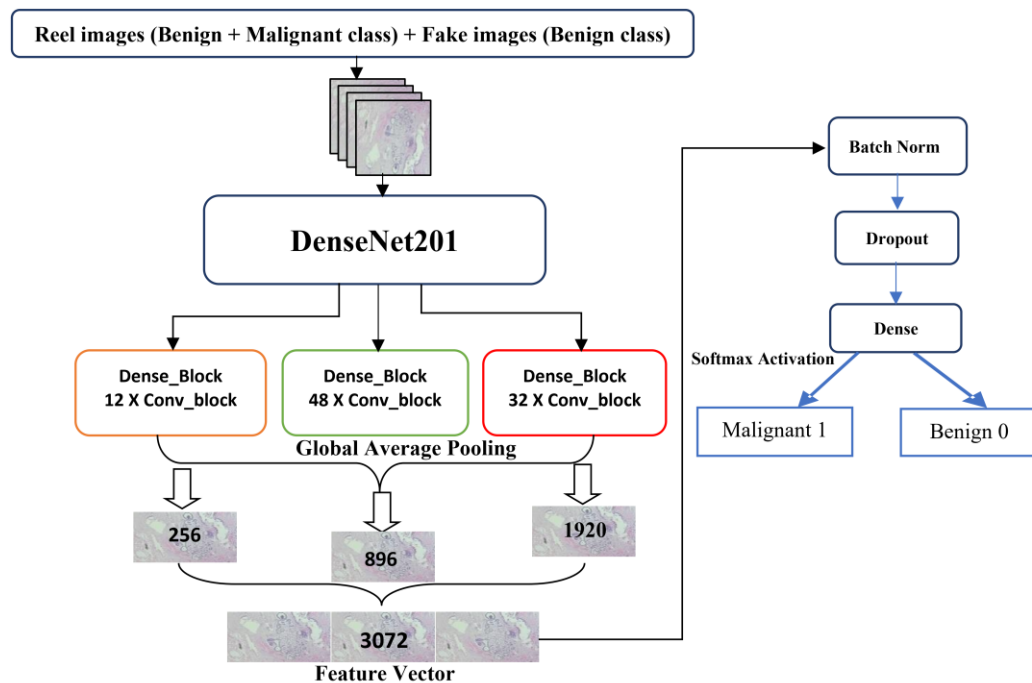


Figure V.10. Architecture proposée pour la classification des images histopathologiques [\[123\]](#)

V.2.5. Paramètres d'entraînement du modèle

- **Poids pré-entraînés** : Utilisation de poids initiaux pré-entraînés sur la base de données ImageNet.
- **Algorithme d'entraînement** : Entraînement avec l'algorithme RMSprop, connu pour sa capacité d'adaptation dynamique aux gradients.
- **Taux d'apprentissage** : Taux d'apprentissage initial fixé à $10e-4$, progressivement réduit vers zéro pour améliorer la convergence.
Le taux d'apprentissage est initialement configuré avec l'optimiseur RMSprop. Ensuite, le ReduceLROnPlateau ajuste le taux d'apprentissage pendant l'entraînement, basé sur l'évolution de la précision de validation. Si la précision ne s'améliore pas, le taux d'apprentissage est réduit par un facteur de 0.2, ce qui peut aider à améliorer la convergence du modèle.
- **Epoques et taille de lot** : Entraînement réalisé sur 40 époques avec une taille de lot de 35, sans convergence significative constatée après ces 40 époques.
- **Fonction de perte** : Utilisation de la cross-entropie catégorique, adaptée pour quantifier la différence entre distributions de probabilité dans un contexte multi-classe.
- **Fonction d'activation** : Application de la fonction d'activation SOFTMAX pour transformer les scores bruts en probabilités, facilitant l'interprétation des résultats.

- Principe de Fonctionnement de SOFTMAX

La fonction SOFTMAX prend en entrée un vecteur de scores produits par le modèle, typiquement dans la dernière couche d'un réseau de neurones. Ces scores peuvent être n'importe quelle valeur réelle, positive ou négative.

La fonction commence par faire l'exponentielle de chaque score, ce qui permet de transformer tous les scores en valeurs positives. La formule pour chaque élément Z_i du vecteur d'entrée est : e^{Z_i} (24)

Ensuite, on calcule la somme des exposants de tous les scores :

$$S = \sum_j e^{Z_j} \quad (25)$$

Enfin, chaque score obtenu dans le calcul des exposants est divisé par cette somme. Cela donne une probabilité pour chaque classe, qui varie entre 0 et 1, et la somme de toutes les probabilités est égale à 1 :

$$\text{Softmax}(Z_i) = \frac{e^{Z_i}}{S} \quad (26)$$

V.2.6. Métriques de performance

Pour évaluer la performance du modèle proposé, plusieurs métriques basées sur les éléments de la matrice de confusion ont été utilisés, à savoir :

- **Vrai positif (TP)** : Le modèle identifie correctement un cas positif de cancer.
- **Vrai négatif (TN)** : Le modèle identifie correctement un cas négatif de cancer.
- **Faux positif (FP)** : Le modèle classe incorrectement un cas négatif comme positif.
- **Faux négatif (FN)** : Le modèle classe incorrectement un cas positif comme négatif.

➤ Définition des métriques :

- **Précision globale (Accuracy)** : Mesure le pourcentage de bonnes prédictions parmi le total des prédictions effectuées. Elle peut cependant être trompeuse en cas de déséquilibre des classes.

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

- **Rappel (Recall)** : Evalue la capacité du modèle à détecter correctement les cas positifs.

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

- **Précision (Precision)** : Mesure la proportion des prédictions positives correctes parmi l'ensemble des prédictions positives.

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

- **F1-Score** : Combine la précision et le rappel pour fournir une évaluation globale de la performance du modèle.

$$F1 \text{ score} = \frac{2 \times (\text{precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$$

En recherche médicale, la réduction des faux positifs et des faux négatifs est primordiale.

Classer à tort un cancer malin comme bénin peut entraîner une fausse assurance, poussant la patiente à ignorer des symptômes inquiétants. Inversement, diagnostiquer un cancer bénin comme malin conduit à des traitements invasifs et inutiles (chirurgies ou chimiothérapies), provoquant un stress évitable.

La Table V.3 présente les mesures de performance calculées à partir de la matrice de confusion binaire.

	Confusion Matrix			Facteurs d'agrandissement	Algorithmes DenseNet	Métriques de performance						
Predit→ Réal ↓	Benin	Malin	Support			Precision		Recall		F1-Score		Accuracy
						B	M	B	M	B	M	
Benin	66	10	76	40x	Block2	0.98	0.92	0.84	0.99	0.91	0.96	0.94
					Block3	1.00	0.84	1.00	0.91	0.96	0.95	0.95
Malin	3	144	Block4		0.96	0.94	0.87	0.98	0.91	0.96	0.94	
			Concaténation		0.96	0.95	0.91	0.98	0.93	0.97	0.96	
Benin	76	3	79	100x	Block2	0.86	0.99	0.97	0.91	0.91	0.95	0.93
					Block3	0.90	0.95	0.91	0.95	0.91	0.95	0.93
Malin	9	141	Block4		0.88	0.98	0.96	0.93	0.92	0.96	0.94	
			Concaténation		0.89	0.98	0.96	0.94	0.93	0.96	95	
Benin	76	1	77	200x	Block2	0.78	1.00	1.00	0.85	0.88	0.92	0.90
					Block3	0.80	0.98	0.96	0.87	0.87	0.92	0.90
Malin	25	122	Block4		0.76	0.96	0.94	0.84	0.84	0.90	0.88	
			Concaténation		0.75	0.99	0.99	0.83	0.85	0.90	0.88	
Benin	70	2	72	400x	Block2	0.74	0.99	0.97	0.89	0.84	0.94	0.91
					Block3	0.71	0.99	0.97	0.87	0.82	0.93	0.90
Malin	22	201	Block4		0.77	0.98	0.94	0.91	0.85	0.94	0.92	
			Concaténation		0.76	0.99	0.97	0.90	0.85	0.94	0.92	

Table V.3. Résultats de la classification.

V.2.7. Discussion des résultats

- Comparaison des blocs individuels et de la concaténation

- La concaténation des blocs produit de meilleures performances, notamment pour l'Accuracy, le rappel (Recall) , et le F1-score.

- Cette approche semble compenser les limitations des blocs individuels en combinant leurs caractéristiques extraites, ce qui enrichit l'apprentissage du modèle.
- Les différences de performance entre les blocs individuels et la concaténation sont plus prononcées à des niveaux de grossissement plus faibles (40x et 100x), ce qui montre que l'intégration des informations globales et locales est particulièrement bénéfique pour ces niveaux.

- **Analyse des performances selon les métriques**

- Le rappel pour la classe maligne (M) est généralement élevé pour tous les niveaux de zoom, ce qui indique que le modèle est performant pour détecter les échantillons malins, une caractéristique souhaitable pour éviter les faux négatifs.
- En revanche, la précision pour la classe bénigne (B) est plus variable et tend à diminuer aux niveaux de zoom plus élevés (200x et 400x), suggérant une difficulté à identifier correctement les échantillons bénins.
- Le score F1 pour la classe bénigne (B) est souvent plus faible que pour la classe maligne (M), ce qui reflète cette asymétrie de performance et pourrait nécessiter des ajustements des méthodes de traitement des données supplémentaires.

Globalement, le modèle DenseNet201, utilisant la concaténation des trois blocs internes, offre les meilleures performances pour la classification des images histopathologiques de cancer du sein, en particulier aux niveaux de grossissement 40x et 100x. Les résultats montrent une robustesse dans la détection des échantillons malins, mais des difficultés à distinguer les échantillons bénins à des niveaux de zoom plus élevés. Cette analyse souligne l'importance de la fusion des caractéristiques pour améliorer la robustesse du modèle et la nécessité d'explorer des techniques supplémentaires pour optimiser la classification des images à des grossissements plus détaillés.

V.3. Deuxième partie

V.3.1. Résolution du déséquilibre des classes avec les ondelettes

Dans cette seconde partie, en décomposant les images bénignes à l'aide de la transformée en ondelettes discrète, nous avons pu équilibrer les classes et

enrichir la qualité des données. Cette méthode améliore l'extraction des caractéristiques essentielles tout en atténuant le bruit, contribuant ainsi à optimiser la classification des images histopathologiques du cancer du sein.

V.3.2. Transformée en ondelettes discrète (DWT)

La transformée en ondelettes discrète (DWT) est un outil d'analyse mathématique qui décrit une image en termes de caractéristiques spatiales et de fréquences. Elle utilise des filtres avec des fréquences de coupure différentes. L'image est soumise à un filtre passe-bas, connu sous le nom de « fonction en échelon », et à un filtre passe-haut, appelé « fonction d'ondelette ». Ces deux filtres sont appliqués successivement à l'ensemble de l'image. A la sortie, nous obtenons quatre bandes de fréquence : la première bande de basse fréquence (LL) représente une sorte de moyenne du signal original, appelée image approximative, ainsi qu'une version réduite et lissée de l'image originale. La deuxième partie comprend trois sous-bandes à haute fréquence, caractérisées par leur orientation spatiale : HL (horizontale), LH (verticale) et HH (diagonale). Les images détaillées (HH, HL et LH) sont généralement appelées coefficients d'ondelettes et mettent en évidence les contours des régions de l'image. Ce processus peut être répété autant de fois que nécessaire [130]. Alfred Haar a introduit les ondelettes en 1909 et les a appliquées à la représentation de signaux unidimensionnels. Stéphane G. Mallat a étendu l'application des transformées en ondelettes aux images. Il a donc introduit un algorithme rapide de décomposition/reconstruction des ondelettes. Cet algorithme est récursif et repose principalement sur deux opérations [130].

Filtrage : Convolution d'un signal avec un filtre passe-bas (h_0) ou un filtre passe-haut (g_0).

Sous-échantillonnage : Réduit le nombre d'échantillons du signal. En fait, le sous-échantillonnage horizontal (1:2) d'une image équivaut à supprimer une colonne sur deux, réduisant ainsi le nombre de pixels par ligne de moitié. La figure V.11 montre l'algorithme de MALLAT.

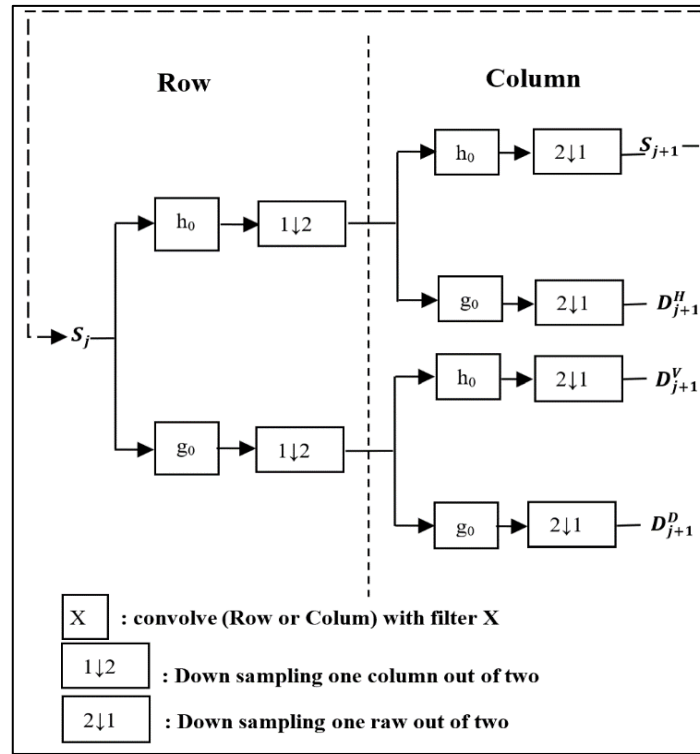


Figure V.11. Algorithme de MALLAT [123]

L'algorithme de MALLAT peut être expliqué comme suit :

Soit S_j représentant l'image approximative au niveau de résolution j , et D_j^X représentant la sous-bande à l'orientation x (où $x \in \{H, V, D\}$) qui est extraite au niveau de résolution j . Dans l'algorithme, l'image d'entrée S_j est d'abord soumise à un filtrage passe-haut et passe-bas. Les images résultantes sont ensuite sous-échantillonnées sur les lignes, et chacune des images sous-échantillonnées est à nouveau filtrée par des filtres passe-haut et passe-bas, ce qui donne un total de quatre images. Ces quatre images sont à nouveau sous-échantillonnées, résultant en quatre images de la même taille : une image d'approximation S_{j+1} et trois images de détail D_{j+1}^X , où $x \in \{H, V, D\}$ [129], [130].

Pour générer des images en ondelettes à partir d'images histopathologiques de cancers du sein bénins, nous avons suivi plusieurs étapes. Tout d'abord, les images d'entrées au format PNG ont été converties en niveaux de gris. Cette conversion est essentielle car la décomposition en ondelettes est plus efficace sur des images unidimensionnelles (niveaux de gris) que sur des images en couleur (tridimensionnelles avec canaux rouge, vert, et bleu).

En niveaux de gris, les coefficients d'ondelettes permettent de mieux capter les variations d'intensité, ce qui renforce la visibilité des contours, textures et structures de l'image. Nous avons spécifiquement choisi le filtre d'ondelettes « bior1.3 », qui appartient à la famille des ondelettes biorthogonales, parce qu'il offre une meilleure performance en termes de réduction du bruit et de préservation des détails dans les signaux. Cela le rend particulièrement efficace pour le traitement d'images médicales. Le filtre « bior1.3 » se compose d'un filtre passe-bas (h_0) et d'un filtre passe-haut (g_0).

- h_0 (**Filtre passe-bas**) : capture les composantes d'approximation (basses fréquences), qui représentent les variations globales et structures principales de l'image.
- g_0 (**Filtre passe-haut**) : détecte les détails (hautes fréquences), mettant en évidence les contours, textures et fines structures.

Les coefficients d'ondelettes obtenus sont :

- **LL (approximation).**
- **LH, HL, HH (détails).**

La figure V.12 illustre l'image histologique après avoir effectué une transformée en ondelettes (DWT).

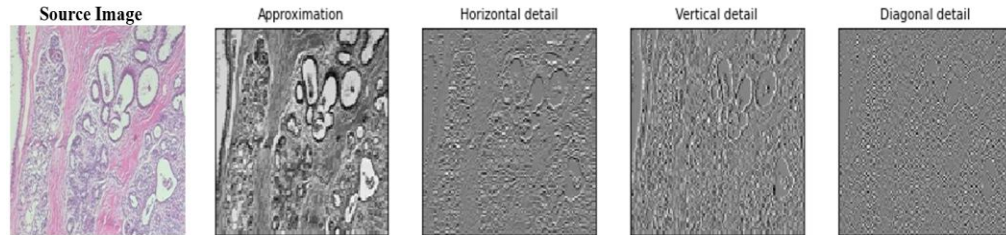


Figure V.12. Exemple Image histologique après transformation en ondelettes (DWT) [123]

V.3.3. Méthodologie

Pour classer les images histopathologiques de l'ensemble de données BreakHis, deux approches complémentaires ont été adoptées : une classification tenant compte des facteurs de grossissement (afin d'intégrer les variations visuelles liées à l'agrandissement) et une classification indépendante du grossissement (visant à renforcer la capacité de généralisation du modèle). Ces deux stratégies sont illustrées dans la Figure V.13.

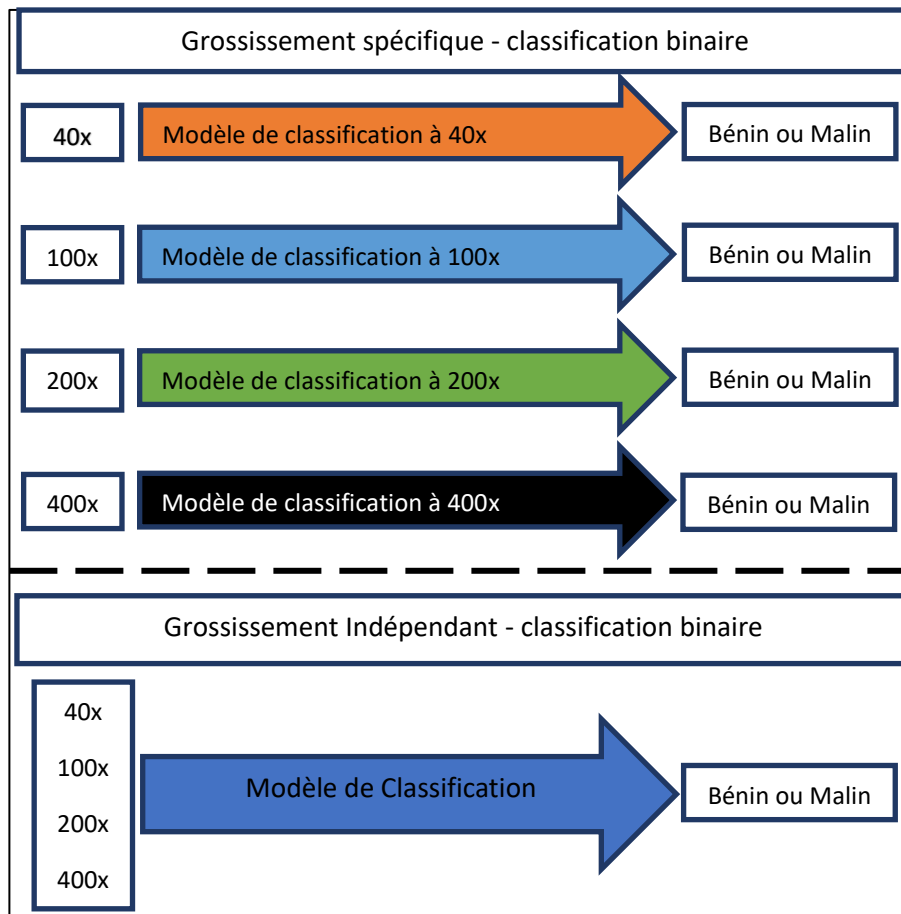


Figure V.13. Approche multi-résolutions pour la classification

Dans un premier temps, 4 % de l'ensemble de données ont été réservés afin de valider la fiabilité du réseau entraîné. Ensuite, une augmentation des données a été effectuée sur les 96 % restants à l'aide de la décomposition en ondelettes, afin de corriger le déséquilibre entre les classes en égalisant la proportion entre la classe minoritaire et la classe majoritaire. L'ensemble de données a ensuite été mélangé aléatoirement et divisé en ensembles d'apprentissage et de validation, avec des proportions respectives de 84 % et 16 %. Enfin, des transformations géométriques ont été appliquées pour générer plusieurs versions des images, utilisées pour l'extraction des caractéristiques par DenseNet201.

Le réseau pré-entraîné DenseNet201 a été choisi pour ses capacités d'extraction de caractéristiques robustes et sa précision supérieure par rapport aux autres modèles d'apprentissage profond. Pour l'adapter à la tâche de classification des images histopathologiques, un fine-tuning a été réalisé en plusieurs étapes. Tout d'abord, les caractéristiques ont été extraites et concaténées à différents niveaux

du modèle (pool3_pool, pool4_pool, conv5_block32_conca) grâce à une mise en commun des moyennes globales (GAP), produisant un vecteur de taille 3072 où chaque composante représente une caractéristique de l'image. Cette opération a permis d'exploiter pleinement les informations du modèle, de réduire la dimensionnalité et d'éviter le surajustement lié à l'utilisation de couches Flatten. La figure V.14 présente une visualisation des différentes cartes de caractéristiques obtenues à partir des différents niveaux du modèle .

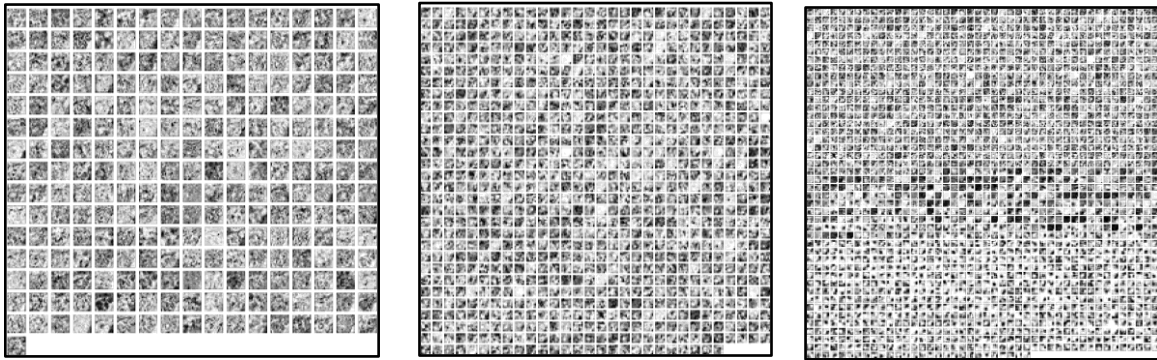


Figure V.14. Illustration des cartes de caractéristiques obtenues à partir des couches (pool3_pool, pool4_pool, conv5_block32_conca)

Pour renforcer la régularisation et minimiser le risque de surajustement, des couches de dropout et de batchNormalization ont été intégrées. Une stratégie de fine-tuning partiel a été appliquée pour adapter DenseNet aux images de la base BreakHis. Les couches situées avant le bloc conv5_block32 ont été gelées afin de préserver les caractéristiques générales apprises lors de l'entraînement initial. Seules les couches finales ont été rendues entraînaibles afin d'ajuster le modèle aux spécificités de la classification binaire. La couche de sortie d'origine, conçue pour 1000 classes, a été remplacée par une nouvelle couche dense adaptée aux deux catégories ciblées. Le modèle a ensuite été entraîné avec un taux d'apprentissage réduit. Ce choix a permis d'ajuster les poids des dernières couches sans altérer les représentations utiles déjà acquises. Sur un total de 18 332 738 paramètres, seuls 294 850 (environ 1.12 Mo) ont été mis à jour. Les 18 037 888 autres (environ 68.81 Mo) ont été conservés. Cette méthode a contribué

à limiter le surapprentissage tout en améliorant la capacité du modèle à distinguer les images bénignes et malignes.

La figure V.15 illustre l'architecture proposée pour la classification des images histopathologiques, mettant en évidence les étapes principales de l'approche adoptée.

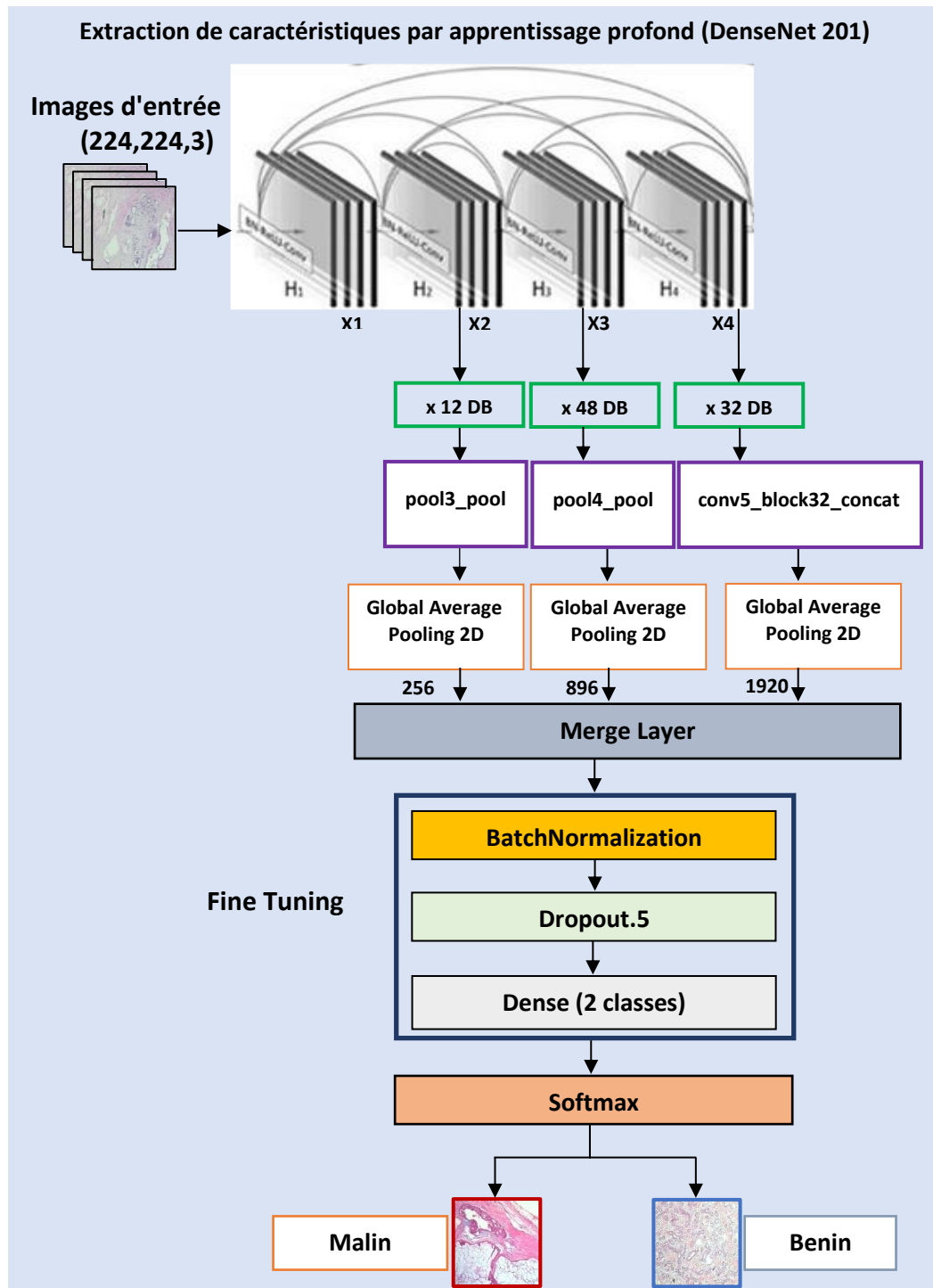


Figure V.15. Modèle proposé pour la classification des images histopathologiques.

Ce processus est détaillé de manière formelle dans l'algorithme, qui décrit méthodiquement les différentes phases de traitement, depuis la préparation des données jusqu'à l'obtention des résultats de classification.

Algorithme : Classification automatisée des images histopathologiques du cancer du sein.

1: Input:

- Ensemble de données pour l'entraînement : $df1$,
- Ensemble de données pour la validation : $df2$,
- Ensemble de données pour le test : $df3$,
- Ep : Epochs,
- bch: Batch size,
- Lr: Taux d'apprentissage,
- N : Couverture par taille de lot,
- X : Poids du modèle CNN pré-entraîné.

2 : Début : Entraînement du cadre

3 : Redimensionner chaque image microscopique des jeux de données à une taille de 224×224 pixels.

4 : Utiliser des techniques d'augmentation pour augmenter la taille de la base de données.

5 : Normaliser les images pour s'assurer que les valeurs sont dans une plage adaptée au modèle.

6 : Extraire les caractéristiques des couches inférieures de DenseNet201 :

- Charger le modèle DenseNet201 pré-entraîné,
- Extraire les caractéristiques des couches intermédiaires :
 - $pool3_pool$ du bloc dense 2,
 - $pool4_pool$ du bloc dense 3,
 - $conv5_block32_concat$ du bloc 4.

7 : Fusionner les caractéristiques extraites à l'aide d'une couche de concaténation.

8 : Appliquer une normalisation par lot (*batch normalization*), un abandon (*dropout*), et une fonction SOFTMAX aux couches ajustées du CNN.

9 : Initialisation du modèle et Fine-Tuning :

- Geler les premières couches du modèle pré-entraîné pour conserver les caractéristiques générales.
- Débloquer les couches supérieures pour les ajuster aux images histopathologiques.

10 : Entraîner le cadre et déterminer les poids initiaux.

11: Boucle d'entraînement:

for $E_p=1$ to E_p do

1. Sélectionner N échantillons pour entraîner le modèle sur l'ensemble d'entraînement $df1$.
2. Effectuer une propagation avant (*forward propagation*) et calculer le coût.
3. Effectuer une rétropropagation (*backpropagation*) et mettre à jour les poids X .

Fin de la boucle.

12: Validation et Test

- Evaluer les performances du modèle sur **df2** (validation).
- Tester le modèle sur **df3** (test) et mesurer la précision.

13 : Fin de l'algorithme.

V.3.4. Paramètres d'entraînement du modèle

La figure V.16 représente les résultats de la simulation du processus d'apprentissage. Différents hyperparamètres sont utilisés pour former le cadre proposé.

- **Poids formés** : Le modèle proposé utilise des poids ImageNet pré-entraînés au début du réseau.
- **Algorithme d'entraînement** : Entraînement avec l'algorithme RMSprop.
- **Taux d'apprentissage** : Taux d'apprentissage initial fixé à $10e-7$, progressivement réduit vers zéro pour améliorer la convergence.

Le `ReduceLROnPlateau` ajuste le taux d'apprentissage pendant l'entraînement, basé sur l'évolution de la précision de validation. Si la précision ne s'améliore pas, le taux d'apprentissage est réduit par un facteur de 0.2, ce qui peut aider à améliorer la convergence du modèle.

- **Fonction de perte** : Utilisation de la cross-entropie catégorique, adaptée pour quantifier la différence entre distributions de probabilité dans un contexte multi-classe.
- **Epoques et taille de lot** : Entraînement réalisé sur 100 époques avec une taille de lot de 35, sans convergence significative constatée après ces 100 époques.

- **Fonction d'activation** : Application de la fonction d'activation SOFTMAX pour transformer les scores bruts en probabilités, facilitant l'interprétation des résultats.
- **DROPOUT** : Pour améliorer les performances du réseau, une couche d'abandon (DROPOUT) a été appliquée avec une probabilité de $P=0,5$.

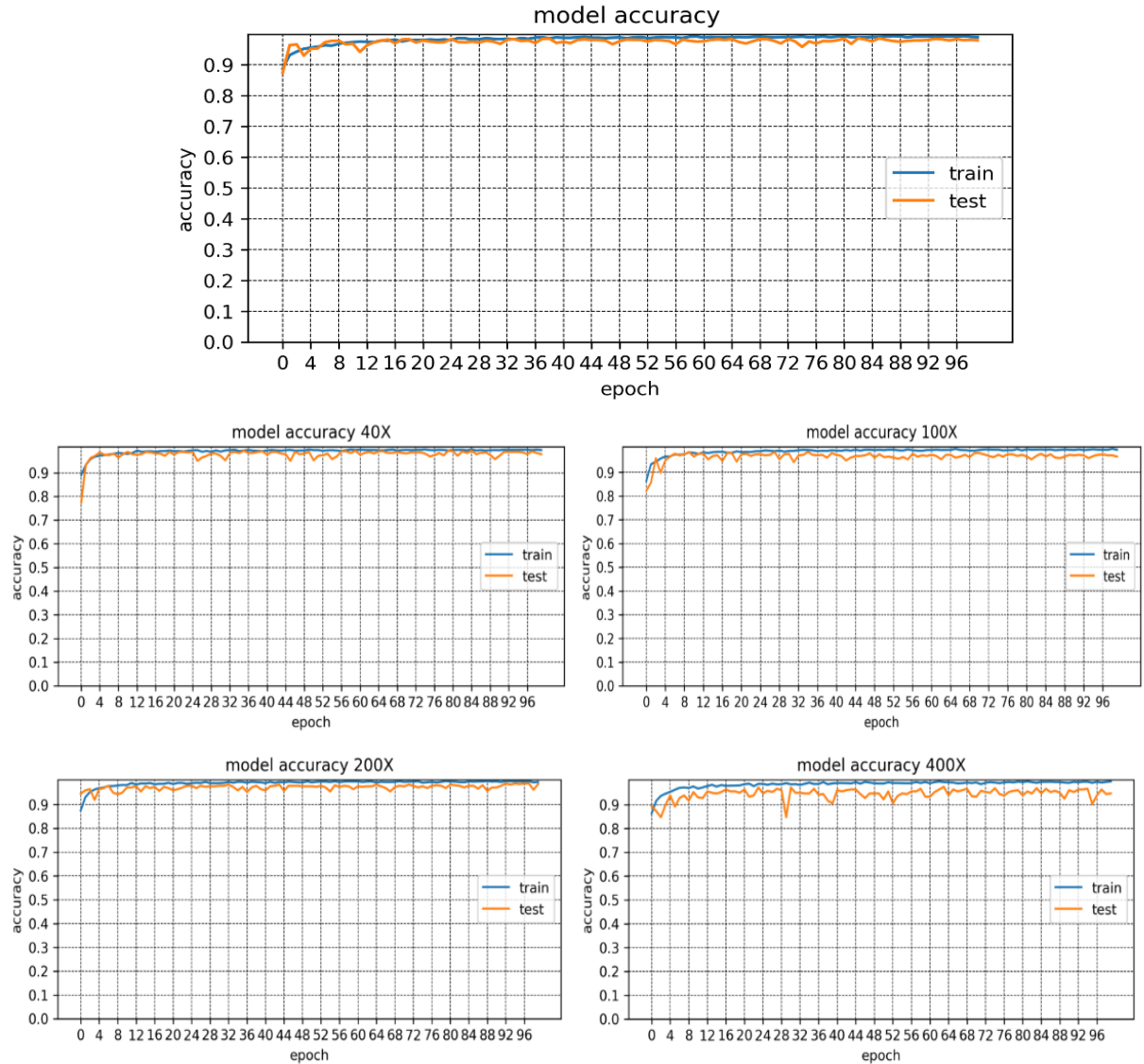


Figure V.16. Evolution de l'apprentissage de la base de données BreakHis avec/sans grossissement.

V.3.5. Analyse des résultats

L'analyse des résultats par facteur de grossissement révèle les excellentes performances du modèle DenseNet201. Selon la matrice de confusion (Tableau V.4), les faux négatifs (cas malins classés comme bénins) et les faux positifs (cas

bénins classés comme malins) sont très négligeables, variant entre 0 et 1. Avec un F1-score dépassant 98 % pour les deux classes, ainsi qu'une précision et un rappel également supérieurs à 96 %, le modèle démontre une capacité remarquable à minimiser les erreurs critiques, susceptibles d'entraîner des conséquences graves pour les patients.

En ce qui concerne la classification indépendante d'agrandissement, les résultats démontrent une efficacité remarquable, avec une précision globale atteignant 99 %. Cette approche vise à évaluer les capacités du modèle à classer correctement les échantillons sans tenir compte des variations de résolution ou de détail offertes par les différents grossissements.

La matrice de confusion associée montre un nombre négligeable de faux négatifs et de faux positifs. Cette observation confirme que le modèle parvient à maintenir des performances élevées tout en réduisant les erreurs graves, telles que la classification erronée d'un échantillon malin comme bénin ou vice-versa.

Les métriques de précision, rappel et F1-score atteignent également 99 %, soulignant la capacité du modèle à extraire efficacement des caractéristiques pertinentes, indépendamment du grossissement. Ces performances stables démontrent que le modèle est non seulement robuste, mais également flexible dans des conditions d'acquisition d'images variées.

La classification indépendante du grossissement présente un intérêt particulier dans des contextes cliniques où les conditions d'acquisition des images peuvent varier en fonction de l'équipement ou des protocoles utilisés. L'absence de la dépendance au grossissement permet également une meilleure généralisation sur des ensembles de données diversifiés.

	Facteurs d'agrandissement	Predict→ Réal↓	Matrice de Confusion			Métriques de performances			
			Benin	Malin	Support	Precision	Recall	F1-Score	Accuracy
Agrandissement spécifique	40x	Benin	25	0	25	0.96	1.00	0.98	0.99
		Malin	1	54	55	1.00	0.98	0.99	
	100x	Benin	25	1	26	1.00	0.96	0.98	0.98
		Malin	0	58	58	0.98	1.00	0.99	
	200x	Benin	25	0	25	0.96	1.00	0.98	0.99
		Malin	1	55	56	1.00	0.98	0.99	
	400x	Benin	24	0	24	0.96	1.00	0.98	0.99
		Malin	1	49	50	1.00	0.98	0.99	
Aggrandissement Independant	Sans facteurs d'agrandissement	Benin	212	5	217	1.00	0.98	0.99	0.99
		Malin	0	218	218	0.98	1.00	0.99	

Tableau V.4. Rapport de classification de l'approche appliquée à l'ensemble de tests du cancer du sein.

La Figure V.17 montre que la précision et le RECALL du modèle restent élevés (96 à 100 %) pour les deux classes, quel que soit le facteur de grossissement, indiquant une excellente capacité à distinguer les échantillons et à détecter les vrais cas, tout en minimisant les faux négatifs.

Les scores F1 atteignent 99 %, démontrant un équilibre optimal entre précision et rappel. Ce résultat confirme la robustesse du modèle de classification.

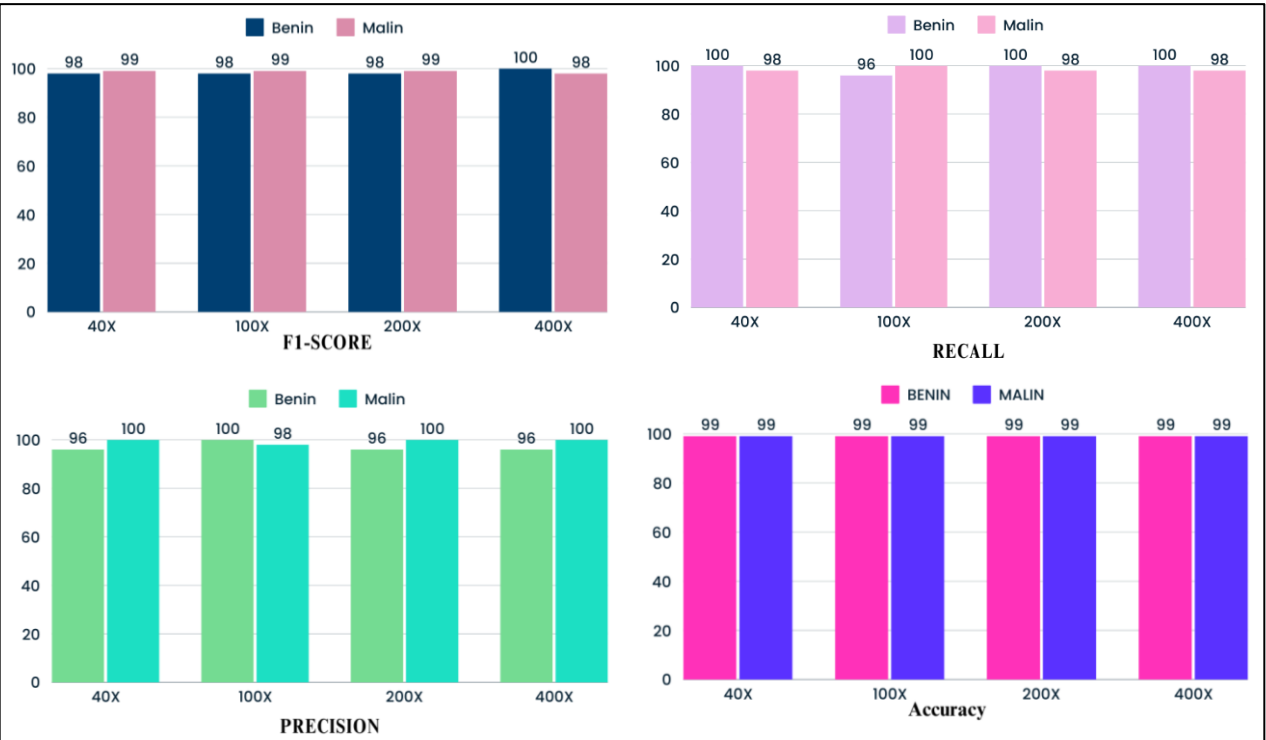


Figure V.17. Performances par métrique et facteur de grossissement pour la classification binaire des images BreakHis

La courbe ROC présentée par la figure V.18, illustre les excellentes performances du modèle DenseNet-201 pour la classification des images histopathologiques, quel que soit le facteur de grossissement (40x, 100x, 200x, 400x) ou l'ensemble global. Les AUC, atteignant 1,0 pour les grossissements 200x et 400x, et dépassant 0,99 pour les autres (40x : 0,9985, 100x : 0,9940), ce qui confirme une discrimination quasi parfaite entre les classes. L'AUC globale de 0,9998 souligne la stabilité et l'efficacité du modèle, en accord avec les métriques élevées observées précédemment.

- **Faible Taux de Faux Positifs (FPR) :** Les courbes montrent que le modèle maintient un très faible taux de faux positifs, avec une True Positive Rate (TPR) élevée dès les premières valeurs de FPR, ce qui est idéal pour les diagnostics médicaux où minimiser les erreurs est primordial.
- **Robustesse aux niveaux 200X et 400X :** Les AUC parfaites pour ces niveaux montrent que le modèle a réussi à extraire efficacement les caractéristiques pertinentes malgré les détails microscopiques complexes présents à ces grossissements.
- **Légère diminution à 100X :** La baisse marginale de l'AUC à 100X peut indiquer une difficulté à discriminer certaines caractéristiques intermédiaires des échantillons bénins ou malins.

Globalement, la courbe ROC confirme l'excellente performance du modèle DenseNet201 pour la classification des images histopathologiques à tous les niveaux de grossissement. Les AUC très élevées (approchant 1) indiquent une capacité exceptionnelle du modèle à distinguer les classes. Toutefois, l'AUC légèrement inférieure à 100X pourrait inciter à des investigations supplémentaires pour améliorer ce niveau spécifique.

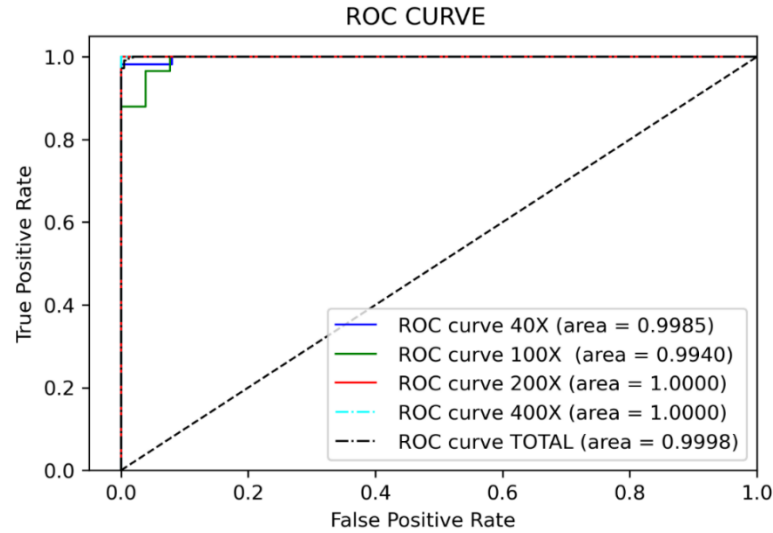


Figure V.18. Performances de la classification binaire de BreakHis avec/sans grossissement.

L'analyse des figures montre la grande fiabilité et stabilité du modèle pour la classification des images histopathologiques du cancer du sein, indépendamment du facteur de grossissement. Le Tableau V.5 compare notre approche à d'autres travaux sur la classification des images BreakHis dans des conditions similaires.

	Reference	Accuracy (%)	Type de Classification	Methode de Classification
1	[1]Khan et al. 2021 [114]	99	Magnification independent-binary classification	Data augmentation+MultiNet
2	Saini and Susan (2020) [115]	40x: 96.5 100x: 94.0 200x: 95.5 400x: 93.0	Magnification specific binary classification	w/BatchNormalization, w/DCGAN samples and w/hyperparameter tuning
3	Liew et al. (2021) [118]	97	Magnification independent-binary classification	data resampling+DenseNet201 and XGBoost
4	Toğaçar et al. (2020) [131]	98.80	Magnification independent-binary classification	Data augmentation/BreastNet
5	Han et al. (2017) [132]	40x: 95.8±3.1 100x: 96.9±1.9 200x: 96.7±2.0 400x: 94.9±1.8	Magnification specific-binary classification	Data over-sampling+CSDCNN model
6	Eshun et al. [129]	40x: 97 100x: 91 200x: 93.6	Magnification specific binary classification	DCGAN+ReseNet50

		400x: 90.2		
7	Proposed Without data augmentation	40x: 92 100x: 91 200x: 88 400x: 91	Magnification specific binary classification	Without data augmentation
8	Djouima et al. (2022) [113]	40x: 96 100x: 95 200x: 88 400x: 92	Magnification specific binary classification	DCGAN augmentation+densnet201 blocks
9	Proposed [123]	99	Magnification independent binary classification	Wavelet transform data augmentation+densnet201blocks
10	Proposed [123]	40x: 99 100x: 98 200x: 99 400x: 99	Magnification specific- binary classification	Wavelet transform data augmentation+densnet201 blocks

Tableau V.5. Comparaison du cadre proposé avec d'autres méthodes sur BreakHis.

V.4. Conclusion

Dans ce chapitre, nous avons présenté deux approches visant à résoudre le déséquilibre des classes et à améliorer la classification des images histopathologiques du cancer du sein issues de la base de données BreakHis. Ces méthodes combinent des stratégies avancées d'enrichissement des données avec l'utilisation du modèle DenseNet201 comme extracteur de caractéristiques, afin d'atteindre des performances optimales pour des applications médicales sensibles.

- Première méthode : DCGAN et DenseNet201

La première approche utilise les DCGAN pour générer des images synthétiques de la classe bénigne. Ces images, associées à une augmentation supplémentaire, réduisent le déséquilibre des classes et enrichissent la diversité des échantillons. Grâce à la concaténation des blocs internes de DenseNet201, cette approche a amélioré la précision, malgré qu'elle puisse introduire des biais si les images générées ne reflètent pas fidèlement les cas réels.

- Deuxième méthode : transformation par ondelettes et DenseNet201

La transformation par ondelettes a permis de résoudre le problème du déséquilibre des classes, une limitation majeure affectant la performance des réseaux d'apprentissage profond appliqués à la classification d'images.

Contrairement aux travaux précédents, centrées sur la fusion spectrale et spatiale des caractéristiques, notre méthode utilise la décomposition par ondelettes pour augmenter les données et réduire les biais liés aux classes minoritaires. Cela a permis d'optimiser les ressources computationnelles et de réduire le temps de convolution, tout en maintenant une haute précision.

En outre, la stratégie de concaténation des caractéristiques extraites des différents blocs de DenseNet201 a capturé efficacement des informations discriminatives à plusieurs niveaux d'abstraction, améliorant ainsi la généralisation et la robustesse du modèle. Les résultats montrent une précision globale de 99 % pour la classification binaire et des précisions entre 98 % et 99 % pour les classifications spécifiques aux différents facteurs de grossissement.

Ces résultats soulignent l'efficacité de la transformation par ondelettes et de la concaténation des caractéristiques pour améliorer les performances des modèles d'apprentissage profond en analyse d'images histopathologiques, offrant un potentiel prometteur pour le diagnostic assisté des maladies.

Les deux méthodes se révèlent complémentaires : l'approche DCGAN génère un volume significatif de données synthétiques, tandis que la transformation par ondelettes assure une meilleure interprétabilité et préservation des caractéristiques spatiales et spectrales.

Conclusion Générale

Cette thèse a exploré de manière approfondie l'application des techniques de Deep Learning au diagnostic médical, en particulier dans le domaine de la classification d'images histopathologiques du cancer du sein. Les travaux présentés mettent en évidence les capacités remarquables des approches basées sur l'intelligence artificielle à relever des défis complexes, en surmontant les limitations des méthodes traditionnelles.

Initialement, le contexte global a été établi en illustrant le rôle croissant de l'IA dans l'optimisation des systèmes de santé. Cette technologie s'est démontrée essentielle pour améliorer la précision et la rapidité des diagnostics médicaux, notamment grâce à sa capacité à analyser des données massives et variées comme les images médicales. Le Deep Learning s'est affirmé comme un outil important dans ce domaine, préparant le terrain pour les chapitres suivants.

Par la suite, une base théorique solide a été fournie en introduisant les concepts fondamentaux de l'apprentissage profond. Ces concepts, notamment la structure des réseaux neuronaux, les stratégies de régularisation et les méthodes d'optimisation, sont indispensables pour comprendre et appliquer efficacement les modèles avancés dans des contextes variés. Cette compréhension théorique a permis d'éclairer les choix architecturaux développés ultérieurement.

L'étude s'est ensuite concentrée, sur les réseaux de neurones convolutifs (CNN), qui constituent un pilier majeur de la vision par ordinateur. Les CNN se sont distingués par leur aptitude à extraire des informations à différents niveaux de granularité, offrant des performances à la pointe de la technologie pour la classification d'images médicales. Leur efficacité, combinée à des techniques telles que l'apprentissage par transfert et l'augmentation des données, a été démontrée à travers divers modèles, incluant DenseNet, qui a été retenu pour nos travaux ultérieurs.

Parallèlement, l'un des principaux défis en imagerie médicale a été abordé : le déséquilibre des classes. Les stratégies présentées, notamment l'augmentation

de données et l'utilisation de méthodes génératives comme les GAN, ont illustré des approches novatrices pour enrichir les ensembles de données et améliorer la fiabilité des modèles. Ces contributions ont été validées par des publications scientifiques, témoignant de leur impact dans le domaine.

A ce stade de la recherche, deux approches ont été explorées pour répondre aux défis liés à la classification des images histopathologiques du cancer du sein. La première méthode, combinant DCGAN pour la génération d'images synthétiques et DenseNet201 comme extracteur de caractéristiques, a permis de corriger efficacement le déséquilibre des classes. Cette stratégie a montré des performances globales prometteuses, notamment aux niveaux de zoom inférieurs (40x et 100x). Toutefois, certaines limites subsistent aux niveaux de grossissement plus élevés (200x et 400x), en particulier pour la détection des échantillons bénins, en raison de la complexité des détails microscopiques et de la représentativité limitée des données générées.

Dans la seconde méthode, une solution complète a été proposée en utilisant la décomposition en ondelettes pour l'augmentation des données et le modèle DenseNet201 pré-entraîné pour l'extraction des caractéristiques. L'augmentation par décomposition en ondelettes, a permis de traiter le déséquilibre des classes, qui affecte négativement les performances des réseaux d'apprentissage profond en classification d'images. Grâce à son extraction multi-échelle, la décomposition en ondelettes capture les détails à différentes résolutions, conservant ainsi les informations essentielles comme les contours, les textures et les motifs. De plus, elle filtre naturellement le bruit, améliorant la qualité des données d'entrée, et renforçant ainsi la robustesse du diagnostic médical.

L'efficacité d'une approche repose aussi sur l'extraction des caractéristiques les plus pertinentes. Le modèle d'apprentissage par transfert DenseNet201, a été choisi pour ses capacités d'extraction de caractéristiques robustes et sa précision supérieure par rapport aux autres modèles d'apprentissage par transfert profond.

Afin d'exploiter pleinement cette capacité, notre approche consistait à concaténer les caractéristiques extraites des blocs spécifiques du modèle DenseNet201 pré-entraîné (*pool3_pool* , *pool4_pool* et *conv5_block32_conca*) grâce à l'application de GAP (Global Average Pooling).

En intégrant les caractéristiques extraites à différents niveaux, cette stratégie a permis de réduire les biais et d'augmenter la robustesse des prédictions, atteignant une précision globale remarquable de 99%.

Pour éviter le surajustement, des couches de régularisation (dropout et batchNormalization) sont ajoutées pour la régularisation. Enfin, le réseau est adapté à la classification binaire en incorporant une couche dense (ou couche entièrement connectée) avec une fonction Softmax à la fin de l'architecture.

L'approche proposée, a été mise en œuvre selon deux taxonomies appliquées à la classification des images histologiques du cancer du sein colorées à l'H&E : la classification binaire spécifique au grossissement et la classification binaire indépendante du grossissement. Les résultats expérimentaux ont atteint une précision de classification de 99 % pour la classification indépendante du grossissement et respectivement 99 %, 98 %, 99 % et 99 % pour les différentes échelles.

Finalement, cette thèse contribue significativement au domaine de l'imagerie médicale en proposant des solutions robustes et innovantes pour améliorer la classification des images histopathologiques.

Perspectives

Pour de futurs travaux, nous proposons l'évaluation des capacités de généralisation de notre modèle à d'autres ensembles de données d'images histopathologiques tels que l'ensemble de données BACH et les ensembles de données couvrant différents types de cancer.

Références Bibliographiques

- [1] RICHENS, Jonathan G. et BUCHARD, Albert. Artificial intelligence for medical diagnosis. In: Artificial Intelligence in Medicine. Cham: Springer International Publishing, 2022. p. 181-201.
- [2] GRABER, Mark L. The incidence of diagnostic error in medicine. *BMJ quality & safety*, 2013, vol. 22, no Suppl 2, p. ii21-ii27.
- [3] DE DOMBAL, F. T., LEAPER, D. J., STANILAND, John R., et al. Computer-aided diagnosis of acute abdominal pain. *Br Med J*, 1972, vol. 2, no 5804, p. 9-13.
- [4] Shagufta, Quazi., Rudra, P., Saha., Manjari, Singh. (2022). Applications of Artificial Intelligence in Healthcare. *Journal of Experimental Biology and Agricultural Sciences*, doi: 10.18006/2022.10(1).211.226
- [5] Kulikowski, C. A. (2019). Beginnings of artificial intelligence in medicine (AIM): computational artifice assisting scientific inquiry and clinical art—with reflections on present aim challenges. *Yearbook of medical informatics*, 28(01), 249-256.
- [6] Dhar, R., Kumar, A., & Karmakar, S. (2022). Artificial intelligence in healthcare: Setting new algo RHYTHM in medicine. *Asian Journal of Medical Sciences*, 13(11), 1-2.
- [7] Turing, A. M. (1950). *Mind*. *Mind*, 59(236), 433-460.
- [8] Rius, A. D. D. M. (2023). Foundations of artificial intelligence and machine learning. In *Artificial Intelligence in Finance* (pp. 2-18). Edward Elgar Publishing.
- [9] McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27(4), 12-12.
- [10] Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach*. Pearson.
- [11] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- [12] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.
- [13] Schmidhuber, J. (2022). Annotated history of modern ai and deep learning. *arXiv preprint arXiv:2212.11279*.
- [14] Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine*, 25(1), 44-56.
- [15] Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Hardt, M., & Dean, J. (2018). Scalable and accurate deep learning with electronic health records. *NPJ digital medicine*, 1(1), 1-10.
- [16] Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural computation*, 3(1), 79-87.

- [17] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639), 115-118.
- [18] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., & Ng, A. Y. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*.
- [19] Alaa, A. M., Bolton, T., Di Angelantonio, E., Rudd, J. H., & Van der Schaar, M. (2019). Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PloS one*, 14(5), e0213653.
- [20] Krittanawong, C., Zhang, H., Wang, Z., Aydar, M., & Kitai, T. (2017). Artificial intelligence in precision cardiovascular medicine. *Journal of the American College of Cardiology*, 69(21), 2657-2664.
- [21] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639), 115-118.
- [22] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- [23] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42, 60-88.
- [24] Juhn, Y., & Liu, H. (2020). Artificial intelligence approaches using natural language processing to advance EHR-based clinical research. *Journal of Allergy and Clinical Immunology*, 145(2), 463-469.
- [25] Kooli, C., & Al Muftah, H. (2022). Artificial intelligence in healthcare: a comprehensive review of its ethical concerns. *Technological Sustainability*, 1(2), 121-131.
- [26] Mini, Diwakaran., D., Surendran. (2023). Breast Cancer Prognosis Based on Transfer Learning Techniques in Deep Neural Networks. *Information Technology and Control*, doi: 10.5755/j01.itc.52.2.33208
- [27] Habib, Izadkhah., Jaber, Karimpour. (2023). A New Deep-Learning-Based Model for Breast Cancer Diagnosis from Medical Images. *Diagnostics*, doi: 10.3390/diagnostics13111944
- [28] Khan Mamun, M. M. R., & Elfouly, T. (2023). Detection of Cardiovascular Disease from Clinical Parameters Using a One-Dimensional Convolutional Neural Network. *Bioengineering*, 10(7), 796.
- [29] Ahmed, Mohamed, Ali, Omer, Hassan., Mohamed, Elhoseny., Mohammed, Kayed. (2023). A novel and accurate deep learning-based Covid-19 diagnostic model for

- heart patients. *Signal, Image and Video Processing*, doi: 10.1007/s11760-023-02561-8
- [30] Hameed, Z., Zahia, S., Garcia-Zapirain, B., Javier Aguirre, J., & Maria Vanegas, A. (2020). Breast cancer histopathology image classification using an ensemble of deep learning models. *Sensors*, 20(16), 4373.
- [31] Wang, S. H., & Zhang, Y. D. (2020). DenseNet-201-based deep neural network with composite learning factor and precomputation for multiple sclerosis classification. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 16(2s), 1-19.
- [32] Suganya, A., & Aarthy, S. L. (2023). Application of Deep Learning in the Diagnosis of Alzheimer's and Parkinson's disease-A Review. *Current medical imaging*.
- [33] I., Yoo., J.J., Yoo., D., Kim., Inyoung, Youn., H., Kim., M., Youn., J., H., Won., W., Cho., Yoon-Won, Myong., S, Kim., Rebecca, Yu., S.M., Kim., K., Kim., S., B., Lee. (2023). Deep learning-based algorithm versus physician judgement for diagnosis of myopathy and neuropathy from needle electromyography. *medRxiv*, doi: 10.1101/2023.01.13.23284511
- [34] Prajoy, Podder., Fatema, Binte, Alam., M., Rubaiyat, Hossain, Mondal., Md., Junayed, Hasan., Subrato, Bharati. (2023). Rethinking Densely Connected Convolutional Networks for Diagnosing Infectious Diseases. *Computers*, doi: 10.3390/computers12050095
- [35] Alyaa, Hamel, Sfayyih., Ahmad, Sabry., Shymaa, Mohammed, Jameel., N., Sulaiman., Safanah, M., Raafat., Amjad, J., Humaidi. (2023). Acoustic-Based Deep Learning Architectures for Lung Disease Diagnosis: A Comprehensive Overview. *Diagnostics*, doi: 10.3390/diagnostics13101748
- [36] Prajoy, Podder., Sanchita, Rani, Das., M., Rubaiyat, Hossain, Mondal., Subrato, Bharati., Azra, Maliha., Md., Junayed, Hasan., Farzin, Piltan. (2023). LDDNet: A Deep Learning Framework for the Diagnosis of Infectious Lung Diseases. *Sensors*, doi: 10.3390/s23010480
- [37] Alberto, Solano., Marcelino, Martínez-Sober. (2023). Deep Learning Architectures for Diagnosis of Diabetic Retinopathy. *Applied Sciences*, doi: 10.3390/app13074445
- [38] Muhammet, Fatih, Aslan., Kadir, Sabanci. (2023). A Novel Proposal for Deep Learning-Based Diabetes Prediction: Converting Clinical Data to Image Data. *Diagnostics*, doi: 10.3390/diagnostics13040796
- [39] He, Xu., Qunli, Zheng., Jingshu, Zhu., Zuoling, Xie., Haitao, Cheng., Peng, Liu., Yimu, Ji. (2022). A Deep Learning Model Incorporating Knowledge Representation Vectors and Its Application in Diabetes Prediction. *Disease Markers*, doi: 10.1155/2022/7593750

- [40] Theodore, Aptekarev., Vladimir, Sokolovsky., Evgeny, Furman., Natalia, Kalinina., Gregory, B., Furman. (2023). Application of deep learning for bronchial asthma diagnostics using respiratory sound recordings. *PeerJ*, doi: 10.7717/peerj-cs.1173
- [41] Das, A., Mohanty, M. N., Mallick, P. K., Tiwari, P., Muhammad, K., & Zhu, H. (2021). Breast cancer detection using an ensemble deep learning method. *Biomedical Signal Processing and Control*, 70, 103009.
- [42] Khan, S. I., Shahrir, A., Karim, R., Hasan, M., & Rahman, A. (2022). MultiNet: A deep neural network approach for detecting breast cancer through multi-scale feature fusion. *Journal of King Saud University-Computer and Information Sciences*, 34(8), 6217-6228.
- [43] Liew, X. Y., Hameed, N., & Clos, J. (2021). An investigation of XGBoost-based algorithm for breast cancer classification. *Machine Learning with Applications*, 6, 100154.
- [44] Mewada, H. K., Patel, A. V., Hassaballah, M., Alkinani, M. H., & Mahant, K. (2020). Spectral-spatial features integrated convolution neural network for breast cancer classification. *Sensors*, 20(17), 4747.
- [45] Saini, M., & Susan, S. (2020). Deep transfer with minority data augmentation for imbalanced breast cancer dataset. *Applied Soft Computing*, 97, 106759.
- [46] LeCun, Y. (2016). L'apprentissage profond, une révolution en intelligence artificielle. *La lettre du Collège de France*, (41), 13.
- [47] Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik*, 29(2), 102-127.
- [48] Pandey, M., Fernandez, M., Gentile, F., Isayev, O., Tropsha, A., Stern, A. C., & Cherkasov, A. (2022). The transformational role of GPU computing and deep learning in drug discovery. *Nature Machine Intelligence*, 4(3), 211-221.
- [49] Bengio, Y. (2019). La révolution de l'apprentissage profond. *Interstices*.
- [50] Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN computer science*, 2(6), 420.
- [51] Schmidgall, S., Ziaei, R., Achterberg, J., Kirsch, L., Hajiseyedrazi, S., & Eshraghian, J. (2024). Brain-inspired learning in artificial neural networks: a review. *APL Machine Learning*, 2(2).
- [52] Tan, K. H., & Lim, B. P. (2018). The artificial intelligence renaissance: deep learning and the road to human-level machine intelligence. *APSIPA Transactions on Signal and Information Processing*, 7, e6.
- [53] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- [54] Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. *Computer Science Review*, 40, 100379.

- [55] Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). Introduction to artificial neural network. *International Journal of Engineering and Innovative Technology (IJEIT)*, 2(1), 189-194.
- [56] Iqbal, H., Sarker., Iqbal, H., Sarker. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. doi: 10.1007/S42979-021-00815-1
- [57] Hagan, M. T., Demuth, H. B., & Jesús, O. D. (2002). An introduction to the use of neural networks in control systems. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 12(11), 959-985.
- [58] Ghamri, S. E. (2016). Commande Neuronale d'un Robot Mobile Non-Holonyme avec glissement des roues. Mémoire Présenté pour l'obtention du diplôme de MAGISTER en Électronique, Centre universitaire Batna2.
- [59] Pirdashti, M., Curteanu, S., Kamangar, M. H., Hassim, M. H., & Khatami, M. A. (2013). Artificial neural networks: applications in chemical engineering. *Reviews in Chemical Engineering*, 29(4), 205-239.
- [60] Ketkar, N., Moolayil, J., Ketkar, N., & Moolayil, J. (2020). *Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch*. Apress LP.
- [61] Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., & Miao, Y. (2021). Review of image classification algorithms based on convolutional neural networks. *Remote Sensing*, 13(22), 4712.
- [62] Li, Y., Huang, C., Ding, L., Li, Z., Pan, Y., & Gao, X. (2019). Deep learning in bioinformatics: Introduction, application, and perspective in the big data era. *Methods*, 166, 4-21.
- [63] Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). Introduction to artificial neural network. *International Journal of Engineering and Innovative Technology (IJEIT)*, 2(1), 189-194.
- [64] Hinton, G., LeCun, Y., & Bengio, Y. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- [65] Raymond, C., Chen, Q., Xue, B., & Zhang, M. (2023, July). Fast and Efficient Local-Search for Genetic Programming Based Loss Function Learning. In *Proceedings of the Genetic and Evolutionary Computation Conference* (pp. 1184-1193).
- [66] Ho, Y., & Wookey, S. (2019). The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling. *IEEE access*, 8, 4806-4813.
- [67] Kavalerov, I., Czaja, W., & Chellappa, R. (2019). cGANs with multi-hinge loss. *arXiv preprint arXiv:1912.04216*.
- [68] Meyer, G. P. (2021). An alternative probabilistic interpretation of the huber loss. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 5261-5269).

- [69] Admojo, F. T., & Sulistya, Y. I. (2022). Analisis performa algoritma Stochastic Gradient Descent (SGD) dalam mengklasifikasi tahu berformalin. *Indonesian Journal of Data and Science*, 3(1), 1-8.
- [70] Tian, Y., Zhang, Y., & Zhang, H. (2023). Recent advances in stochastic gradient descent in deep learning. *Mathematics*, 11(3), 682.
- [71] Muskan, Agarwal., Kanwarpartap, Singh, Gill., Sonal, Malhotra., Swati, Devliyal. (2024). Analysing Digits with Sequential Convolutional Neural Networks and Adam Optimizer. doi: 10.1109/iciteics61368.2024.10625453
- [72] Diederik, P., Kingma., Jimmy, Ba. (2015). 1. Adam: A Method for Stochastic Optimization.
- [73] Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.
- [74] Chollet, F. (2021). Deep learning with Python. Simon and Schuster.
- [75] Zachary, Rabin., Jim, Davis., Benjamin, Lewis., Matthew, Scherreik. (2024). 1. Overfitting In Contrastive Learning ? arXiv.org, doi: 10.48550/arxiv.2407.15863
- [76] Anil, Kumar, Prajapati, Anil., Umesh, Kumar, Singh. (2023). 5. An Optimal Solution to the Overfitting and Underfitting Problem of Healthcare Machine Learning Models. *Journal of Systems Engineering and Information Technology (JOSEIT)*, Doi: 10.29207/joseit.v2i2.5460
- [77] (2023). Should I Stop or Should I Go: Early Stopping with Heterogeneous Populations. doi : 10.48550/arxiv.2306.11839
- [78] Hinton, G., Srivastava, N., & Swersky, K. (2012). "Dropout: A Simple Way to Prevent Overfitting." *The Journal of Machine Learning Research*, 15(1), 1929-1958.
- [79] Liu, Z., Xu, Z., Jin, J., Shen, Z., & Darrell, T. (2023). "Dropout Reduces Underfitting." *Proceedings of the 40th International Conference on Machine Learning*.
- [80] Ioffe, S. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167.
- [81] Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43-76.
- [82] Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural networks*, 106, 249-259.
- [83] Valkenborg, D., Geubbelmans, M., Rousseau, A. J., & Burzykowski, T. (2023). Supervised learning. *American Journal of Orthodontics and Dentofacial Orthopedics*, 164(1), 146-149.
- [84] Kaur, A. (2016). "An introduction to neural network." *International Journal of Applied Research*, 2(8), 164-167. DOI: 10.22271/allresearchjournal.2016.v2.i8.182.

- [85] Saint-Cirgue, G. (2019). Apprendre le Machine Learning en une semaine. Machine Learnia, Londres, 16.
- [86] Deepanshu, Mehta. (2020). 8. State-of-the-Art Reinforcement Learning Algorithms. International Journal of Engineering Research and, doi: 10.17577/IJERTV8IS120332
- [87] Er., Sandeep, Kaur. (2023). (1) Image Classification Using Deep Learning. Indian Scientific Journal of Research in Engineering and Management, doi: 10.55041/ijjsrem27484
- [88] Zhao, X., Wang, L., Zhang, Y. et al. A review of convolutional neural networks in computer vision. Artif Intell Rev 57, 99 (2024). <https://doi.org/10.1007/s10462-024-10721-6>
- [89] Mohammed, F. A., Tune, K. K., Assefa, B. G., Jett, M., & Muhie, S. (2024). Medical Image Classifications Using Convolutional Neural Networks: A Survey of Current Methods and Statistical Modeling of the Literature. Machine Learning and Knowledge Extraction, 6(1), 699-735.
- [90] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- [91] Chen, L.; Li, S.; Bai, Q.; Yang, J.; Jiang, S.; Miao, Y. Review of Image Classification Algorithms Based on Convolutional Neural Networks. Remote Sens. 2021, 13, 4712. <https://doi.org/10.3390/rs13224712>
- [92] Dumoulin, V.; Visin, F. A guide to convolution arithmetic for deep learning. arXiv 2016, arXiv:1603.07285
- [93] Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G. Recent Advances in Convolutional Neural Networks. Pattern Recognit. 2018, 77, 354–377.
- [94] Nair, V.; Hinton, G. Rectified Linear Units Improve Restricted Boltzmann Machines Vinod Nair. 2010, Volme 27, pp. 807–814. Available online: <https://dl.acm.org/doi/10.5555/3104322.3104425> (accessed on 1 June 2021).
- [95] Boureau, Y.L.; Ponce, J.; Lecun, Y. A Theoretical Analysis of Feature Pooling in Visual Recognition. 2010, pp. 111–118. Available online: <https://dl.acm.org/doi/10.5555/3104322.3104338> (accessed on 1 June 2021).
- [96] Jonghoon, Jin., Aysegul, Dundar., Eugenio, Culurciello. (2014). 3. Flattened Convolutional Neural Networks for Feedforward Acceleration. arXiv: Neural and Evolutionary Computing,
- [97] Hinton, G.E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R.R. Improving neural networks by preventing co-adaptation of feature detectors. arXiv 2012, arXiv:1207.0580

- [98] Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation*, 29(9), 2352-2449.
- [99] Christian, Raymond., Qi, Chen., Bing, Xue., Mengjie, Zhang. (2024). 1. Fast and Efficient Local Search for Genetic Programming Based Loss Function Learning. doi: 10.48550/arxiv.2403.00865
- [100] Nixon, Jeremy., Sohl-Dickstein, Jascha, Narain., Metz, Luke, Shekerjian., Freeman, Christian, Daniel., Maheswaranathan, Niruban. (2020). 2. Training optimizer neural networks.
- [101] Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE* 1998, 86, 2278–2324.
- [102] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
- [103] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [104] Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 15 October 2015; pp. 1–9.
- [105] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).
- [106] Lu, Z., Yang, J., & Liu, Q. (2017). Face image retrieval based on shape and texture feature fusion. *Computational Visual Media*, 3, 359-368.
- [107] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- [108] Gross, S., & Wilber, M. (2016). Training and investigating residual nets. *Facebook AI Research*, 6(3).
- [109] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [110] Hwang, S. Y., & Kim, J. J. (2022). A Deep Learning Hierarchical Ensemble for Remote Sensing Image Classification. *Computers, Materials & Continua*, 72(2).
- [111] Salehi, A. W., Khan, S., Gupta, G., Alabduallah, B. I., Almjally, A., Alsolai, H., ... & Mellit, A. (2023). A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope. *Sustainability*, 15(7), 5930.
- [112] Mukhlif, A. A., Al-Khateeb, B., & Mohammed, M. A. (2023). Incorporating a novel dual transfer learning approach for medical images. *Sensors*, 23(2), 570.

- [113] Djouima, H., Zitouni, A., Megherbi, A. C., & Sbaa, S. (2024). Data Augmentation by Wavelet Transform for Breast Cancer Based on Deep Learning. *Revue d'Intelligence Artificielle*, 38(4).
- [114] Khan, S.I., Shahrir, A., Karim, R., Hasan, M., Rahman, A. (2022). MultiNet: A deep neural network approach for detecting breast cancer through multi-scale feature fusion. *Journal of King Saud University-Computer and Information Sciences*, 34(8): 6217-6228. <https://doi.org/10.1016/j.jksuci.2021.08.004>
- [115] Saini, M., Susan, S. (2020). Deep transfer with minority data augmentation for imbalanced breast cancer dataset. *Applied Soft Computing*, 97: 106759. <https://doi.org/10.1016/j.asoc.2020.106759>
- [116] Yari, Y., Nguyen, H. (2020). A state-of-the-art deep transfer learning-based model for accurate breast cancer recognition in histology images. In *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*, Cincinnati, OH, USA, pp. 900-905. <https://doi.org/10.1109/BIBE50027.2020.00153>
- [117] Vo, D.M., Nguyen, N.Q., Lee, S.W. (2019). Classification of breast cancer histology images using incremental boosting convolution networks. *Information Sciences*, 482: 123-138. <https://doi.org/10.1016/j.ins.2018.12.089>
- [118] Liew, X.Y., Hameed, N., Clos, J. (2021). An investigation of XGBoost-based algorithm for breast cancer classification. *Machine Learning with Applications*, 6: 100154. <https://doi.org/10.1016/j.mlwa.2021.100154>
- [119] LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature*, 521(7553): 436-444. <https://doi.org/10.1038/nature14539>
- [120] Hameed, Z., Zahia, S., Garcia-Zapirain, B., Javier Aguirre, J., Maria Vanegas, A. (2020). Breast cancer histopathology image classification using an ensemble of deep learning models. *Sensors*, 20(16): 4373. <https://doi.org/10.3390/s20164373>
- [121] Prajit, Sengupta., Anant, Mehta., Prashant, Singh, Rana. (2023). 1. Enhancing Performance of Deep Learning Models with a Novel Data Augmentation Approach. doi: 10.1109/icccnt56998.2023.10308298
- [122] Fatima, M., Inamdar., Mahendra, Kumar, Singar., Preeti, Naval., Ashutosh, Pattanaik., M.P., Santhoshkumar., Maheswaran, S. (2023). 4. Improving Accuracy and Speed of Deep Learning Algorithms for Real-Time Data Analysis in Machine Learning. doi: 10.1109/icpsitiags59213.2023.10527455
- [123] Djouima, H., Zitouni, A., Megherbi, A. C., & Sbaa, S. (2022, May). Classification of breast cancer histopathological images using DensNet201. In *2022 7th International Conference on Image and Signal Processing and their Applications (ISPA)* (pp. 1-6). IEEE.

- [124] Aksac, A., Demetrick, D.J., Ozyer, T., Alhajj, R. (2019). BreCaHAD: A dataset for breast cancer histopathological annotation and diagnosis. *BMC Research Notes*, 12(1): 1-3. <https://doi.org/10.1186/s13104-019-4121-7>
- [125] Bayramoglu, N., Kannala, J., Heikkilä, J. (2016). Deep learning for magnification independent breast cancer histopathology image classification. In 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, IEEE, pp. 2440-2445. <https://doi.org/10.1109/ICPR.2016.7900002>
- [126] Bardou, D., Zhang, K., Ahmad, S.M. (2018). Classification of breast cancer based on histology images using convolutional neural networks. *IEEE Access*, 6: 24680-24693.
- [127] Zhu, C., Song, F., Wang, Y., Dong, H., Guo, Y., Liu, J. (2019). Breast cancer histopathology image classification through assembling multiple compact CNNs. *BMC Medical Informatics and Decision Making*, 19(1): 1-17. <https://doi.org/10.1186/s12911-019-0913-x>
- [128] Benhammou, Y., Achchab, B., Herrera, F., Tabik, S. (2020). BreakHis based breast cancer automatic diagnosis using deep learning: Taxonomy, survey and insights. *Neurocomputing*, 375: 9-24. <https://doi.org/10.1016/j.neucom.2019.09.044>
- [129] Eshun, R. B., Islam, A. K., & Bikdash, M. (2024). A deep convolutional neural network for the classification of imbalanced breast cancer dataset. *Healthcare Analytics*, 100330.
- [130] Lipiński, P. (2012). On domain selection for additive, blind image watermarking. *Bulletin of the Polish Academy of Sciences. Technical Sciences*, 60(2): 317-321. <https://doi.org/10.2478/v10175-012-0042-5>
- [131] Toğaçar, M., Özkurt, K.B., Ergen, B., Cömert, Z. (2020). BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer. *Physica A: Statistical Mechanics and its Applications*, 545: 123592. <https://doi.org/10.1016/j.physa.2019.123592>
- [132] Han, Z., Wei, B., Zheng, Y., Yin, Y., Li, K., Li, S. (2017). Breast cancer multi-classification from histopathological images with structured deep learning model. *Scientific Reports*, 7(1): 4172. <https://doi.org/10.1038/s41598-017-04075-z>