



HAL
open science

Détection et classification multi-label de maladies pulmonaires par apprentissage automatique à partir d'images de radiographie thoracique

Aya Hage Chehade

► **To cite this version:**

Aya Hage Chehade. Détection et classification multi-label de maladies pulmonaires par apprentissage automatique à partir d'images de radiographie thoracique. Traitement du signal et de l'image [eess.SP]. Université d'Angers, 2024. Français. NNT : 2024ANGE0020 . tel-04840343

HAL Id: tel-04840343

<https://theses.hal.science/tel-04840343v1>

Submitted on 16 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT

DE
L'UNIVERSITÉ D'ANGERS

ÉCOLE DOCTORALE N° 641
*Mathématiques et Sciences et Technologies du numérique
de l'Information et de la Communication*
Spécialité : *Signal, Image, Vision*

Par

Aya HAGE CHEHADE

**Détection et classification multi-label de maladies pulmonaires par
apprentissage automatique à partir d'images de radiographie tho-
racique**

Thèse présentée et soutenue à Angers, le 20/09/2024

Unité de recherche : Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS)

Thèse N° : « si pertinent »

Rapporteurs avant soutenance :

Chaouki DIAB Professeur des Universités, Institut Supérieur des Sciences Appliquées et Économiques, CNAM Liban
Pierre-Henri CONZE Maître de Conférences (HDR), IMT Atlantique, Brest

Composition du Jury :

Président : Mohamed QUAFAROU Professeur des Universités (HDR), Université Aix-Marseille

Examineurs : Jean-Marie MARION Maître de Conférences, Université Catholique de l'Ouest, Angers
 Mohamed QUAFAROU Professeur des Universités (HDR), Université Aix-Marseille
 Nadine ABDALLAH Maîtresse de Conférences, ISEN OUEST Nantes

Dir. de thèse : Pierre CHAUVET Professeur UCO (HDR), Angers
Co-dir. de thèse : Mohamad OUEIDAT Professeur des Universités, Université Libanaise
Co-encadrant de thèse : Nassib ABDALLAH Docteur (Ingénieur de recherche), LaTIM, INSERM UMR 1101, Brest

Invité :

Ayman AL FALOU Professeur ISEN OUEST (HDR), Brest

REMERCIEMENTS

Je remercie toutes les personnes qui ont contribué à la bonne réalisation de cette thèse.

Je tiens à exprimer ma profonde reconnaissance et mes sincères remerciements à mon directeur de thèse Pr. Pierre CHAUVET, mon co-directeur Pr. Mohamad OUEIDAT, et mes deux encadrants Dr. Nassib ABDALLAH et Dr. Jean-Marie MARION, pour leur disponibilité, leur encadrement continu, ainsi que pour leurs précieux conseils et leurs remarques constructives qu'ils ont pu me prodiguer durant toute la période de ma thèse. Leur rigueur scientifique et leur confiance en moi m'ont toujours motivée et incitée à donner le meilleur de moi-même et à mener à bien mon projet de thèse. Travailler avec eux a été un plaisir et une des expériences les plus enrichissantes. Je leur adresse ma profonde gratitude et reconnaissance.

Je tiens à remercier Pr. Chaouki DIAB et Pr. Pierre-Henri CONZE d'avoir bien voulu accepter d'être rapporteur, ainsi que Pr. Mohamed QUAFAROU, Dr. Jean-Marie MARION, Dr. Nadine ABDALLAH et Pr. Ayman AL FALOU pour avoir accepté d'être membre du jury et d'évaluer ma thèse.

Je tiens aussi à remercier tous les membres du LARIS et tous mes collègues doctorants pour les bons moments partagés. Leur dynamisme et leur engagement envers la recherche ont créé un environnement stimulant et propice à l'apprentissage.

Enfin, je souhaite exprimer ma gratitude envers ma famille, en particulier envers ma mère, ma sœur et mon frère. Leur soutien inconditionnel tout au long de mon parcours académique ainsi que leur encouragement constant ont été des piliers solides durant les moments de doute et de difficulté. Je n'oublie pas non plus mon père qui est toujours vivant dans mon cœur. Je leur suis profondément reconnaissante et je leur dédie ce succès avec toute ma gratitude.

TABLE DES MATIÈRES

Introduction	11
I Etat de l’art : Pathologies pulmonaires et leur classification par des modèles d’Intelligence Artificielle (IA)	15
1 Pathologies pulmonaires et techniques d’imagerie	17
1.1 Introduction	17
1.2 Anatomie des poumons humains	17
1.3 Pathologies pulmonaires	18
1.3.1 La pneumonie	19
1.3.2 L’épanchement pleural	21
1.3.3 La consolidation	23
1.4 Techniques d’imagerie pulmonaire	24
1.4.1 Radiographie thoracique (CXR)	25
1.4.2 Tomodensitométrie (CT)	26
1.5 Diagnostic des pathologies pulmonaires à partir des CXRs	28
1.5.1 Diagnostic de la pneumonie	28
1.5.2 Diagnostic de l’épanchement pleural	28
1.5.3 Diagnostic de la consolidation	29
1.6 Conclusion	30
2 Bases de données, prétraitement des images et classification des pathologies pulmonaires par les modèles d’apprentissage profond	31
2.1 Introduction	31
2.2 Acquisition des images (Bases de données)	32
2.3 Prétraitement des images	35
2.3.1 Amélioration de l’image	36
2.3.2 Segmentation sémantique de l’image	42
2.3.3 Suppression des structures osseuses	43

2.4	Modèles d'Intelligence Artificielle	43
2.4.1	Réseaux de neurones convolutifs	45
2.4.2	Architecture U-Net pour la segmentation sémantique d'images	51
2.4.3	Architectures d'attention	52
2.5	Techniques face à la limitation des données	54
2.6	Evaluation du modèle d'apprentissage	58
2.7	Classification binaire des pathologies spécifiques	60
2.8	Classification multi-classe des pathologies	64
2.9	Classification multi-label des pathologies	66
2.10	Conclusion	69

II Contributions : Classification des maladies pulmonaires par des modèles d'apprentissage automatique 71

3	Amélioration du diagnostic des maladies pulmonaires : Analyse d'images et exploitation des données cliniques pour une meilleure classification 73
3.1	Introduction 73
3.2	Approche proposée 75
3.3	Identification des groupes d'images avec l'Analyse Hiérarchique en Clusters 76
3.4	Prétraitement des images avec le Masque Flou 79
3.5	Modèle de classification 80
3.5.1	Réseaux de neurones convolutifs 80
3.5.2	Mécanisme d'attention 81
3.6	Résultats et discussion 82
3.7	Conclusion 87
4	Amélioration de la classification des maladies pulmonaires : Une nouvelle approche de prétraitement basée sur CycleGAN pour la réduction des artefacts 89
4.1	Introduction 89
4.2	Approche proposée 91
4.3	Prétraitement des images avec l'égalisation d'histogramme 92
4.4	Séparation des images avec et sans artefacts à l'aide du clustering 94
4.4.1	Extraction des caractéristiques 95

4.4.2	K-means clustering	96
4.5	Réduction des artefacts avec CycleGAN	99
4.6	Modèle de classification	103
4.6.1	Caractéristiques cliniques	104
4.6.2	Convolutional Block Attention Module (CBAM)	104
4.7	Résultats et discussion	107
4.8	Conclusion	113
5	Nouvelle approche de segmentation d'images de radiographie thoracique basée sur CycleGAN pour une meilleure classification multi-label des maladies thoraciques	115
5.1	Introduction	115
5.2	Base de données	116
5.3	Approche proposée	118
5.4	Génération des masques	119
5.4.1	U-Net pour la segmentation et génération des masques	120
5.4.2	Génération avancée de masques de pathologie	121
5.5	Extraction des caractéristiques	125
5.6	Sélection des caractéristiques	128
5.6.1	Clustering avec l'Analyse Hiérarchique en Clusters	128
5.6.2	Sélection des caractéristiques statistiquement significatives	130
5.7	Modèles de classification	133
5.7.1	Forêt aléatoire (Random Forest)	134
5.7.2	Extreme Gradient Boosting (XGBoost)	135
5.8	Résultats et discussion	136
5.9	Conclusion	143
	Conclusion et Perspectives	145
	Liste de publications	151
	Liste des abréviations	153
	Bibliographie	155
	Table des figures	182

Table des tableaux

185

RÉSUMÉ

Les maladies pulmonaires représentent une cause majeure de décès dans le monde. Le diagnostic précoce est crucial pour prévenir la détérioration de l'état du patient et améliorer les chances de rétablissement et de survie. La radiographie thoracique (CXR) est la technique d'imagerie médicale la plus couramment utilisée pour le diagnostic de nombreuses anomalies thoraciques, en raison de sa simplicité, son accessibilité, et son coût-efficacité. Cependant l'interprétation correcte des CXRs constitue un défi majeur pour le médecin. En effet, chaque radiographie contient de nombreuses pathologies pulmonaires potentielles dont beaucoup partagent des caractéristiques visuelles similaires, ce qui peut entraîner des difficultés d'interprétation. D'où l'intérêt d'automatiser la détection et la classification des maladies pulmonaires à l'aide des technologies d'Intelligence Artificielle qui ont ouvert des voies prometteuses pour relever ces défis, jouant un rôle important dans le domaine biomédical.

Ainsi, l'objectif principal de cette thèse est de développer des modèles d'apprentissage automatique pour améliorer la classification des maladies pulmonaires à partir des images de radiographie thoracique.

Pour ce faire, différentes étapes sont nécessaires et ont été appliquées dans ce travail de recherche : l'acquisition des données, le prétraitement des images, la construction des modèles d'IA et enfin l'évaluation de ces modèles. La base de données ChestX-ray14, largement utilisée dans la littérature et contenant des images de radiographie thoracique ainsi que des données cliniques, est utilisée dans notre recherche. Le prétraitement des images a pour but d'améliorer la qualité des images et de supprimer le bruit afin d'accroître les performances de classification et donc l'efficacité des modèles d'apprentissage.

Cette thèse présente plusieurs contributions. Au-delà de l'analyse des images, les données cliniques associées sont également examinées, puisqu'elles peuvent considérablement influencer l'interprétation des images et la prise de décision concernant la présence d'une pathologie. En effet, prendre en compte ces métadonnées contribue à affiner le modèle de classification et donc le diagnostic selon le profil du patient, ce qui améliore l'efficacité du diagnostic.

Ensuite, une nouvelle approche de prétraitement basée sur CycleGAN est développée

afin de minimiser l'effet du bruit produit par les artefacts dans les images de radiographie thoracique. En effet, les images CXR peuvent souvent contenir des artefacts tels que des objets étrangers, des dispositifs médicaux, des fils, et des électrodes. Ces éléments non pertinents introduisent du bruit qui peut compromettre la visibilité des structures anatomiques essentielles et rendre difficile la distinction des caractéristiques pathologiques pertinentes. La réduction de l'impact de ces éléments non pertinents permet au modèle de se concentrer plus précisément sur les caractéristiques diagnostiques pertinentes, améliorant ainsi la précision de la classification des maladies pulmonaires.

Dans ce même contexte, pour le prétraitement, une nouvelle approche de segmentation basée sur CycleGAN est développée afin de générer des masques qui incluent les zones affectées par la pathologie dans la région d'intérêt. En effet, les approches traditionnelles telles que U-Net, bien que largement utilisées pour la segmentation des images médicales, échouent à inclure les régions spécifiquement affectées par les pathologies, essentielles pour une classification précise, dans les masques. Notre approche de génération de masques de pathologies assure que les caractéristiques radiomiques sont extraites des zones les plus indicatives des pathologies, permettant ainsi de concentrer l'analyse sur les aspects pathologiques des images et améliorant de manière significative la performance de la classification multi-label des maladies pulmonaires par rapport à l'utilisation du modèle U-Net.

Finalement, une nouvelle approche de sélection de caractéristiques est développée pour identifier a priori les caractéristiques statistiquement les plus significatives pour les maladies avant de faire la classification. Cette étape permet d'éliminer les caractéristiques redondantes ou non essentielles, garantissant que seules les caractéristiques les plus pertinentes sont retenues. Cette méthode améliore l'efficacité du modèle de classification en ciblant et en analysant précisément les caractéristiques clés qui distinguent les différentes conditions pathologiques, optimisant ainsi la performance du modèle et renforçant la fiabilité des diagnostics.

Pour conclure, quatre contributions ont été développées au cours de cette thèse. Les approches proposées présentent des résultats prometteurs améliorant de manière significative la performance de la classification binaire et multi-label des maladies pulmonaires, ce qui contribuera à optimiser la qualité des diagnostics et à améliorer la prise en charge des patients.

INTRODUCTION

Les maladies pulmonaires représentent un ensemble de troubles qui compromettent la santé et le fonctionnement des poumons, menant potentiellement à une diminution significative de la capacité respiratoire. Les maladies pulmonaires ont un impact significatif sur la santé mondiale, se plaçant comme la troisième cause principale de décès dans le monde [1] et comptant pour environ cinq millions de décès chaque année [2].

Le diagnostic et le traitement précoces sont cruciaux pour réduire les risques potentiellement mortels des pathologies pulmonaires, tout en optimisant les chances de rétablissement et de survie des individus déjà affectés. La détection et la prise en charge à temps sont essentielles pour non seulement arrêter la progression de la maladie pulmonaire mais également pour minimiser les risques de complications graves.

L'imagerie médicale joue un rôle crucial dans le diagnostic des maladies pulmonaires, offrant plusieurs techniques pour une évaluation précise et guidant vers des stratégies de traitement adaptées. Les modalités d'imagerie comme la radiographie thoracique ou chest X-ray (CXR), la tomographie par ordinateur (CT) et l'imagerie par résonance magnétique (IRM) servent d'outils diagnostiques essentiels, permettant la visualisation des poumons et l'identification des anomalies.

La radiographie thoracique est la technique d'imagerie médicale la plus couramment utilisée pour le diagnostic de nombreuses anomalies thoraciques [3], en raison de sa simplicité, son accessibilité, sa non-invasivité et son coût-efficacité. En plus, la CXR est capable de fournir des résultats rapides, essentiels dans le traitement d'urgence des patients, et permet l'évaluation de la gravité des maladies ainsi que l'identification de complications potentielles [4].

Cependant, l'examen manuel des images de radiographie thoracique par les radiologues est une tâche ardue et prend beaucoup de temps, notamment en raison de la nature complexe des images CXR [5]. En effet, identifier avec précision les différents types de maladies pulmonaires à partir des images CXR représente un défi significatif pour les médecins à cause du grand nombre de pathologies pulmonaires potentielles qui peuvent y être présentes et dont beaucoup partagent des caractéristiques visuelles similaires. Cette complexité peut entraîner des erreurs d'interprétation, des diagnostics manqués et poten-

tiellement des erreurs diagnostiques critiques. Cela met en évidence l'importance d'adopter des méthodes avancées pour améliorer la précision et l'efficacité du processus diagnostique à partir des images de radiographie thoracique.

Récemment, l'intérêt pour les technologies d'Intelligence Artificielle (IA) dans le domaine médical a connu une augmentation marquée, grâce à leur capacité d'accélérer l'analyse des données et d'améliorer le diagnostic par imagerie médicale. L'IA, notamment via l'apprentissage automatique (ML) et l'apprentissage profond (DL), a amélioré l'efficacité des évaluations des radiographies thoraciques, offrant une classification précise des maladies pulmonaires et augmentant de manière significative la précision des diagnostics cliniques [6]. Grâce à l'analyse automatisée d'images pathologiques, les systèmes de Diagnostic Assisté par Ordinateur (CAD) ont le potentiel de réduire considérablement la charge de travail des radiologues et aider à la prise de décision clinique [7, 8]. Les avancées récentes dans le domaine de l'apprentissage profond ont capté une attention considérable en raison de leur vaste potentiel pour automatiser l'extraction et la classification des caractéristiques, des processus cruciaux pour résoudre des problèmes complexes nécessitant une analyse fine et détaillée. Les algorithmes basés sur l'apprentissage profond ont démontré un succès significatif et une efficacité dans plusieurs tâches de vision par ordinateur, y compris la classification d'images [9], la détection de maladies [10, 11], le diagnostic médical [12, 13] et l'analyse de soins de santé [14].

L'objectif principal de cette thèse est de développer des modèles d'IA pour améliorer la classification et le diagnostic des maladies pulmonaires à partir des images de radiographie thoracique. Pour atteindre ce but, des techniques de prétraitement sont utilisées pour améliorer la qualité des images pour une classification automatique plus précise des pathologies.

Ce manuscrit de thèse est composé de deux parties.

La première partie se concentre sur une étude bibliographique sur le sujet de thèse. Cette partie se divise en deux chapitres.

Le chapitre 1 présente l'anatomie des poumons humains et les pathologies qui affectent les poumons. Ce chapitre examine également les techniques d'imagerie pulmonaire utilisées pour diagnostiquer ces pathologies, mettant l'accent sur la radiographie thoracique.

Le chapitre 2 aborde les bases de données d'images CXR, les techniques de prétraitement des images et les modèles de classification des pathologies pulmonaires.

La deuxième partie présente nos contributions dans la construction des modèles d'IA pour la classification des pathologies pulmonaires à partir d'images de radiographie tho-

racique. Cette partie se compose de trois chapitres.

Le chapitre 3 présente une approche qui exploite les données cliniques disponibles dans l'ensemble de données afin d'identifier les groupes présentant des caractéristiques cliniques communes. Ensuite, pour chaque groupe, le contraste des images a été amélioré et un modèle d'apprentissage profond intégrant une couche d'attention a été développé pour la classification binaire des maladies pulmonaires.

Le chapitre 4 présente une méthode de prétraitement qui se base sur l'application du modèle CycleGAN sur les images bruitées pour minimiser les artefacts dus à la présence de bruits non liés à la pathologie. Suite à cette étape de prétraitement, le modèle DenseNet-121 est appliqué, intégrant les caractéristiques cliniques et un mécanisme d'attention, pour effectuer la classification binaire des maladies pulmonaires.

Enfin, dans ce même contexte de prétraitement, le chapitre 5 présente une méthode de segmentation pour générer des masques précis, visant à inclure les zones affectées par la pathologie dans la région d'intérêt, afin d'extraire les caractéristiques de Radiomiques de ces régions d'intérêt. Des modèles d'apprentissage automatique tel que les forêts aléatoires et XGBoost sont ensuite utilisés pour effectuer la classification multi-label des pathologies thoraciques.

PREMIÈRE PARTIE

**Etat de l'art : Pathologies
pulmonaires et leur classification par
des modèles d'Intelligence Artificielle
(IA)**

PATHOLOGIES PULMONAIRES ET TECHNIQUES D'IMAGERIE

1.1 Introduction

Outre le cœur et le cerveau, les poumons sont les organes les plus actifs dans l'organisme puisqu'ils travaillent en permanence tout au long de la vie depuis le premier souffle jusqu'au dernier pour fournir à l'organisme l'Oxygène nécessaire à son fonctionnement. Etant en contact direct avec le milieu extérieur, les poumons sont exposés en continu à un grand nombre de particules et de micro-organismes qui sont nocifs la plupart du temps pour le corps humain. Plusieurs maladies peuvent en découler et seront discutées par la suite dans ce chapitre ainsi que les principaux types d'imageries médicales qui aident à diagnostiquer ou à identifier ces maladies.

1.2 Anatomie des poumons humains

L'être humain possède deux poumons logés dans la cage thoracique, situés au-dessus du diaphragme. Ce dernier étant une vaste nappe musculaire qui sépare le thorax de la cavité abdominale. Les poumons sont protégés par les côtes et appartiennent au système respiratoire.

Les poumons gauche et droit, sont enveloppés par la plèvre, membrane constituée de deux feuillets qui les maintient collés à la paroi thoracique. Ils sont distincts et séparés en leur centre par le médiastin, à travers lequel passe l'oesophage, la trachée et les bronches, et où se trouvent le cœur, les gros vaisseaux ainsi que des vaisseaux lymphatiques et des ganglions [15]. Chaque poumon est constitué de lobes, trois du côté droit et seulement deux du côté gauche en raison de la place occupé par le cœur. Chacun est relié à la trachée, par laquelle l'air transite, via la bronche principale. Chaque bronche se divise dans le poumon en bronches plus petites qui se dirigent vers les lobes pulmonaires, et ces dernières se

ramifient de plus en plus pour se terminer en bronchioles. Au bout des bronchioles se trouvent les alvéoles qui forment de petits sacs, comme l'illustre la Figure 1.1. Chacune de ces alvéoles est entourée de vaisseaux très fins, appelés capillaires, lesquelles sont délimités par une membrane mince. C'est à travers cette membrane que se déroulent les échanges gazeux : l'oxygène présent dans l'air respiré passe des alvéoles au sang dans les capillaires, tandis que le dioxyde de carbone est éliminé du sang vers les alvéoles pour être expiré. Ce processus est crucial pour l'apport d'oxygène aux tissus du corps et l'élimination du dioxyde de carbone, soutenant ainsi les fonctions vitales [16].

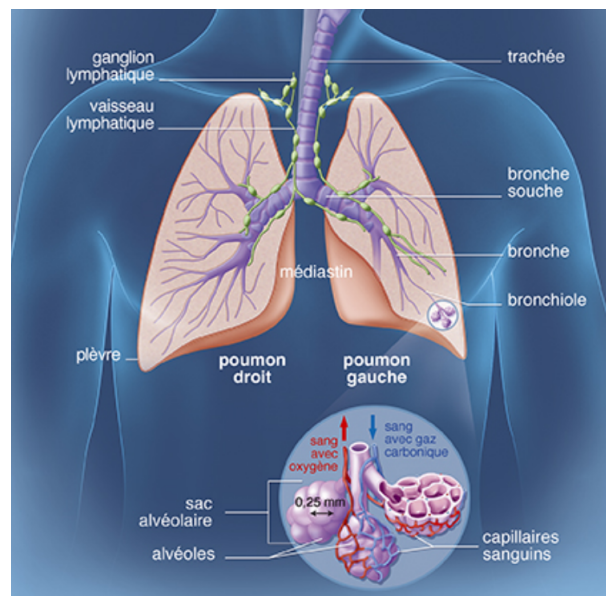


FIGURE 1.1 – Anatomie des poumons humains (extrait de[17]).

1.3 Pathologies pulmonaires

Les poumons interagissent directement avec l'environnement extérieur, inspirant de l'air qui peut être chargé de nombreux micro-organismes et particules. Ces derniers sont acheminés directement au système respiratoire, où ils entrent en contact prolongé avec les tissus pulmonaires. Ceci peut entraîner différentes pathologies graves, particulièrement lorsque le système immunitaire n'est pas en mesure de les éliminer efficacement par phagocytose.

1.3.1 La pneumonie

La pneumonie est une maladie infectieuse qui provoque l'inflammation des alvéoles pulmonaires, affectant un ou les deux poumons. Cette condition peut entraîner l'accumulation de liquide ou de pus dans les sacs aériens (Figure 1.2), ce qui se manifeste par une toux productive de mucosités ou de pus, de la fièvre, des frissons et de difficultés respiratoires. Divers organismes, notamment des bactéries, des virus et des champignons, peuvent provoquer une pneumonie. Son degré de gravité varie, allant de formes bénignes à des cas extrêmement sévères, voire mortels. D'après l'Organisation mondiale de la santé, la pneumonie représente la cause infectieuse de mortalité majeure chez les enfants dans le monde, entraînant le décès d'environ 740 180 enfants âgés de moins de cinq ans en 2019 [18].

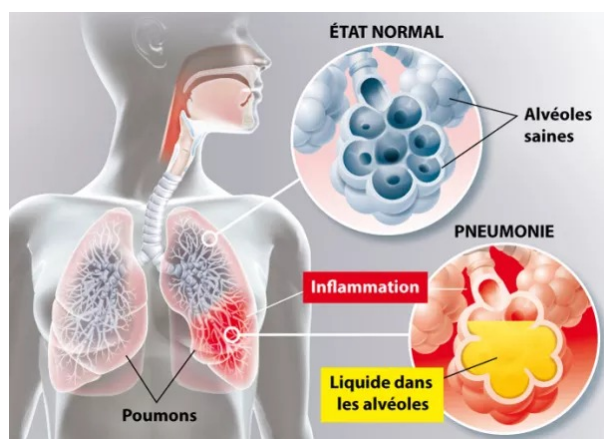


FIGURE 1.2 – Illustration de la pneumonie (extrait de [19]).

1. Causes de la pneumonie :

La majorité des cas de pneumonie sont causés par des micro-organismes infectieux, notamment des bactéries, des virus, des mycoplasmes et des champignons microscopiques. Les pneumonies se classent principalement en deux catégories en fonction de leur cause : les pneumonies bactériennes, qui sont les plus courantes, et les pneumonies virales.

De nombreuses bactéries peuvent être à l'origine de pneumonies, parmi lesquelles les plus couramment rencontrées incluent le pneumocoque (*Streptococcus pneumoniae*), *Hæmophilus influenzae* type B et le staphylocoque doré (*Staphylococcus aureus*). Toutefois, la mortalité hospitalière pour la pneumonie s'élève à

30,3% chez les patients de plus de 75 ans, avec le pneumocoque représentant 41,2% des bactéries isolées dans les cas de pneumonie [20].

Près de la moitié des pneumonies sont causées par des virus, en particulier chez les enfants. Les pneumonies virales sont souvent moins graves que les pneumonies bactériennes. Cependant, elles peuvent entraîner des complications comme la bronchite ou la bronchiolite, et dans certains cas, se compliquer d'une surinfection bactérienne, nécessitant une hospitalisation.

2. **Symptômes de la pneumonie :**

Les symptômes initiaux de la pneumonie comprennent souvent une toux sèche qui devient productive au fil des jours. La fièvre monte rapidement, pouvant atteindre jusqu'à 41°C. Les personnes malades ressentent un essoufflement et une faiblesse même pour des activités quotidiennes, accompagnés de douleurs thoraciques à l'inspiration ou lors de la toux, de douleurs musculaires et de céphalées. Une mauvaise haleine est également fréquente, et un bleuissement des lèvres et des ongles peut apparaître, indiquant une oxygénation insuffisante du sang [21].

En cas d'aggravation, la respiration du patient s'accélère (plus de 30 inspirations par minute) et le pouls au repos devient anormalement rapide (plus de 120 battements par minute). Le patient peut expérimenter des épisodes de délire ou de confusion.

Dans les pneumonies virales, contrairement aux formes bactériennes, les alvéoles pulmonaires ne se remplissent pas de liquide ou de pus ; c'est l'inflammation des parois des alvéoles et des bronchioles qui perturbe la respiration. Par conséquent, la toux demeure généralement sèche, sans production d'expectorations (crachats).

3. **Diagnostic de la pneumonie :**

Pour poser le diagnostic de pneumonie, le médecin s'appuie sur l'examen clinique, l'écoute des sons respiratoires à l'aide d'un stéthoscope et l'imagerie pulmonaire via radiographie ou tomodensitométrie (scanner). L'analyse des expectorations ou du sang pour identifier l'agent pathogène responsable n'est effectuée que lors d'une hospitalisation [21].

4. **Traitement de la pneumonie :**

Le traitement des pneumonies bactériennes est basé sur l'administration d'antibiotiques dès la confirmation du diagnostic. La durée de ce traitement est généralement d'une à deux semaines, et il est crucial de le suivre scrupuleusement et jusqu'à son terme pour éviter toute rechute. Environ quatre à six semaines après la fin du trai-

tement antibiotique, une radiographie de suivi est souvent réalisée par le médecin afin de s'assurer de la guérison complète des poumons.

Pour les pneumonies virales, telles que celles causées par la grippe ou le COVID-19, le traitement vise principalement à surveiller l'évolution de la maladie et à alléger les symptômes tels que la fièvre et les douleurs à l'aide de paracétamol. Dans certains cas, un antibiotique peut être prescrit en prévention d'une possible surinfection bactérienne [21].

1.3.2 L'épanchement pleural

L'épanchement pleural ou *Effusion*, également connu sous le nom de pleurésie, se caractérise par l'accumulation anormale de liquide dans la cavité pleurale, qui est l'espace entre les deux couches de la plèvre enveloppant les poumons (Figure 1.3). Selon sa composition, ce fluide peut être classé en transsudat, généralement aqueux, ou en exsudat, riche en protéines [22]. L'épanchement peut affecter le côté gauche ou droit, en fonction du poumon concerné.

La nature du liquide de l'épanchement pleural dépend de sa cause. Il peut être constitué de sang (hémothorax) en cas de lésion thoracique, de liquide lymphatique (chylothorax) suite à une lésion ou une obstruction du canal lymphatique principal, de pus (empyème) dans les cas de pneumonie ou d'abcès pulmonaire, ou encore d'urine (urinorax) en présence d'uretères obstrués [23].

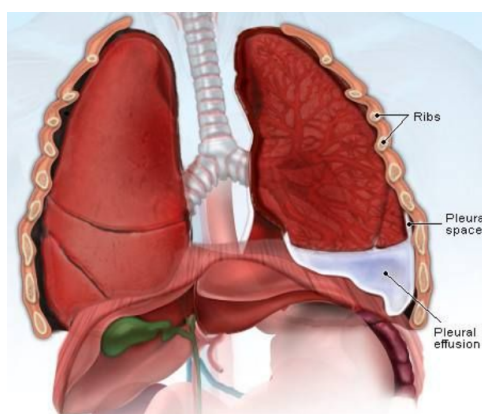


FIGURE 1.3 – Illustration de l'épanchement pleural (extrait de [24]).

1. Causes de l'épanchement pleural :

Les causes de l'épanchement pleural peuvent être diverses [22, 25] :

- Une insuffisance cardiaque, rénale, pulmonaire (embolie pulmonaire) ou hépatique (cirrhose) ;
- Un cancer : une tumeur cancéreuse avec des métastases au niveau de la plèvre ;
- Un traumatisme : fracture des côtes ;
- Les effets secondaires de certains traitements médicamenteux ;
- Une inflammation ou une infection : polyarthrite rhumatoïde, tuberculose, lupus érythémateux systémique, pneumonie, etc. ;
- Une pancréatite ;
- Une dialyse péritonéale.

2. Symptômes de l'épanchement pleural :

Les symptômes associés à l'épanchement pleural dépendent largement du volume de liquide accumulé dans la cavité pleurale. Ainsi, lorsque la quantité de liquide est faible, il est courant que la personne ne présente aucun symptôme. Les manifestations cliniques deviennent apparentes lorsque l'épanchement pleural devient plus important. Lorsqu'un épanchement pleural atteint un stade modéré à important, deux symptômes principaux sont fréquemment observés : des douleurs thoraciques et une difficulté à respirer (dyspnée).

La douleur, de nature pleurétique, peut être constante et s'aggraver avec l'effort physique, la toux, ou en position allongée. Elle peut aussi n'apparaître qu'en cas d'effort ou lors de la toux. Cette douleur peut être ressentie au niveau de la paroi thoracique ou se propager à l'épaule, au cou et à l'abdomen. De plus, l'accumulation significative de liquide dans la cavité pleurale entraîne une compression des poumons, engendrant des difficultés respiratoires, en particulier lors d'efforts (dyspnée). D'autres symptômes peuvent accompagner l'épanchement pleural, tels que la fatigue, due aux réveils fréquents provoqués par l'essoufflement, la fièvre en cas d'infection, ou encore une toux sèche [26].

3. Diagnostic de l'épanchement pleural :

Pour établir un diagnostic d'épanchement pleural, le pneumologue initie par un examen physique du patient, suivi de la prescription de plusieurs tests diagnostiques, dont la radiographie des poumons. En effet cette dernière constitue la première approche diagnostique. Elle peut révéler une accumulation significative de liquide dans la cavité pleurale en cas de pleurésie. Cet examen peut être complété par une échographie et/ou un scanner des poumons pour identifier de faibles quantités de liquide. La thoracentèse est un examen médical supplémentaire pouvant être réa-

lisé pour plus d'information sur la nature du liquide. Cette procédure implique le prélèvement d'une petite quantité de liquide pleural à l'aide d'une aiguille insérée dans la cage thoracique. Le liquide prélevé est ensuite examiné en laboratoire pour déterminer sa nature (pus, sang, liquide lymphatique, urine) et pour détecter la présence éventuelle de bactéries ou de cellules cancéreuses [27].

4. **Traitement de l'épanchement pleural :**

La prise en charge de l'épanchement pleural est conditionnée par l'identification préalable de sa cause. Dans les cas où l'épanchement est mineur, il se peut qu'aucun traitement spécifique ne soit nécessaire, hormis la prise d'antalgiques pour soulager la douleur. Une simple surveillance médicale peut suffire à vérifier que le liquide pleural se résorbe naturellement. Néanmoins, il reste essentiel d'identifier et de traiter la cause sous-jacente de cet épanchement léger. Pour des épanchements pleuraux modérés à sévères, la mise en place d'un drainage thoracique s'avère indispensable. Pour ce faire, un tube de drainage est inséré entre deux côtes sous anesthésie locale.

Lorsque l'épanchement pleural est causé par une pneumonie, le traitement inclut, outre le drainage, l'administration d'antibiotiques par voie intraveineuse, et parfois une assistance respiratoire. Dans le cas où l'épanchement est lié à un cancer de la plèvre, des complications peuvent survenir telles que la réaccumulation rapide du liquide malgré le drainage, nécessitant potentiellement le maintien permanent d'un drain. Des traitements spécifiques au cancer sont alors prescrits.

Dans certains cas, une intervention chirurgicale, comme la thoracotomie, peut être considérée. Cette procédure vise à fusionner les deux couches de la plèvre (pleurodèse) pour empêcher la réformation de liquide entre elles. Si cette mesure s'avère insuffisante, une résection d'une portion des tissus pulmonaires peut être nécessaire (pleurectomie) [28].

1.3.3 **La consolidation**

La consolidation pulmonaire se produit lorsque l'air habituellement présent dans les alvéoles pulmonaires est remplacé par un liquide, un solide ou une autre matière telle que du pus, du sang, de l'eau, des contenus gastriques ou des cellules [29].

Les causes de la consolidation pulmonaire sont variées, mais la pneumonie, une infection des poumons, est l'une des raisons les plus courantes. D'autres causes incluent

l'œdème pulmonaire (accumulation de liquide), l'hémorragie alvéolaire (saignement dans les alvéoles), ou certains types de cancer du poumon.

Les symptômes associés à une consolidation pulmonaire dépendent de sa cause, mais peuvent inclure de la toux, de la fièvre, une respiration sifflante, une difficulté à respirer et une douleur dans la poitrine.

Le diagnostic et la gestion de la consolidation pulmonaire nécessitent une évaluation clinique détaillée, comprenant souvent l'histoire médicale du patient, un examen physique, et des tests diagnostiques supplémentaires pour identifier la cause sous-jacente et déterminer le traitement approprié.

Le traitement de la consolidation dépendra de la condition sous-jacente identifiée, allant des antibiotiques pour la pneumonie bactérienne à des interventions plus complexes pour le cancer ou d'autres affections graves.

La consolidation est donc un signe radiologique important, reflétant la présence d'une anomalie pulmonaire qui nécessite une attention médicale pour résoudre la condition sous-jacente et restaurer la fonction pulmonaire normale [30].

1.4 Techniques d'imagerie pulmonaire

L'imagerie médicale joue un rôle crucial dans le diagnostic des pathologies et notamment les pathologies pulmonaires, offrant un aperçu de l'anatomie et de la pathophysiologie pulmonaires. La radiographie thoracique (CXR), le scanner (CT) et l'imagerie par résonance magnétique (IRM) sont les modalités d'imagerie à disposition des professionnels de la santé et les plus connues pour la détection et la caractérisation des anomalies pulmonaires. Ces techniques d'imagerie fournissent une visualisation détaillée des structures pulmonaires et des potentielles anomalies, jouant ainsi un rôle essentiel dans la détermination des stratégies de traitement. Elles permettent non seulement de diagnostiquer la présence d'une pathologie, mais également de suivre son évolution et de surveiller la réponse au traitement, guidant ainsi les décisions thérapeutiques vers les options de traitement les plus appropriées. La radiographie thoracique et la tomographie par ordinateur sont les techniques les plus fréquemment utilisées.

1.4.1 Radiographie thoracique (CXR)

La radiographie thoracique ou *Chest X-Ray* (CXR) constitue un outil de diagnostic essentiel offrant une vision détaillée du cœur, des poumons, des vaisseaux sanguins et les os. Cette technique consiste à projeter des rayons X à travers le corps, capturant ainsi des images détaillées sur une plaque de capture numérique. La quantité de radiation absorbée varie selon les parties du corps, ce qui permet la distinction entre les différentes structures. Les os, qui absorbent une grande quantité de radiation, apparaissent en blanc, tandis que les tissus moins denses, qui absorbent une quantité plus faible de radiation, comme le cœur, apparaissent en nuances de gris. Les espaces aériens, principalement ceux des poumons ressortent en noir, soulignant l'absence de matière dense. Cette imagerie est indispensable pour le diagnostic et la surveillance de diverses affections, notamment l'hypertrophie cardiaque (augmentation du volume du cœur), signalant une possible insuffisance cardiaque, ou l'œdème pulmonaire, indiquant une accumulation anormale de liquide dans les poumons ainsi que pour identifier d'autres anomalies pulmonaires [31].

Les radiographies thoraciques peuvent être réalisées selon deux orientations principales : postéro-antérieure (PA) et antéro-postérieure (AP) [32].

1. Radiographie thoracique PA (Postéro-Antérieur) :

La radiographie thoracique PA est l'orientation standard pour les examens de radiographie du thorax. Dans cette configuration, le patient est positionné de sorte que sa poitrine soit contre la plaque de radiographie, avec le dos tourné vers la source de rayons X. Les rayons X traversent le corps du patient de l'arrière vers l'avant avant de frapper la plaque de radiographie (Figure 1.4). Cette orientation est privilégiée lorsque le patient est capable de se tenir debout, car elle offre plusieurs avantages diagnostiques, notamment une représentation plus précise de la taille et de la forme du cœur. Dans la vue PA, le cœur et le médiastin subissent moins de distorsion car ils se trouvent plus près de la plaque de radiographie, ce qui réduit l'amplification de leur taille due à la perspective.

2. Radiographie thoracique AP (Antéro-Postérieur) :

En contraste avec la vue PA, la radiographie thoracique AP est effectuée lorsque le patient est face à la source de rayons X, le dos étant orienté vers la plaque de radiographie. Dans cette orientation, les rayons X pénètrent le corps de l'avant vers l'arrière (Figure 1.4). La vue AP est souvent utilisée dans des contextes où le patient est incapable de se tenir debout ou de se déplacer aisément, comme dans

les unités de soins intensifs ou lors d'examens d'urgence. Bien que cette orientation permette de réaliser des radiographies dans des situations moins idéales, elle peut induire une certaine distorsion, particulièrement une amplification apparente de la taille du cœur, due à sa position plus éloignée de la plaque de radiographie, rendant parfois l'interprétation des images plus délicate.

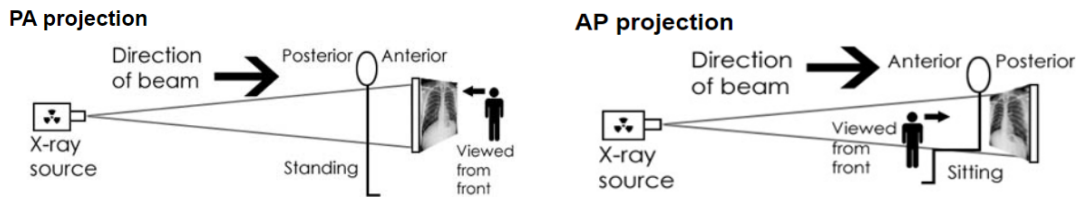


FIGURE 1.4 – Procédure d'acquisition des radiographies thoraciques PA et AP (extrait de [32]).

1.4.2 Tomodensitométrie (CT)

Le scanner ou tomodensitométrie ou *Computed Tomography* (CT scans) est une technique d'imagerie médicale qui utilise des rayons X pour créer des images détaillées en coupe du corps. Contrairement à une radiographie standard qui fournit une vue bidimensionnelle, le scanner CT combine une série de prises de vue radiographiques prises sous différents angles pour générer des images en coupe transversale.

Le scanner est composé d'un anneau équipé d'un tube à rayons X et de détecteurs numériques. Ce dispositif réalise un balayage circulaire autour de la région corporelle ciblée, émettant des rayons X à faible dose vers la zone à examiner (Figure 1.5). Il mesure l'absorption des rayons X, différente suivant la nature des tissus qu'ils traversent. Cette différence d'absorption peut être mesurée par une cellule photoélectrique ou détecteur. Les variations dans l'absorption des rayons, liées à la densité des tissus, sont ensuite converties en données numériques par l'appareil. Le scanner produit des images en coupes anatomiques fines et en série (c'est-à-dire en tranches fines). Ces images sont traitées par un ordinateur qui permet une reconstitution en deux ou en trois dimensions (2D ou 3D) des zones du corps étudiées permettant ainsi aux médecins de visualiser avec précision et en détail les organes, les os, les tissus mous et les vaisseaux sanguins. Pour améliorer la visibilité de certains organes, un produit de contraste iodé peut être administré, accentuant le contraste sur les images et facilitant l'interprétation diagnostique. Ce processus rend le scanner particulièrement efficace pour explorer en détail l'anatomie interne et détecter

des anomalies ou des maladies [33].

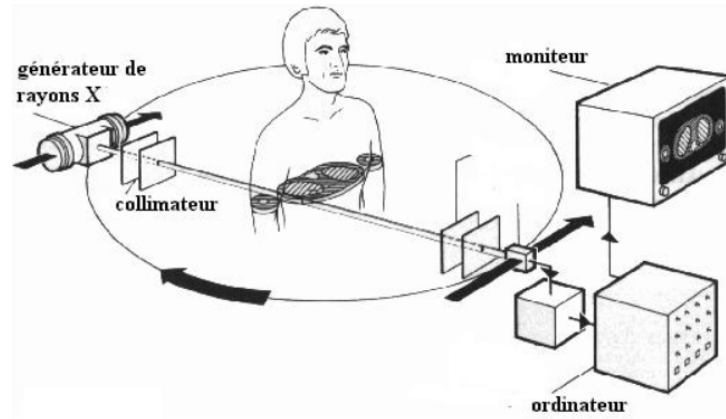


FIGURE 1.5 – Procédure d'acquisition d'une image scanner (extrait de [34]).

La radiographie thoracique est souvent privilégiée par rapport au scanner pour le diagnostic initial des pathologies pulmonaires principalement en raison de sa simplicité, de son coût moindre et de sa rapidité d'exécution. Étant une méthode moins coûteuse et plus accessible dans de nombreux contextes de soins, la radiographie thoracique permet un dépistage efficace et une évaluation préliminaire des affections pulmonaires sans imposer une charge financière ou logistique importante. De plus, la radiographie thoracique expose le patient à une dose de radiation significativement plus faible par rapport au scanner CT, ce qui en fait une option plus sûre pour le suivi répété ou chez des populations sensibles, comme les enfants ou les femmes enceintes. Bien que le CT offre une résolution d'image supérieure et une meilleure visualisation des structures internes, la radiographie thoracique reste un outil diagnostique de première ligne efficace pour détecter de nombreuses anomalies pulmonaires, réservant l'utilisation du CT pour des cas où une analyse plus approfondie est nécessaire suite aux résultats initiaux. Dans notre travail, nous nous concentrons spécifiquement sur le diagnostic des pathologies pulmonaires à partir de la radiographie thoracique.

1.5 Diagnostic des pathologies pulmonaires à partir des CXRs

1.5.1 Diagnostic de la pneumonie

Les signes les plus courants de pneumonie sur une CXR incluent des zones d'opacification, où le tissu pulmonaire normalement aéré est remplacé par des fluides infectieux tels que du pus ou du mucus. Ces zones peuvent varier en taille et en forme. Un bronchogramme aérien, où les bronches remplies d'air deviennent visibles dans la zone consolidée, peut également être présent et est considéré comme un signe diagnostique important de la pneumonie alvéolaire.



FIGURE 1.6 – Visualisation de la pneumonie sur la radiographie thoracique (extrait de [35]).

1.5.2 Diagnostic de l'épanchement pleural

L'épanchement pleural est caractérisé par l'accumulation anormale de liquide dans l'espace pleural. La localisation de ce liquide varie en fonction de la posture du patient. Lorsque le patient se tient debout, le liquide s'accumule au niveau inférieur du thorax, enveloppant la base du poumon et créant un "ménisque" - une courbe concave qui obscurcit l'angle costophrénique et, dans certains cas, une portion ou la totalité de l'hémi-diaphragme. En revanche, si le patient est allongé pendant la prise de la radiographie, l'épanchement pleural a tendance à se répartir le long de la paroi postérieure du thorax, rendant sa détection via une radiographie thoracique plus difficile [36].

Comme démontré dans la Figure 1.7, le diagnostic de l'épanchement pleural révèle les caractéristiques suivantes :

- La zone inférieure gauche est uniformément blanche.
- Au sommet de cette zone blanche se trouve une surface concave - signe du ménisque.
- Le bord gauche du cœur, l'angle costophrénique et l'hémidiaphragme sont masqués.

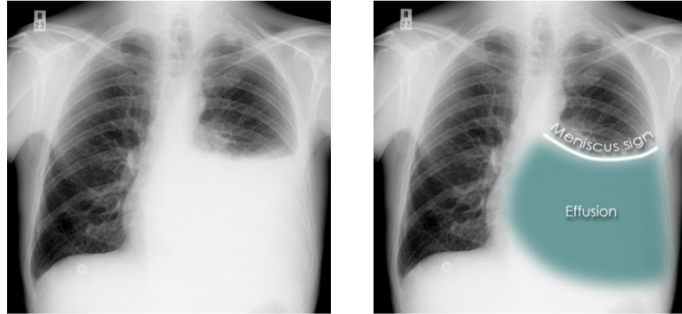


FIGURE 1.7 – Visualisation de l'épanchement pleural sur la radiographie thoracique (extrait de [36]).

1.5.3 Diagnostic de la consolidation

Sur une image radiologique, une consolidation apparaît comme une zone d'opacification (dense et blanche), indiquant la zone où le tissu pulmonaire a perdu son aspect aéré typique (Figure 1.8). Un signe diagnostique important de la consolidation est le bronchogramme aérien, où les bronches remplies d'air deviennent visibles à travers la zone consolidée. Cela se produit parce que les bronches ne sont pas obstruées et apparaissent comme des structures linéaires sombres au sein de l'opacification [37].



FIGURE 1.8 – Visualisation de la consolidation sur la radiographie thoracique (extrait de [38]).

1.6 Conclusion

Dans ce chapitre, nous avons présenté les différentes pathologies qui affectent les poumons, en abordant leurs causes, symptômes, et traitements. Nous avons également introduit les principales techniques d'imagerie pulmonaire utilisées pour le diagnostic des maladies pulmonaires, en mettant particulièrement l'accent sur la radiographie thoracique. Cette dernière, en tant que méthode de diagnostic, a été examinée, soulignant son importance et son application dans l'identification des pathologies pulmonaires.

Cependant, l'examen manuel des images de radiographie thoracique par les radiologues est une tâche complexe et qui demande du temps, en raison de la complexité inhérente des images CXR. La distinction précise entre les différents types de pathologies pulmonaires sur ces images constitue donc un défi majeur pour les professionnels de la santé, d'autant plus que de nombreuses affections pulmonaires partagent des caractéristiques visuelles similaires. Cette situation peut entraîner des erreurs d'interprétation, des diagnostics manquants et des erreurs diagnostiques potentiellement graves. Ainsi, il est impératif de recourir à des méthodes avancées pour améliorer la précision et l'efficacité du processus diagnostique à partir des radiographies thoraciques.

Dans le chapitre suivant, nous présenterons les méthodes de prétraitement des images de radiographie thoracique et les modèles d'apprentissage automatique pour la classification des pathologies pulmonaires à partir de ces images.

BASES DE DONNÉES, PRÉTRAITEMENT DES IMAGES ET CLASSIFICATION DES PATHOLOGIES PULMONAIRES PAR LES MODÈLES D'APPRENTISSAGE PROFOND

2.1 Introduction

Les progrès récents dans les technologies d'Intelligence Artificielle (IA), incluant les algorithmes d'apprentissage automatique (ML) et d'apprentissage profond (DL), ont ouvert des voies prometteuses pour relever les défis associés à l'analyse des radiographies thoraciques dans le diagnostic des maladies pulmonaires. Contrairement aux algorithmes de ML qui nécessitent une extraction manuelle de caractéristiques à partir des images, les algorithmes de DL ont la capacité à extraire automatiquement des caractéristiques pertinentes des images. Ce processus est particulièrement utile en imagerie médicale, où l'interprétation des images peut être subjective et dépendre de l'expertise du radiologue.

Ces dernières années, l'application des techniques de DL aux ensembles de données d'imagerie médicale a démontré une efficacité remarquable. Ces techniques, connues pour leur capacité à acquérir des représentations détaillées, ont prouvé qu'elles surpassaient les professionnels médicaux dans des tâches telles que le diagnostic de la pneumonie [39] et la classification de diverses pathologies pulmonaires [40]. Ces avancées significatives dans la technologie DL offrent un potentiel considérable dans la réduction de la charge de travail des radiologues, l'augmentation de l'efficacité du diagnostic médical, contribuant à une meilleure prise de décision clinique et à l'amélioration des soins aux patients.

Ce chapitre présente le processus d'application d'un algorithme d'apprentissage profond pour la classification des pathologies pulmonaires à partir d'images de radiographie thoracique. Comme le montre la Figure 2.1, le cadre méthodologique de cette application

comporte plusieurs étapes, qui seront détaillées dans les sections ultérieures.

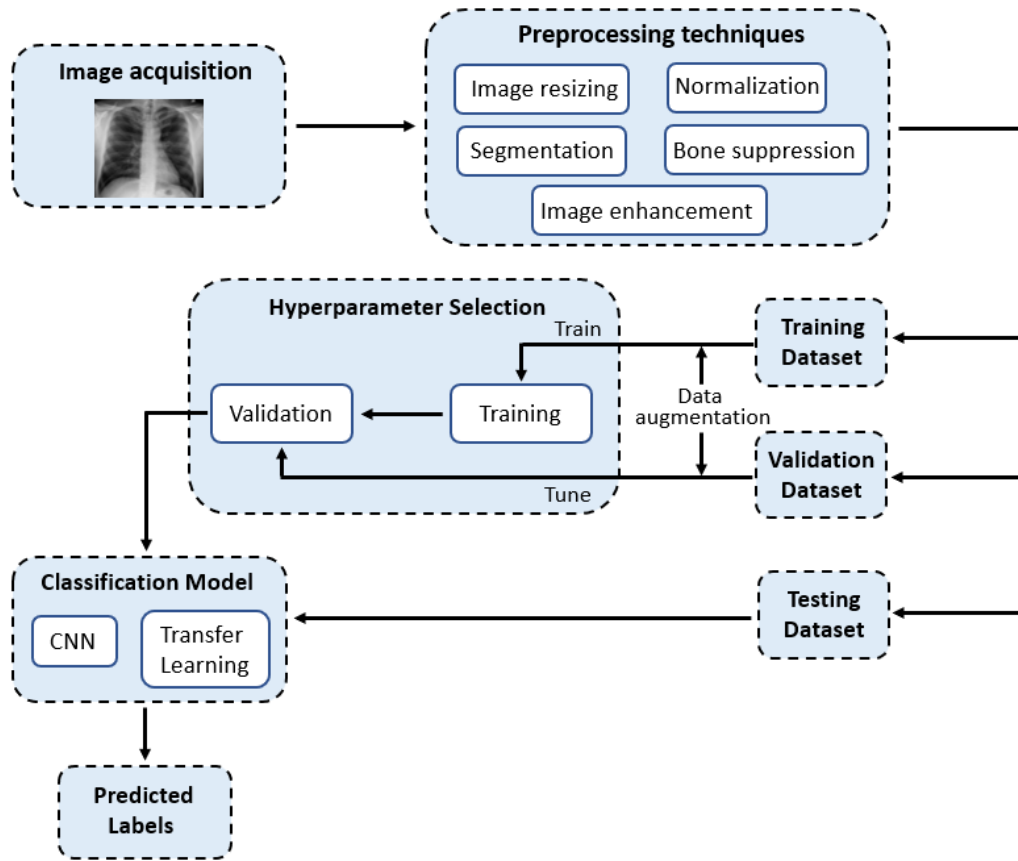


FIGURE 2.1 – Organigramme de l'application de l'apprentissage profond pour la classification des pathologies pulmonaires à partir des images de radiographie thoracique (extrait de [41]).

2.2 Acquisition des images (Bases de données)

L'étiquetage des images médicales nécessite des connaissances et un niveau élevé d'expertise clinique, typiquement celles des radiologues, et exige beaucoup de travail, d'efforts et de temps. Afin de faciliter la recherche sur les données médicales, notamment l'analyse des radiographies thoraciques et le développement de modèles d'IA pour la classification des différentes pathologies pulmonaires à partir de ces images, plusieurs bases de données ont été développées et rendues accessibles au public ainsi qu'aux chercheurs et à la communauté scientifique.

La Table 2.1 présente un résumé des bases de données les plus utilisées dans la litté-

rature. Pour chaque base de données, une description est fournie, soulignant ses caractéristiques principales, les types de pathologies pulmonaires qu'elle comprend, ainsi que le nombre d'images qu'elle contient.

TABLE 2.1 – Comparaison des bases de données.

Base de données	Centre d'acquisition	Type des images	Position de vue	Nombre d'images	Nombre des patients	Classes
Montgomery County X-ray (MC)	Département de la santé et des services sociaux, comté de Montgomery, Maryland, États-Unis	X-ray	Frontale	138	-	Tuberculose et normale
Shenzhen	Hôpital n°3 de Shenzhen, Collège médical du Guangdong, Shenzhen, Chine	X-ray	Frontale	662	-	Tuberculose et normale
Chest X-ray Images (Pneumonia)	Centre médical pour femmes et enfants de Guangzhou	X-ray	Frontale	5 856	-	Pneumonie et normale
Japanese Society of Radiological Technology (JSRT)	13 centres médicaux au Japon et une institution aux États-Unis	X-ray	Frontale	247	-	Nodules pulmonaires et absence de nodules
CheXpert	Centre Médical, Université de Stanford	X-ray	Frontale et latérale	224 316	65 240	12 pathologies thoraciques, dispositifs de support et absence de pathologie
MIMIC-CXR	Centre Médical Beth Israel Deaconess à Boston, MA	X-ray	Frontale et latérale	377 110	65 379	12 pathologies thoraciques, dispositifs de support et absence de pathologie
ChestX-ray14	Hôpitaux affiliés au Centre Clinique des Instituts Nationaux de la Santé	X-ray	Frontale	112 120	30 805	14 pathologies thoraciques et absence de pathologie

1. Montgomery County X-ray dataset (MC) :

La base de données MC [42] a été collectée en collaboration avec le Département de la Santé et des Services Sociaux, comté de Montgomery, Maryland, États-Unis. Cette base de données contient 138 radiographies thoraciques, en vue frontale, issues du programme de dépistage de la tuberculose (TB) du comté de Montgomery. Parmi celles-ci, 80 sont des cas normaux et 58 sont des cas présentant des manifestations de la TB. Les radiographies ont été capturées avec une machine à rayons X stationnaire Eureka (CR) et sont fournies au format PNG en tant qu'images en niveaux de gris de 12 bits. Les images ont une taille soit de 4020×4892 , soit de 4892×4020 pixels. L'ensemble de données comprend également des informations cliniques pour chaque image de radiographie, telles que l'âge et le genre du patient.

2. Shenzhen dataset :

La base de données Shenzhen [42] a été collectée en collaboration avec l'Hôpital n°3 de Shenzhen, Collège médical du Guangdong, Shenzhen, Chine. Les radiographies thoraciques ont été capturées en utilisant un système Philips DR Digital Diagnost. Cette base de données contient 662 radiographies thoraciques frontales, dont 326 sont des cas normaux et 336 sont des cas présentant des manifestations de la TB. Les radiographies sont fournies au format PNG. La taille des images peut varier mais elle est approximativement de 3000×3000 pixels. L'ensemble de données inclut aussi des informations cliniques pour chaque image, telles que l'âge et le genre du patient.

3. Chest X-ray Images (Pneumonia) dataset :

La base de données Chest X-ray Images (Pneumonia) [43] contient 5 856 images de radiographie thoracique en vue frontale (antéro-postérieur). Les images ont été sélectionnées à partir de cohortes rétrospectives de patients pédiatriques âgés de un à cinq ans du Centre médical pour femmes et enfants de Guangzhou. Parmi celles-ci, 4 273 images montrent une pneumonie et 1 583 sont normales. Les images sont fournies au format JPEG.

4. Japanese Society of Radiological Technology dataset (JSRT) :

La base de données JSRT [44] comprend 247 images de radiographie thoracique en vue frontale (postéro-antérieure) provenant de 13 centres médicaux au Japon et d'une institution aux États-Unis. Cette base de données inclut 93 images sans nodules et 154 avec des nodules pulmonaires (100 malins et 54 bénins). Toutes les images ont une taille de 2048×2048 pixels. La base de données fournit des détails cliniques pour chaque patient, y compris la dimension du nodule, l'âge, le genre, le diagnostic final, le degré de subtilité, la position anatomique du nodule, ainsi que les coordonnées x et y marquant le centre du nodule sur l'image numérique (avec le coin supérieur gauche établi comme point d'origine pour les axes x et y). Elle inclut également la classification du nodule, distinguant entre les nodules malins et bénins.

5. CheXpert dataset :

La base de données CheXpert [45] contient 224 316 images de radiographie thoracique, à la fois frontales et latérales, provenant de 65 240 patients. Chaque image est étiquetée pour la présence de 14 observations radiographiques thoraciques courantes, classées comme positives, négatives ou incertaines. Ces observations incluent

atelectasis, cardiomégalie, consolidation, œdème, épanchement pleural, pneumonie, pneumothorax, cardiomédiastin élargi, lésion pulmonaire, opacité pulmonaire, autre affection pleurale, fracture, dispositifs de support et absence de pathologie.

6. MIMIC-CXR dataset :

La base de données MIMIC-CXR [46] comprend 377 110 images de radiographie thoracique en vues frontales et latérales, correspondant à 227 835 études radiographiques effectuées au Centre Médical Beth Israel Deaconess à Boston, MA, pour 65 379 patients. Les images sont disponibles au format DICOM et sont accompagnées de descriptions radiologiques en texte libre. Elles ont été catégorisées en 14 classes : Absence de pathologie, cardiomédiastin élargi, cardiomégalie, opacité alvéolaire, lésion pulmonaire, œdème, consolidation, pneumonie, atélectasie, pneumothorax, épanchement pleural, autres affection pleurale, fracture, dispositifs de support et absence de pathologie.

7. ChestX-ray14 dataset :

La base de données ChestX-ray14 [47] est extrait des bases de données PACS cliniques des hôpitaux affiliés au Centre Clinique des Instituts Nationaux de la Santé (NIH) et comprend 112 120 images de radiographie thoracique en vue frontale de 30 805 patients uniques. Les images sont étiquetées avec 14 pathologies thoraciques courantes (chaque image peut avoir plusieurs étiquettes). L'image est étiquetée comme "no finding" si aucune de ces pathologies n'a été détectée. Les 14 pathologies sont les suivantes : atelectasis, cardiomégalie, consolidation, œdème, épanchement, emphysème, fibrose, hernie, infiltration, masse, nodule, épaissement pleural, pneumonie, pneumothorax. Toutes les images sont au format PNG et ont une taille de 1024×1024 pixels. L'ensemble de données contient également d'autres données telles que le genre et l'âge du patient et la position de vue de la radiographie (c'est-à-dire si les images ont été capturées en position postéro-antérieure (PA) ou antéro-postérieure (AP)).

2.3 Prétraitement des images

Le prétraitement des images constitue une étape cruciale précédant le développement d'un modèle de classification. Les techniques de prétraitement visent à accentuer les caractéristiques de l'image par l'amélioration de l'image et la suppression des effets de distorsion indésirables, ce qui permet d'obtenir une image améliorée plus adéquate pour

le modèle d'apprentissage [48]. Par conséquent, cette étape de prétraitement s'avère indispensable pour accroître l'efficacité des modèles d'apprentissage, et donc les performances de classification.

Dans le contexte de l'analyse des radiographies thoraciques, le prétraitement est effectué pour éliminer le bruit de l'image et l'améliorer de manière à ressortir les informations et les détails pertinents pour une analyse approfondie.

Les techniques typiques de prétraitement d'image sont l'amélioration de l'image, la segmentation de l'image, et la suppression des structures osseuses en vue de la détection d'objets pour des applications spécifiques dans les images CXR. Le redimensionnement et la normalisation des images comptent parmi les méthodes de prétraitement les plus utilisées. Le redimensionnement d'image modifie les dimensions de l'image, dans le but de réduire la taille du fichier image par rapport à l'original. La normalisation de l'image ou la mise à l'échelle des caractéristiques vise à uniformiser les valeurs de données sur une échelle standard.

2.3.1 Amélioration de l'image

Le contraste, la clarté des bords et la présence de bruit dans les images influencent considérablement la classification et l'identification des lésions [49]. Afin d'extraire des détails précis dans les zones sombres et à faible contraste des radiographies thoraciques, il est nécessaire d'améliorer ces images pour accentuer les informations structurelles et éliminer le bruit. L'amélioration des radiographies thoraciques implique donc l'augmentation du contraste, la réduction du bruit et l'accentuation des bords [9, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62]

Plusieurs techniques sont utilisées pour fournir une meilleure qualité des images. Les techniques d'amélioration des images de radiographie thoracique les plus utilisées sont le masque flou, l'égalisation d'histogramme, l'égalisation d'histogramme adaptative à contraste limité et le filtrage à accentuation des hautes fréquences.

1. Masque Flou :

Le masque flou ou *Unsharp Masking* (UM) est une technique d'amélioration des détails de bords qui repose sur le principe de la soustraction pour indirectement affiner les informations à haute fréquence de l'image tels que les textures, les bords, les motifs fins et d'autres détails complexes dans une image. Concrètement, cette méthode consiste à soustraire l'image originale de sa version filtrée par un filtre

passer-bas afin d'obtenir la partie haute fréquence. Ensuite, cette partie haute fréquence est amplifiée par un coefficient, puis ajoutée de nouveau à l'image originale. Par ce processus, les informations sur les détails de bord sont renforcées de manière significative, ce qui aboutit à une amélioration effective de la netteté des bords dans l'image d'origine [63, 64]. La Figure 2.2 montre l'image CXR originale et le résultat de son amélioration à l'aide de la méthode Masque flou.

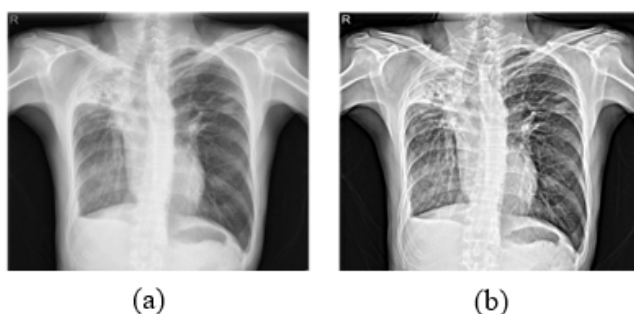


FIGURE 2.2 – (a) CXR originale et (b) son amélioration en utilisant la méthode Unsharp Masking (extrait de [56]).

2. Égalisation d'histogramme :

L'égalisation d'histogramme ou *Histogram Equalization* (HE) est une technique reconnue pour l'amélioration du contraste global dans les images, qui vise à optimiser la visibilité des détails. Cette technique répartit de manière plus uniforme les intensités lumineuses ou les niveaux de gris sur toute la plage disponible, en ajustant la distribution des valeurs des pixels afin qu'elles s'étendent sur toute l'échelle de gris, favorisant ainsi un contraste amélioré [55, 65, 66, 67, 68].

L'avantage d'utiliser la technique d'égalisation d'histogramme dans le traitement préalable des images médicales, avant l'application d'un algorithme d'apprentissage automatique ou profond, est qu'elle fournit une image au contraste amélioré dans laquelle les bords seront plus visibles pour les systèmes automatisés, facilitant l'extraction de caractéristiques distinctives cruciales pour la tâche de classification [69]. Ainsi, cela permet au système de classification de diagnostiquer les images médicales plus rapidement et avec une plus grande précision. La Figure 2.3 montre un exemple de l'image CXR originale et l'image améliorée en utilisant la méthode de l'égalisation d'histogramme, ainsi que leurs histogrammes correspondants. L'histogramme de l'image égalisée montre que les valeurs sont redistribuées

sur tous les pixels par rapport à l'histogramme de l'image originale.

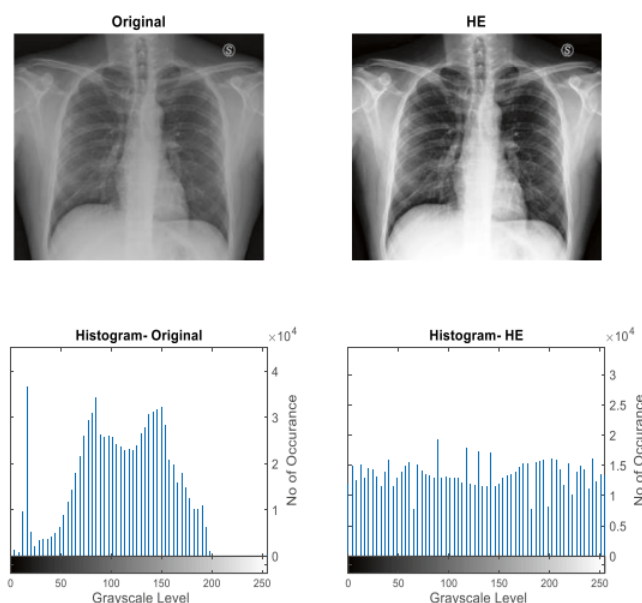


FIGURE 2.3 – CXR originale et son amélioration en utilisant la méthode Histogram Equalization et leurs histogrammes respectifs (extrait de [70]).

3. Égalisation d'histogramme adaptative à contraste limité :

L'égalisation d'histogramme adaptative à contraste limité ou *Contrast Limited Adaptive Histogram Equalization* (CLAHE) est utilisée pour améliorer le contraste des images [9, 56, 58, 60, 62, 70, 71]. Elle diffère de l'égalisation d'histogramme ordinaire par le fait qu'elle divise l'image en petits blocs ou tuiles et applique l'égalisation d'histogramme séparément à chaque bloc, ce qui permet de tenir compte des variations locales du contraste. Cependant, pour éviter une amplification excessive du contraste (qui pourrait rendre le bruit plus visible ou rendre l'image non naturelle), un seuil est défini pour limiter le contraste. Si des bins de l'histogramme dépassent ce seuil, les pixels excédentaires sont redistribués uniformément sur tous les bins de l'histogramme. Cette redistribution a pour effet de lisser les pics trop élevés dans la distribution des intensités, évitant ainsi de renforcer le contraste dans les zones homogènes. Ainsi l'histogramme redistribué est différent d'un histogramme ordinaire car l'intensité de chaque pixel est limitée à un maximum sélectionné [72].

La taille de la tuile et la valeur limite du clip sont des hyper-paramètres critiques pour cette méthode [65]. La taille des tuiles correspond à la taille des petites

régions dans lesquelles l'image est divisée pour appliquer l'égalisation adaptative de l'histogramme. Chaque tuile agit comme une région locale où l'égalisation de l'histogramme est effectuée. La taille des tuiles détermine donc l'étendue de l'amélioration locale du contraste. La limite du clip est une valeur seuil qui limite le degré d'égalisation de l'histogramme dans chaque région locale (tuile). Elle empêche l'amplification extrême du contraste dans les régions où la distribution des valeurs de pixels est uniforme. Si le processus d'égalisation de l'histogramme augmente les valeurs des pixels au-delà de cette limite, celles-ci sont réduites pour garantir que l'amélioration reste contrôlée et n'entraîne pas de résultats extrêmes [73]. Après l'égalisation des histogrammes de chaque bloc, les bords entre les blocs peuvent présenter des discontinuités. Pour éviter cet effet de bordure, CLAHE utilise une interpolation bilinéaire entre les blocs adjacents pour lisser la transition.

La Figure 2.4 présente le résultat de la technique CLAHE. L'image CLAHE présente un histogramme en forme de cloche, tandis que l'image HE (Figure 2.3) présente un histogramme avec une distribution uniforme. Contrairement à l'image résultante de l'HE (Figure 2.3) où certaines régions présentent une forte différence de luminosité, l'image CLAHE (Figure 2.4) présente une transition douce des intensités pour les pixels adjacents.

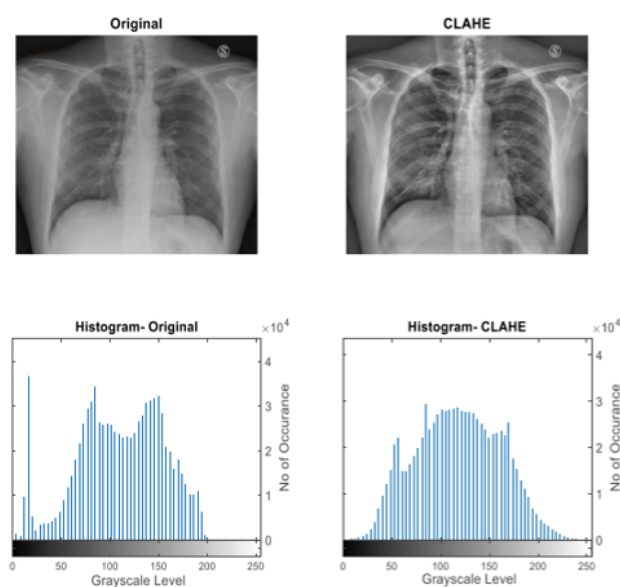


FIGURE 2.4 – CXR originale et son amélioration en utilisant la méthode Contrast Limited Adaptive Histogram Equalization et leurs histogrammes respectifs (extrait de [70]).

4. Filtrage à accentuation des hautes fréquences :

Le filtrage à accentuation des hautes fréquences ou *High-Frequency Emphasis Filtering* (HEF) est une technique qui vise à accentuer les détails fins et les textures d'une image en augmentant la contribution des composantes à haute fréquence.

La Figure 2.5 illustre le processus de la méthode HEF. Tout d'abord, l'image originale est transformée pour séparer les différentes fréquences. Cela est souvent réalisé à l'aide de la Transformée de Fourier, qui décompose l'image en ses composantes fréquentielles. Dans ce domaine fréquentiel, les composantes sont classées en basses fréquences (associées aux variations d'intensité lentes, ou aux grandes structures de l'image) et en hautes fréquences (associées aux changements rapides d'intensité, ou aux détails fins et aux textures). L'étape de filtrage de l'algorithme consiste à appliquer le filtre gaussien passe-haut pour accentuer les bords. Les bords ont tendance à s'exprimer dans le spectre des hautes fréquences car ils présentent des changements d'intensité plus importants. Enfin, l'image est retransformée dans son domaine spatial d'origine, et ainsi les détails fins sont plus visibles [56].

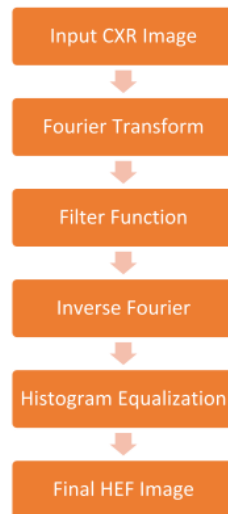


FIGURE 2.5 – Le processus de la méthode High-Frequency Emphasis Filtering (extrait de [56]).

Cette technique produit une image à faible contraste et peu lumineuse et qui a un histogramme de niveau d'intensité à faible plage dynamique dû au fait que la largeur de l'histogramme est étroite par rapport à l'échelle de gris de l'image. L'utilisation de l'égalisation d'histogramme est nécessaire pour augmenter la plage dynamique de l'histogramme, ce qui permet d'avoir un niveau d'intensité unifor-

mément réparti et ainsi des images avec un contraste plus élevé [74].

La Figure 2.6 montre un exemple de l'image CXR originale, le résultat HFF, le résultat HFEF (avec l'égalisation d'histogramme) et de leurs histogrammes. La Figure 2.6 (a) illustre une radiographie thoracique originale qui apparaît légèrement floue. Son histogramme associé, visible dans la Figure 2.6 (b), montre une distribution étroite des intensités lumineuses avec des extrémités où les fréquences sont nulles, indiquant une limitation dans la dynamique des niveaux de gris. L'application de HEF à la Figure 2.6 (c) aboutit à une image plus sombre que l'originale, rendant les détails de la radiographie thoracique moins évidents, mais la tonalité des niveaux de gris due à la basse fréquence est préservée. Son histogramme correspondant, présenté dans la Figure 2.6 (d), confirme toujours la présence d'une gamme étroite des niveaux d'intensité de gris. L'application de l'égalisation d'histogramme, comme le montre la Figure 2.6 (e), produit une image à contraste amélioré où la structure et les détails de la radiographie thoracique deviennent nettement plus visibles et plus clairs. Son histogramme à la Figure 2.6 (f) montre une extension significative de la plage des niveaux d'intensité de gris.

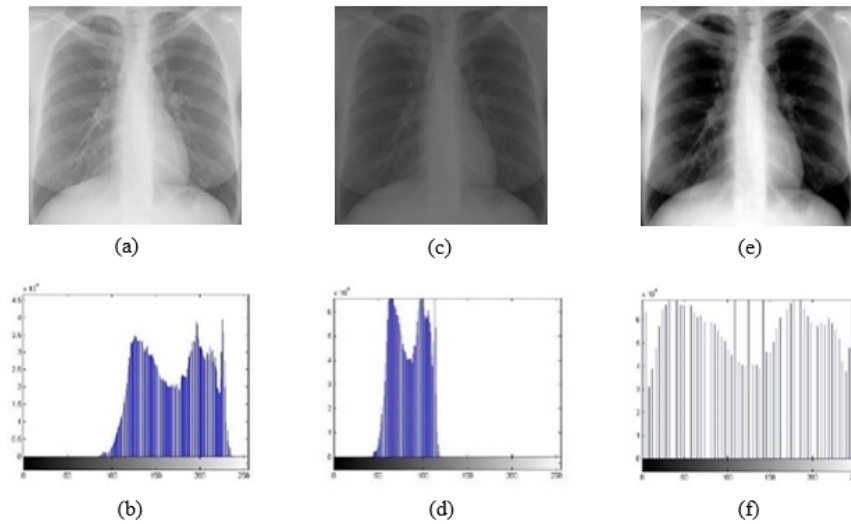


FIGURE 2.6 – (a) CXR originale, (c) le résultat de la méthode High-Frequency Emphasis Filtering, (e) le résultat de la méthode HEF et l'égalisation d'histogramme, et (b), (d) et (f) leurs histogrammes respectifs (extrait de [74]).

2.3.2 Segmentation sémantique de l'image

La segmentation d'images consiste à diviser une image en plusieurs segments ou régions, chacun identifiant un objet ou une partie différente de l'image.

Particulièrement dans l'analyse des radiographies thoraciques, la segmentation du champ pulmonaire est une étape importante. Elle permet de définir avec précision les régions d'intérêt (ROI) du champ pulmonaire dans l'image, facilitant ainsi l'identification de signes radiologiques spécifiques, tels que les opacités pulmonaires, la consolidation et les nodules. Cette séparation des images de leur arrière-plan augmente la précision et l'exactitude de l'analyse radiologique. La Figure 2.7 montre l'image CXR originale, le masque produit par le modèle U-Net et l'image pulmonaire segmentée finale.

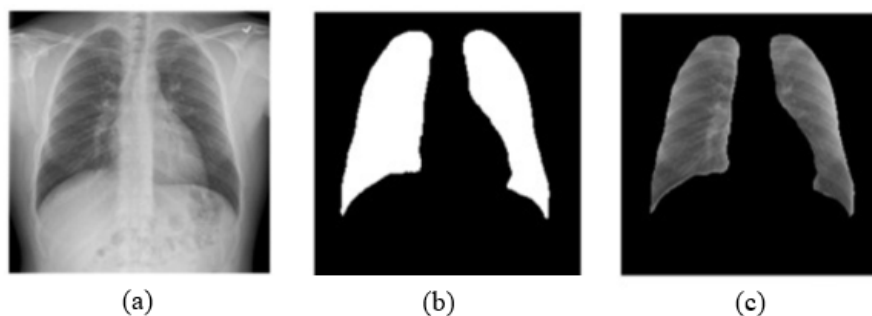


FIGURE 2.7 – Résultats de la segmentation : (a) CXR originale, (b) masque généré par U-Net et (c) régions pulmonaires segmentées (extrait de [75]).

La segmentation des radiographies thoraciques a fait l'objet de nombreuses recherches [58, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 75, 90, 91]. La segmentation des images est souvent effectuée en utilisant des réseaux de neurones convolutifs de type U-Net, entraînés sur de grands ensembles de données d'images étiquetées. Narayanan et al. [85] ont employé la segmentation d'image pour déterminer si une pneumonie diagnostiquée résultait d'agents bactériens ou viraux. L'intégration de la segmentation a permis d'améliorer les performances du modèle de classification de 3%. Rahman et al. [86] ont montré que la classification à l'aide d'images pulmonaires segmentées était plus performante que la classification à l'aide d'images radiographiques entières.

2.3.3 Suppression des structures osseuses

La suppression des structures osseuses consiste à éliminer les structures squelettiques, en particulier les côtes et la clavicule, des images radiographiques, comme le montre la Figure 2.8. Ces structures osseuses peuvent masquer des anomalies pulmonaires, rendant plus difficile la phase d'extraction des caractéristiques d'un système de diagnostic assisté par ordinateur (CAD).

La suppression des structures osseuses permet d'augmenter la visibilité de la densité des tissus mous, ce qui peut également améliorer la capacité du système à détecter les nodules pulmonaires et d'autres anomalies pulmonaires [92, 93, 94, 95, 96, 97, 98]. Yang et al. [94] ont présenté une méthode de suppression de la structure osseuse dans les CXRs en utilisant des réseaux de neurones convolutifs profonds, et ont constaté que cette méthode produisait des images de haute résolution et de haute qualité des tissus mous. Gordienko et al. [95] ont démontré l'efficacité de la technique de suppression des structures osseuses dans la détection du cancer du poumon à l'aide d'une approche d'apprentissage profond, atteignant une meilleure précision avec l'ensemble de données prétraité sans les structures osseuses.

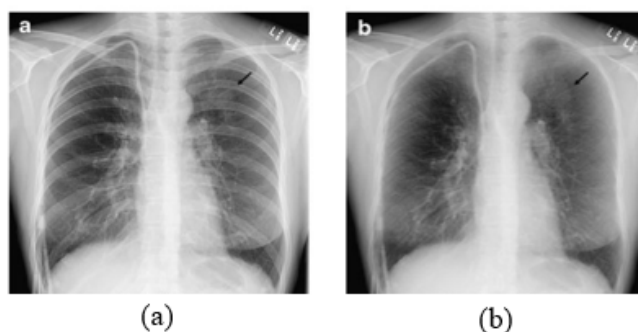


FIGURE 2.8 – (a) CXR originale, et (b) CXR après suppression des structures osseuses montrant plus clairement l'opacité de l'espace aérien (extrait de [92]).

2.4 Modèles d'Intelligence Artificielle

L'intelligence artificielle ou *Artificial Intelligence* (AI) est une branche de l'informatique qui vise à donner aux machines la capacité de raisonner comme un être humain, et qui implique le développement des tâches normalement associées à l'intelligence humaine, telles que la reconnaissance vocale, la reconnaissance d'images, la prise de décisions, etc.

L'apprentissage automatique ou apprentissage machine, *Machine Learning* (ML), est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'apprendre à partir des données et de s'améliorer grâce à un processus d'apprentissage. De nombreux modèles d'apprentissage machine ont été développés pour résoudre des tâches variées telles que la régression, la classification, la détection, la prédiction, etc. Il existe deux principaux types d'algorithmes de machine learning : l'apprentissage supervisé et l'apprentissage non supervisé. L'apprentissage supervisé implique l'entraînement d'un modèle sur un ensemble de données étiquetées, où chaque exemple d'entraînement comprend des entrées et la sortie souhaitée (étiquette). Le modèle apprend à prédire la sortie à partir des entrées. Au contraire, l'apprentissage non supervisé traite des données non étiquetées, cherchant à identifier des structures ou des motifs cachés dans les données. Il ne repose pas sur des sorties connues mais tente de regrouper les données en fonction de leurs similarités ou de trouver des distributions sous-jacentes dans les données.

La disponibilité de grandes quantités de données et l'amélioration significative des puissances de calcul ont contribué au développement de l'apprentissage profond [99].

L'apprentissage profond ou *Deep Learning* (DL) est un sous-domaine de l'apprentissage machine basé sur des réseaux de neurones, inspiré par la structure et le fonctionnement du cerveau humain [100]. Le processus d'apprentissage est qualifié de profond en raison de la configuration de son architecture, qui intègre de multiples couches : des couches d'entrée et de sortie, ainsi que plusieurs couches cachées intermédiaires. Chaque couche est composée d'unités neuronales qui traitent les données entrantes et génèrent des informations utilisables par la couche suivante pour une tâche spécifique. Cette structure multicouche permet à la machine d'apprendre de façon autonome, en extrayant et en traitant les informations contenues dans les données sans intervention humaine directe. Les modèles basés sur l'apprentissage profond se distinguent par leur capacité à apprendre automatiquement et à extraire des caractéristiques pertinentes directement à partir des données. Cette capacité d'apprentissage en profondeur élimine le besoin d'extraire manuellement les caractéristiques, étape traditionnellement cruciale dans les méthodes classiques d'apprentissage automatique.

Ces dernières années, les algorithmes d'apprentissage profond sont de plus en plus utilisés dans la détection et le diagnostic de pathologies à travers l'analyse d'images médicales, notamment pour les infections respiratoires [101, 102]. L'objectif de l'apprentissage profond dans l'analyse des images médicales est de faciliter l'extraction des informations

de manière efficace, de stocker ces connaissances et de les réutiliser afin de renforcer la précision du diagnostic clinique [103]. Le succès des méthodes d'apprentissage profond dans le domaine médical est due à leur capacité à traiter de grandes quantités de données, à automatiser les processus et à accroître l'efficacité des systèmes de soins. Dans les sections suivantes, nous décrivons les architectures d'apprentissage profond les plus couramment utilisées pour la classification des pathologies pulmonaires et pour les tâches de segmentation.

2.4.1 Réseaux de neurones convolutifs

Les réseaux de neurones convolutifs ou *Convolutional Neural Networks* (CNN) sont une catégorie de réseaux de neurones d'apprentissage profond spécifiquement conçus pour traiter des données avec des topologies en grille, telles que les images [104]. Les CNNs portent leur nom de l'opération mathématique de convolution qu'ils utilisent pour traiter les données. En mathématiques, la convolution est une opération effectuée sur deux fonctions pour produire une troisième fonction. Cette dernière décrit la manière dont la forme de l'une est modifiée par l'autre. Le terme convolution désigne à la fois la fonction résultat et son mécanisme de calcul [105].

Les CNNs, grâce à leur architecture, assurent à la fois l'extraction des caractéristiques et la classification, comme le montre la Figure 2.9 qui représente une architecture typique de CNN.

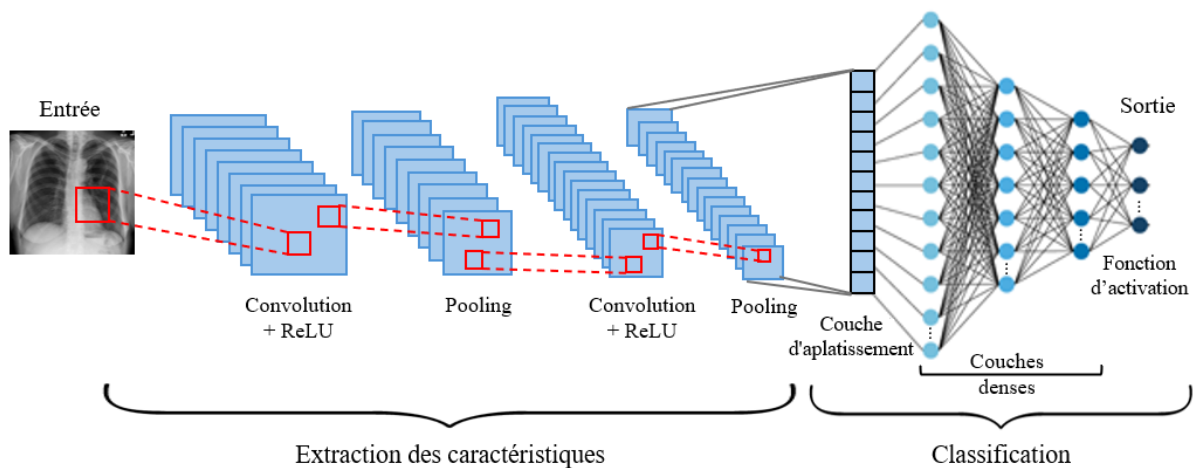


FIGURE 2.9 – Architecture d'un réseau de neurones convolutifs.

Le CNN comporte deux parties distinctes :

- Une partie convolutive : Son objectif est d'extraire des caractéristiques de chaque image d'entrée en les compressant de façon à réduire leur taille initiale. L'image fournie en entrée passe à travers des couches de convolution générant des cartes de caractéristiques qui sont par la suite concaténées en un vecteur de caractéristiques.
- Une partie classification : Elle utilise le vecteur de caractéristiques produit par la partie convolutive comme entrée. Cette partie est composée de couches entièrement connectées, appelées perceptron multicouche ou *Multi Layers Perceptron* (MLP). Son objectif est de combiner les caractéristiques extraites par la partie convolutive pour classer l'image.

Le CNN est composé de trois principaux types de couches : couche de convolution, couche de pooling et couche entièrement connectée.

1. Couches de convolution :

Les couches de convolution sont la partie centrale du CNN où certains calculs sont effectués. Chaque image fournie en entrée passe dans des couches de convolution qui lui appliquent une succession de filtres afin d'extraire des caractéristiques. Un filtre est constitué d'une série de noyaux de convolution (également appelés kernel) qui sont de petites matrices bidimensionnelles contenant des poids. Le filtre est glissé sur toute l'étendue de l'image, calculant à chaque position un produit scalaire entre les valeurs du filtre et les valeurs des pixels de l'image. Ainsi, pour une image $2D$ notée I et un filtre $2D$ noté K , l'opérateur de convolution, représenté par l'équation 2.1 [99], s'applique en tout point (i, j) de l'image pour produire une sortie appelée carte de caractéristiques, comme présenté dans la Figure 2.10.

$$(I * K)(i, j) = \sum_{m=1}^M \sum_{n=1}^N I(m, n) \cdot K(i - m, j - n) \quad (2.1)$$

Généralement, les premières couches de convolution sont dédiées à l'extraction des caractéristiques de bas niveau, notamment les couleurs, les bords, les orientations de gradient, etc. Avec l'ajout de couches de convolution successives, la richesse des informations extraites de l'image s'accroît, permettant de saisir des caractéristiques de niveau plus élevé et ainsi de fournir une représentation plus détaillée et complète de l'image.

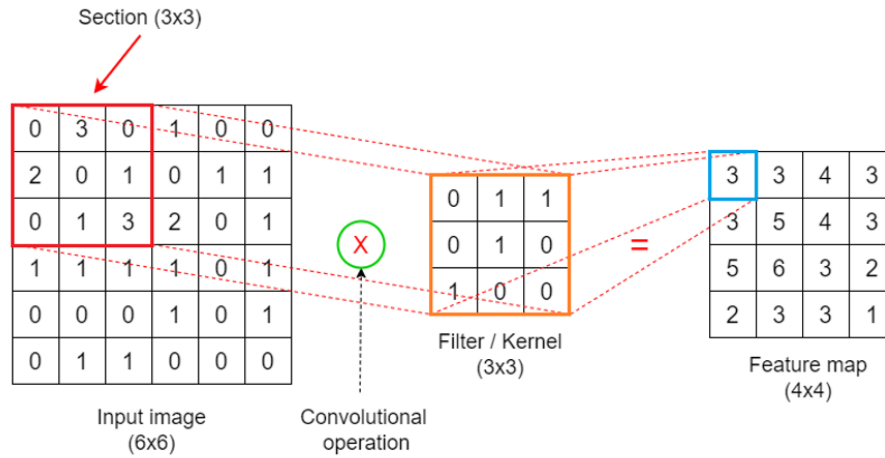


FIGURE 2.10 – Principe de l'opérateur de convolution (extrait de [106]).

Dans les couches de convolution, il est essentiel de prendre en compte deux hyperparamètres qui influencent directement la taille de la carte de caractéristiques.

- Le "padding" consiste à ajouter des valeurs spécifiques (généralement des zéros) autour des bords de l'image d'entrée, permettant ainsi au filtre de glisser sur tous les pixels originaux. Ainsi, la dimension de sortie (celle de la carte des caractéristiques) après convolution reste identique à celle de l'image d'entrée, comme illustré dans la Figure 2.11 (a). À l'inverse, et comme le montre la Figure 2.11 (b), l'absence de padding entraîne une réduction de la dimension de l'image ou de la carte de caractéristiques résultante. En effet, sans padding, les bords de l'image ne sont pas totalement couverts par le filtre durant sa convolution à travers l'image. Par conséquent, la convolution s'effectue uniquement dans les zones où le filtre peut se superposer entièrement à l'image.
- Le "stride" désigne le nombre de pixels sur lesquels le filtre (ou le noyau) se déplace à travers l'image ou la carte de caractéristiques lors de chaque opération de convolution. Comme illustré dans la Figure 2.12, l'utilisation d'un stride de 1 implique que le filtre se déplace d'un pixel à chaque étape, permettant au filtre de couvrir toutes les positions possibles sur l'image. Un stride plus élevé, tel que 2 ou supérieur, entraîne le saut de pixels entre chaque déplacement du filtre. Ceci a pour effet de diminuer le nombre de positions où le filtre est appliqué et, ainsi, de réduire la dimension de la carte de caractéristiques résultante.

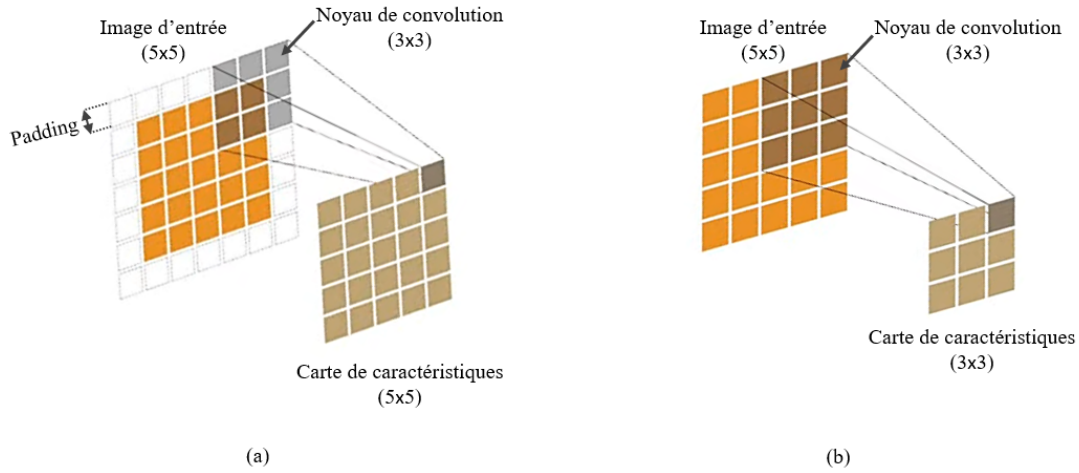


FIGURE 2.11 – L'impact du padding sur la dimension de la carte de caractéristiques. (a) convolution avec padding, et (b) convolution sans padding.

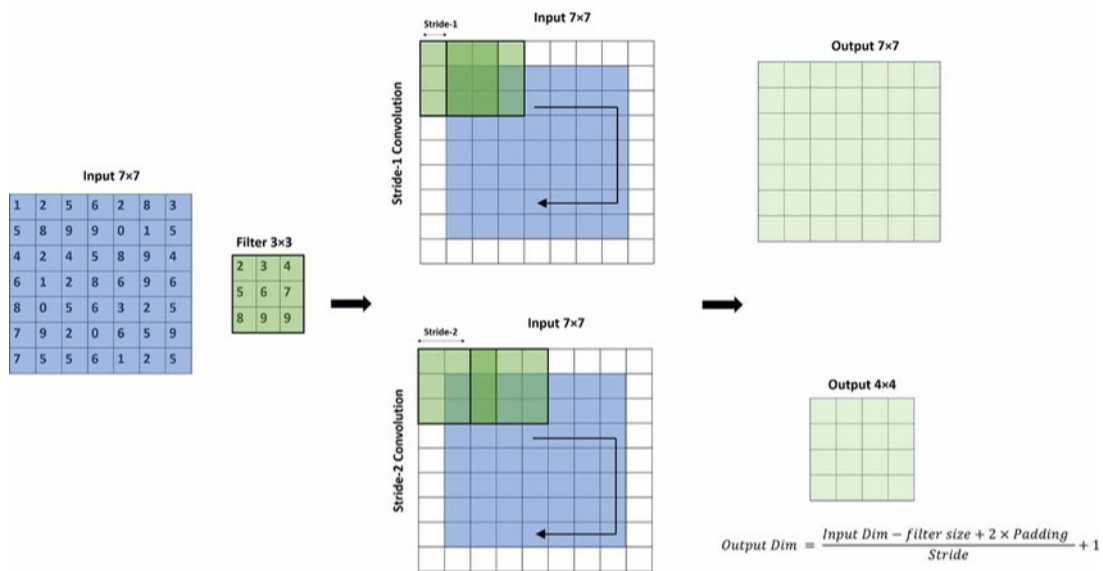


FIGURE 2.12 – L'effet du stride sur la convolution et la dimension de la carte de caractéristiques. (extrait de [107]).

2. Couches de pooling :

Les couches de pooling, placées entre les couches de convolution, réduisent progressivement la dimension spatiale (la hauteur et la largeur) des cartes de caractéristiques entrantes, tout en conservant les informations les plus importantes. Cette réduction de la taille des cartes de caractéristiques permet de diminuer le nombre de paramètres et de calculs dans le réseau. Comme la couche convolutive, la couche de pooling fait glisser un filtre, appelé fenêtre de pooling, sur la carte de caractéristiques avec un certain stride. La principale différence est que le filtre utilisé dans la couche de pooling ne contient pas de poids.

Il existe deux principaux types de pooling (Figure 2.13) :

- Max pooling : Cette technique sélectionne la valeur maximale parmi les pixels couverts par la fenêtre de pooling, ce qui la rend particulièrement efficace pour mettre en évidence les caractéristiques les plus importantes.
- Average pooling : Cette technique calcule la valeur moyenne des pixels dans la fenêtre de pooling, ayant pour effet potentiel d'atténuer les caractéristiques importantes par rapport au max pooling.

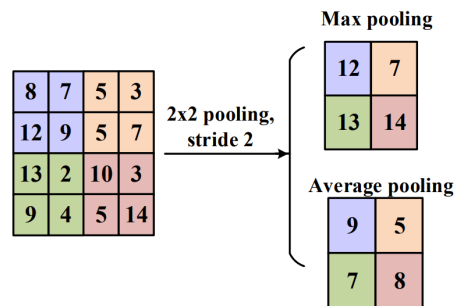


FIGURE 2.13 – Types de pooling. (extrait de [108]).

3. Couche entièrement connectée :

Dans la couche entièrement connectée, chaque neurone de la couche de sortie est entièrement connecté à tous les neurones de la couche précédente. Ces couches interviennent après que les couches convolutives et de pooling ont extrait et condensé les informations spatiales des données d'entrée. Le rôle des couches entièrement connectées est de combiner ces caractéristiques puis y appliquer une fonction d'activation afin de classer l'image d'entrée. Le résultat est transmis à la couche de sortie sous forme d'un vecteur dont la taille est égale au nombre de classes ciblées,

où chaque composante du vecteur indique le score que l'image appartienne à une classe spécifique.

Les fonctions d'activation sont des composants cruciaux dans les modèles de réseaux de neurones, y compris les CNNs. Elles déterminent la sortie d'un neurone en appliquant une transformation non linéaire à sa somme pondérée d'entrée. Les fonctions d'activation introduisent des non-linéarités dans le modèle, ce qui est essentiel pour permettre au réseau de neurones d'apprendre des relations complexes entre les entrées et les sorties.

Plusieurs fonctions d'activation sont couramment utilisées dans les réseaux de neurones, comme illustré dans la Figure 2.14.

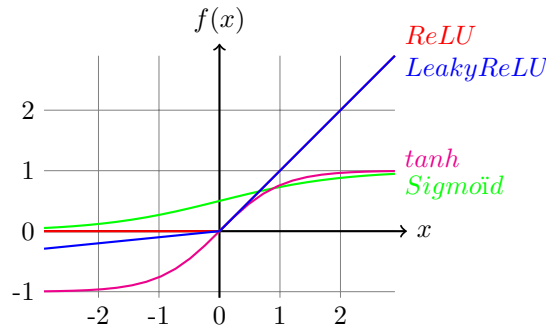


FIGURE 2.14 – Fonctions d'activation.

- La fonction Sigmoide, qui convertit les valeurs d'entrée en une sortie allant de 0 à 1, est typiquement employée dans les dernières couches des modèles de classification. Cependant, elle est moins utilisée dans les couches cachées en raison du problème de disparition des gradients, où les gradients deviennent très petits, empêchant l'ajustement efficace des poids pendant l'entraînement. Elle est définie par l'équation 2.2 :

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.2)$$

- La fonction tangente hyperbolique, souvent abrégée tanh, fournit une sortie comprise entre -1 et 1, et est définie par l'équation 2.3 :

$$f(x) = \tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (2.3)$$

- La fonction ReLU est largement utilisée en raison de sa simplicité et de son efficacité. Elle permet une activation rapide des neurones avec des valeurs positives,

tout en annulant les valeurs négatives. Elle est donnée par l'équation 2.4 :

$$f(x) = \max(0, x) \quad (2.4)$$

- La fonction Leaky ReLU, une variation de ReLU, introduit un petit gradient même pour les valeurs négatives, ce qui permet la mise à jour des poids du réseau. Elle est définie par l'équation 2.5 :

$$y = \begin{cases} x & \text{si } x \geq 0 \\ \alpha x & \text{sinon} \end{cases} \quad (2.5)$$

où α est un petit coefficient positif qui permet une activation légère pour les entrées négatives, facilitant ainsi le flux de gradient dans le réseau même pour des valeurs négatives de x .

- La fonction Softmax est utilisée principalement dans la couche de sortie pour les problèmes de classification multiclass car elle convertit efficacement les scores de chaque classe en probabilités. Sa capacité à fournir une distribution probabiliste sur plusieurs classes en fait le choix idéal pour les tâches de classification où chaque instance doit être classée dans une seule catégorie. Elle est donnée par l'équation 2.6 :

$$\sigma(y_i) = \left(\frac{e^{y_i}}{\sum_j e^{y_j}} \right) j = 1, \dots, n \quad (2.6)$$

où n représente le nombre total de classes possibles dans le problème de classification, et y_i représente les scores pour chaque classe.

2.4.2 Architecture U-Net pour la segmentation sémantique d'images

U-Net est une architecture de réseau de neurones convolutifs (CNN) [109] largement utilisée dans les tâches de segmentation sémantique, en particulier dans le domaine biomédical pour la segmentation d'images médicales [110].

Comme le montre la Figure 2.15, U-Net se distingue par sa structure en « U », composée de deux parties principales : l'encodeur (sur la partie gauche) et le décodeur (sur la partie droite). L'encodeur se compose de plusieurs couches de convolution et de couches de max pooling qui réduisent progressivement la résolution spatiale de l'image en extrayant des caractéristiques à différentes échelles de l'image. Cette partie de l'architecture fonc-

tionne comme un extracteur de caractéristiques. Le décodeur, quant à lui, est responsable de la reconstruction de l'image segmentée à partir des caractéristiques extraites par l'encodeur. Il est constitué de couches de convolution transposée (upconvolution) et de couches de concaténation qui permettent de restaurer progressivement la résolution spatiale de l'image. Cette partie de l'architecture agit comme un générateur d'image segmentée.

Un aspect fondamental de l'architecture U-Net est l'intégration des connexions de saut (skip connections) entre les couches de l'encodeur et du décodeur. Ces connexions permettent de transmettre des informations à des résolutions spatiales supérieures, en permettant au décodeur d'accéder à des caractéristiques de bas niveau provenant de l'encodeur. Cela aide à conserver les informations contextuelles nécessaires à la précision de la segmentation.

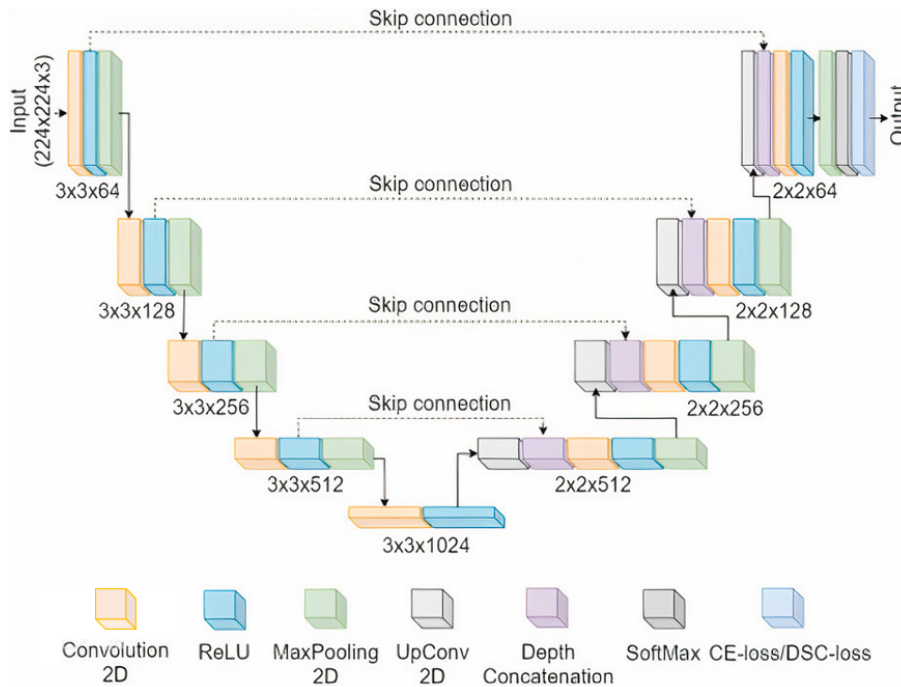


FIGURE 2.15 – Architecture U-Net (extrait de [75]).

2.4.3 Architectures d'attention

Les mécanismes d'attention ont suscité un grand intérêt dans le domaine de l'apprentissage profond en raison de leur capacité à permettre aux modèles de se concentrer sur des parties spécifiques des données d'entrée, ce qui améliore significativement leur

performance. Le modèle Transformer, architecture composée de couches d'attention, a initialement révolutionné le domaine du traitement du langage naturel (NLP) par sa capacité à effectuer des analyses fines et contextuelles des séquences de mots [111]. Cette approche a ensuite été étendue au domaine de la vision par ordinateur avec les Vision Transformers (ViT), qui utilisent des mécanismes d'attention pour analyser les images [112, 113].

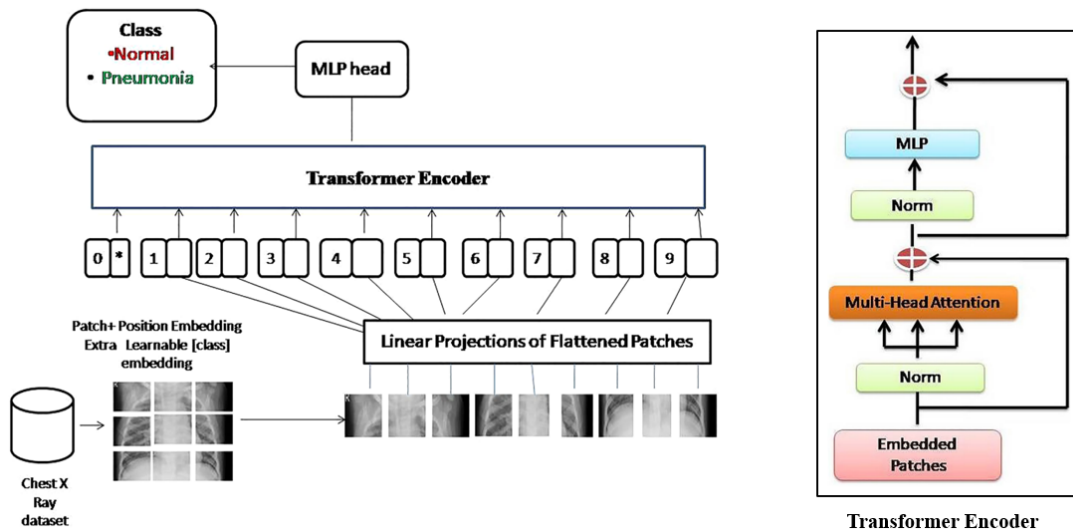


FIGURE 2.16 – Exemple d'architecture du Vision Transformer (extrait de [114]).

Comme le montre la Figure 2.16, ViT divise une image en plusieurs patches qui sont ensuite linéarisés et projetés dans un espace d'embeddings afin de saisir les relations entre les caractéristiques des images. Pour préserver l'information spatiale inhérente à l'image, un token de position est ajouté à chaque patch. La séquence résultante de vecteurs d'embeddings est alors introduite dans un encodeur transformer qui utilise des couches d'attention et des réseaux de neurones. Le mécanisme d'attention permet au modèle de pondérer différemment les parties de l'image, facilitant ainsi la capture des relations spatiales globales entre les différentes parties de l'image.

En raison de la capacité supérieure des Transformers à capturer des relations globales de l'image surpassant celle des CNNs, l'intégration des couches d'attention avec les couches convolutives est souvent adoptée pour bénéficier des avantages de chaque approche [115]. Par conséquent, les architectures qui intègrent des couches d'attention entre les opérations de convolution obtiennent de bonnes performances dans l'analyse d'images [116]. Dans

la tâche de classification des pathologies pulmonaires à partir d'images de radiographie thoracique, l'application des ViT et des mécanismes d'attention facilite l'identification automatique de motifs et d'anomalies présents dans les radiographies thoraciques [114, 117, 118].

2.5 Techniques face à la limitation des données

Les modèles d'apprentissage profond nécessitent une grande quantité de données pour l'entraînement afin d'éviter le risque de surapprentissage et de garantir une précision optimale. Dans le domaine médical, l'un des défis majeurs réside dans l'acquisition de larges bases de données annotées de manière fiable et précise. Pour remédier à cette contrainte, l'augmentation des données et l'apprentissage par transfert sont les techniques les plus utilisées.

1. L'augmentation des données

L'augmentation des données est une technique courante utilisée pour augmenter la taille de l'ensemble de données d'entraînement en créant des variations des données existantes [61, 119, 120, 121, 122]. Cette méthode est souvent utilisée dans l'apprentissage profond pour rendre le modèle plus robuste, en l'entraînant à identifier les motifs ou caractéristiques importantes sous des conditions variées, diminuant ainsi le risque de surajustement et améliorant sa capacité à se généraliser.

Dans le contexte de la classification des pathologies pulmonaires à partir d'images de radiographie thoracique, l'augmentation des données peut être utilisée pour accroître la diversité des images dans l'ensemble de données, ce qui peut aider le modèle à mieux se généraliser à de nouvelles images.

Il existe plusieurs méthodes pour effectuer l'augmentation des données sur les images de radiographie thoracique. Par exemple, une approche consiste à appliquer des transformations géométriques comme la rotation, le redimensionnement et la translation aux images [123, 124, 125, 126, 127, 128, 129]. Cela peut aider le modèle à identifier les pathologies pulmonaires quelle que soit l'orientation ou la taille de l'image. Par exemple, Brunese et al. ont utilisé l'augmentation des données en appliquant une rotation aléatoire des images jusqu'à 15 degrés, dans le sens horaire ou antihoraire pour garantir la généralisation des modèles proposés [130]. Une autre approche consiste à appliquer des transformations basées sur l'image, telles que la modification du contraste, l'application d'un bruit ou d'un flou [131]. Cela peut

aider le modèle à identifier les pathologies pulmonaires dans des images potentiellement altérées par des conditions d'éclairage variables ou des artefacts d'imagerie. Sirazitdinov et al. [131] ont appliqué plusieurs techniques d'augmentation des données pour améliorer le processus d'entraînement de leur réseau, incluant le retournement vertical et horizontal, la rotation aléatoire, l'ajustement aléatoire de la luminosité, les transformations gamma, l'introduction de bruit gaussien aléatoire et l'application de flou.

Au cours des dernières années, la génération d'images CXR artificielles à travers l'utilisation de réseaux antagonistes génératifs ou Generative Adversarial Networks (GANs) a suscité un grand intérêt dans la communauté scientifique. En 2014, Goodfellow et al. ont introduit les GANs comme moyen pour augmenter la taille des ensembles de données existants en générant de nouvelles images artificielles [132]. Les GANs constituent un modèle dans lequel l'objectif est d'apprendre à générer des objets si réalistes qu'ils sont difficiles à différencier des objets réels existants. Autrement dit, les GANs capturent la distribution des données d'apprentissage pour ensuite générer de nouveaux exemples à partir de la même distribution.

Comme illustré dans la Figure 2.17, les GANs se composent de deux réseaux de neurones distincts : le générateur et le discriminateur. Le rôle du générateur est de créer de nouvelles données qui ressemblent le plus possible aux données réelles présentes dans l'ensemble de données. Alors que le discriminateur a pour tâche de distinguer entre les données réelles et les données générées par le générateur. Son objectif est de s'améliorer dans l'identification précise des faux exemples produits par le générateur. Au début du processus de génération, le générateur reçoit un vecteur de bruit aléatoire, généralement issu d'une distribution normale (gaussienne). Le générateur, qui est un réseau de neurones, prend ce vecteur de bruit comme entrée et le transforme en données qui ont la même structure des données réelles que le modèle cherche à imiter (par exemple une image), en effectuant des transformations complexes via ses multiples couches durant l'entraînement. Au début, les données générées à partir du bruit aléatoire ne ressemblent probablement pas beaucoup aux données réelles. Cependant, à travers le processus d'entraînement des GANs qui est un processus dynamique où le générateur et le discriminateur s'améliorent l'un par rapport à l'autre, le générateur apprend à produire des données de plus en plus réalistes, tandis que le discriminateur s'efforce de devenir meilleur à les identifier comme fausses. L'évaluation des données fournies par le discrimina-

teur guide le générateur pour améliorer sa capacité à produire des données réalistes semblables aux données réelles.

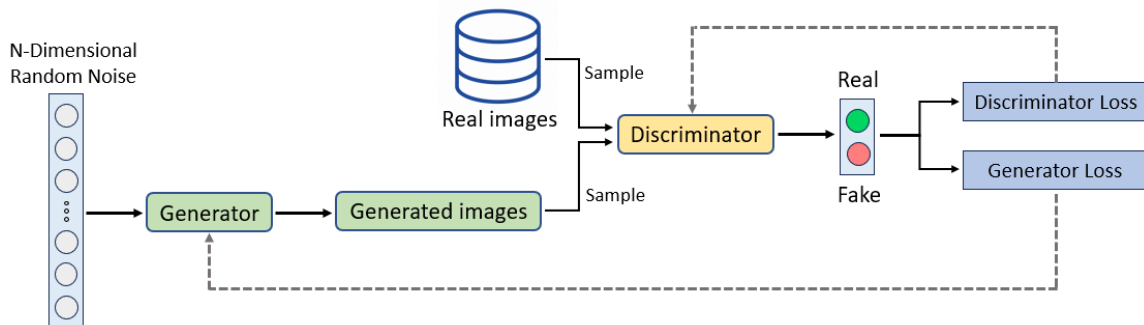


FIGURE 2.17 – Architecture du réseau antagoniste génératif (GAN).

L'augmentation des données peut être particulièrement utile dans le cas d'ensembles de données déséquilibrés, où certaines classes sont moins représentées que d'autres. En générant de nouveaux échantillons pour ces classes minoritaires, l'augmentation des données permet d'équilibrer l'ensemble de données, améliorant ainsi les performances du modèle pour ces classes.

Dans le contexte de la classification des pathologies pulmonaires, cette approche peut être particulièrement utile pour les pathologies peu fréquentes, telles que le cancer du poumon et autres pathologies rares, qui ne sont présentes que dans un nombre limité d'images dans la base de données. Plusieurs études ont utilisé les GANs comme méthode d'augmentation des données pour accroître le nombre d'images CXR, ce qui permet d'améliorer les performances du modèle de classification [133, 134, 135, 136]. Salehinejad et al. [133] ont démontré que l'ajout à l'ensemble de données original, initialement déséquilibré, des images générées par un GAN améliore la performance de la classification des pathologies thoraciques par le réseau de neurones convolutif profond (DCNN) proposé, par rapport à ce même DCNN entraîné uniquement sur l'ensemble de données d'origine. Cette amélioration des performances est largement attribuée à l'équilibrage de l'ensemble de données grâce aux images générées par le GAN, en se concentrant en particulier sur l'augmentation des classes peu représentées dans les images existantes. Albahli et al. ont généré de nouvelles images CXR synthétiques pour traiter le problème du déséquilibre des données en utilisant le Deep Convolutional Generative Adversarial Network (DCGAN) [136]. Ils ont entraîné plusieurs modèles d'apprentissage

profond, avec et sans augmentation des données. Les résultats obtenus montrent que l'ajout des images générées par DCGAN améliore l'accuracy du modèle de 5%. Kora et al. ont procédé à une augmentation du nombre des images de la classe sous-représentée, à savoir les images CXR normales, en utilisant le Deep Convolutional Generative Adversarial Network (DCGAN) [135]. Après l'entraînement du modèle DCGAN, des images de haute qualité de type CXR ont été générées. Ensuite, ils ont formé un classificateur de réseau de neurones en utilisant l'ensemble de données augmenté par le DCGAN ainsi que l'ensemble de données augmenté traditionnel pour la classification de la pneumonie. Les résultats montrent que les méthodes basées sur le GAN ont obtenu les meilleurs résultats parmi les autres méthodes d'augmentation des données.

En somme, l'augmentation des données représente une technique efficace pour améliorer les performances des modèles d'apprentissage profond dans la classification des maladies pulmonaires à partir d'images de radiographie thoracique. En augmentant la variété des images dans l'ensemble de données, cette technique permet au modèle de mieux se généraliser à de nouvelles images et à améliorer son efficacité face à des ensembles de données déséquilibrés. Il est cependant important de choisir soigneusement les techniques d'augmentation les plus adaptées, en tenant compte des caractéristiques spécifiques à l'ensemble de données et de la tâche à accomplir.

2. L'apprentissage par transfert

L'apprentissage par transfert est une autre technique largement utilisée dans le domaine médical où les ensembles de données annotées sont limités et difficiles à obtenir. Cette technique compense la nécessité d'une grande base de données pour l'entraînement efficace des réseaux de neurones profonds tout en évitant le sur-apprentissage des modèles. L'apprentissage par transfert consiste à transférer les connaissances acquises par un modèle sur un grand ensemble de données pour un problème spécifique (source) pour traiter un autre problème (cible) [137].

En utilisant des modèles pré-entraînés au départ, l'apprentissage par transfert permet d'améliorer les performances des modèles, surtout lorsque les données d'entraînement sont limitées. De plus, il peut contribuer à réduire le temps et les ressources nécessaires pour entraîner un modèle à partir de zéro.

L'apprentissage par transfert s'est avéré efficace dans diverses tâches médicales, y compris la classification des pathologies pulmonaires [119, 138, 139, 140, 141, 142].

Dans le cadre de l'analyse d'images de radiographie thoracique, l'apprentissage par transfert consiste à utiliser un modèle d'apprentissage profond déjà entraîné sur un grand ensemble de données et à l'affiner sur un ensemble de données spécifique, généralement plus petit. L'apprentissage par transfert peut également être utilisé comme extracteur de caractéristiques, tirant parti des caractéristiques apprises sur une tâche similaire [143]. En utilisant l'apprentissage par transfert comme extracteur de caractéristiques, il suffit de ré-entraîner uniquement la couche finale du modèle dédiée à la classification, tout en gelant toutes les couches précédentes [144].

Plusieurs études ont démontré que l'apprentissage par transfert améliore l'efficacité de la classification par rapport à la construction et à l'entraînement d'un réseau à partir de zéro [145]. Pour classer les images CXR en pneumonie ou Covid-19, Choudhuri et al. ont utilisé un modèle VGG-16 pré-entraîné tout en affinant ses couches supérieures [146]. L'ensemble de données utilisé comprenait 1 368 images CXR obtenues à partir de trois bases de données. Avec un CNN construit à partir de zéro, ils ont obtenu une accuracy de 96,6%, tandis que le modèle basé sur le VGG-16 a atteint une accuracy de 98,3%.

2.6 Evaluation du modèle d'apprentissage

Suite à la phase d'entraînement, le modèle est évalué sur un ensemble de test. Plusieurs métriques sont utilisées pour évaluer la performance des modèles d'apprentissage, y compris la matrice de confusion et les métriques associés telles que l'accuracy, la précision, le rappel, la spécificité, le F1-score et l'AUC (Aire sous la courbe ROC).

Dans le cadre de cette évaluation, quatre termes principaux sont généralement définis : les vrais positifs (VP), les vrais négatifs (VN), les faux positifs (FP) et les faux négatifs (FN). Les termes VP et VN désignent les prédictions correctes effectuées par le modèle pour les cas respectivement labellisés comme positifs et négatifs. Les FP et FN représentent les erreurs de classification du modèle. Les FP concernent les cas où le modèle a incorrectement prédit une entrée comme positive alors qu'elle était réellement négative, tandis que les FN se réfèrent aux instances où le modèle a prédit de manière erronée une entrée comme étant négative alors qu'elle est effectivement positive.

- L'Accuracy est une mesure de la capacité du classificateur à prédire correctement les cas dans leur bonne classe. Elle est calculée comme la proportion des échantillons correctement classés par rapport au nombre total d'échantillons (Eq. 2.7).

$$Accuracy = \frac{VP + VN}{VP + FP + VN + FN} \quad (2.7)$$

- Le Précision (*Precision*) est définie comme le rapport entre les échantillons positifs correctement détectés et les échantillons qui ont été détectés comme positifs (Eq. 2.8).

$$Précision = \frac{VP}{VP + FP} \quad (2.8)$$

- Le Rappel (*Recall*), également appelé sensibilité, est le pourcentage d'instances positives d'une classe particulière qui sont correctement détectées. Il est défini comme le rapport entre les vrais échantillons positifs et le nombre total d'échantillons positifs (Eq. 2.9).

$$Rappel = \frac{VP}{VP + FN} \quad (2.9)$$

- La Spécificité (*Specificity*) mesure la capacité du modèle à identifier correctement les observations négatives. Elle est définie comme le rapport entre les vrais échantillons négatifs et le nombre total d'échantillons négatifs (Eq. 2.10).

$$Spécificité = \frac{VN}{VN + FP} \quad (2.10)$$

- Le F1-score est défini comme la moyenne harmonique de la précision et du rappel (Eq. 2.11).

$$F1_score = \frac{2 \times Précision \times Rappel}{Précision + Rappel} = \frac{2 \times VP}{2 \times VP + FP + FN} \quad (2.11)$$

- L'aire sous la courbe ROC (*Area under ROC curve* – AUC) mesure l'aire sous la courbe obtenue en traçant le taux de vrais positifs (Rappel) par rapport au taux de faux positifs à différents seuils de classification. Il s'agit d'une mesure de la capacité d'un modèle à faire la distinction entre les classes.

2.7 Classification binaire des pathologies spécifiques

Plusieurs pathologies respiratoires peuvent affecter les patients, parmi lesquelles la tuberculose, la pneumonie et le cancer du poumon qui sont considérées comme les maladies pulmonaires graves les plus courantes. En plus, la pandémie de COVID-19 a récemment touché la population mondiale, infectant et causant la mort des millions d'individus. L'intelligence artificielle a joué un rôle important dans le domaine biomédical, notamment dans l'identification et la classification des pathologies pulmonaires. De nombreuses recherches ont utilisé des modèles basés sur l'apprentissage profond pour le diagnostic des pathologies pulmonaires à partir d'images de radiographie thoracique. Les Tables 2.2, 2.3, 2.4 et 2.5 présentent les modèles et les résultats obtenues par différentes études dans la classification de la tuberculose, de la pneumonie, des nodules pulmonaires et du COVID-19, respectivement, en utilisant des images de radiographie thoracique.

Plusieurs études ont utilisé des CNNs et des modèles pré-entraînés sur la base de données ImageNet, avec diverses architectures telles que VGGNet, ResNet, DenseNet, etc., pour la classification des différentes pathologies pulmonaires à partir des CXRs. Par exemple, Rohilla et al. [147] ont utilisé les architectures AlexNet et VGGNet pour classer la tuberculose pulmonaire à partir d'images de radiographie thoracique provenant des ensembles de données Shenzhen et MC. Ils ont effectué des techniques d'augmentation d'image incluant une rotation de 10 degrés, des translations de 0,1 en largeur et hauteur, ainsi qu'un zoom de 0,1. Le modèle basé sur VGGNet a démontré une accuracy de 81,6%, surpassant celui basé sur AlexNet, principalement grâce à sa structure plus profonde favorisant l'apprentissage d'un plus grand nombre de caractéristiques distinctives. Rajpurkar et al. [39] ont développé CheXNet, un modèle DenseNet à 121 couches, pour identifier la pneumonie à partir de la base de données ChestX-ray14. Ce modèle a surpassé la performance moyenne des radiologues pour la tâche de détection de la pneumonie, atteignant une AUC de 76,8%. Rahman et al. [148] ont entraîné plusieurs réseaux profonds, y compris DenseNet-201, AlexNet, ResNet-18, DenseNet-201 et SqueezeNet, pour identifier la pneumonie à partir de la base de données Chest X-ray Images (Pneumonia). Ils ont appliqué trois techniques d'augmentation : la rotation de 315 degrés (45 degrés dans le sens contraire des aiguilles d'une montre), le redimensionnement et la translation. Les résultats ont montré que DenseNet-201 dépassait les autres architectures, atteignant une accuracy de 98% pour la classification des images de pneumonie et des images normales.

Certaines études ont utilisé plusieurs architectures des modèles pré-entraînés pour développer des modèles d'ensemble. La plupart de ces modèles d'ensemble intègrent des variations des architectures ResNet, AlexNet, GoogleNet, EfficientNet et XceptionNet. Par exemple, Hooda et al. [139] ont utilisé un modèle d'ensemble composé de trois architectures, notamment AlexNet, GoogleNet et ResNet-34, pour la détection de la tuberculose à partir des images CXR obtenues en combinant quatre ensembles de données. Ce modèle a atteint une accuracy de 88,24% et une AUC de 93%. Kumar et al. [129] ont développé un modèle d'ensemble en combinant plusieurs architectures, notamment EfficientNet, GoogleNet et XceptionNet, pour le diagnostic précoce du COVID-19. Ils ont appliqué des techniques d'augmentation des données telles que le retournement horizontal et vertical, une transformation en cisaillement (avec un angle de 0,2) et une rotation de 45 degrés. Ce modèle d'ensemble a obtenu une accuracy de 98,95% pour la classification binaire de l'absence et la présence du COVID-19.

Dans une étude unique, Huy et al. ont développé un modèle basé sur DenseNet, intégrant un module d'attention pour identifier la tuberculose, atteignant une accuracy de 98,8% [149].

TABLE 2.2 – Performance des méthodes de classification de la tuberculose utilisées dans la littérature.

Référence	Année	Base de données	Modèle	Accuracy	AUC	Rappel	Spécificité
[150]	2016	Shenzhen	CNN	90.3%	96.7%	-	-
[147]	2017	Shenzhen + MC	VGGNet	81.6%	-	-	-
[138]	2017	Shenzhen	Ensemble model	90%	-	88%	92%
[151]	2017	MC + Shenzhen + 2 Private datasets	AlexNet + GoogLeNet	-	99%	97.3%	100%
[139]	2019	MC + JSRT + Shenzhen + Private dataset	AlexNet + GoogleNet + ResNet-34	88.24%	93%	88.42%	88%
[105]	2020	Shenzhen	CNN	84%	-	-	-
[56]	2020	Shenzhen	EfficientNet-B4	89.92%	94.8%	-	-
[86]	2020	Six different datasets	Nine pre-trained CNN	98.6%	-	-	-
[60]	2022	MC + Shenzhen	CNN + DenseNet-169	98.98%	-	98.98%	-
[149]	2023	Tuberculosis chest X-ray images + MC + Shenzhen	CBAMWDNet	98.8%	-	94.28%	95.7%
[88]	2023	Shenzhen + MC	Dual convolution blocks + DenseNet-169	98.98%	-	98.98%	99.03%
[75]	2023	Tuberculosis chest X-ray images dataset (Six different datasets)	Xception	99.29%	99.9%	98.29%	99.3%

TABLE 2.3 – Performance des méthodes de classification de la pneumonie utilisées dans la littérature.

Référence	Année	Base de données	Modèle	Accuracy	AUC	Rappel
[39]	2017	ChestX-ray14	DenseNet-121	-	76.8%	-
[152]	2020	Chest X-ray Images (Pneumonia)	CNN	98.46%	98.48%	-
[153]	2020	Unspecified Kaggle dataset	Deep Siamese based neural network	-	97%	-
[148]	2020	Chest X-ray Images (Pneumonia)	DenseNet-201	98%	-	99%
[128]	2020	Chest X-ray Images (Pneumonia)	ResNet-152 v2	99.22%	99.77%	99.44%
[127]	2020	Kaggle dataset	CNN	90.68%	-	-
[135]	2020	Pediatric CXRs	GAN + CNN	95.5%	93.6%	97.7%
[119]	2021	Pediatric CXRs	ResNet-50 + DenseNet-121	-	99.6%	98.12%
[154]	2022	Kaggle dataset	VGG-16	95.6%	-	-
[155]	2023	Chest X-ray Images	VGG-16 + NN	92.15%	97.4%	93.08

TABLE 2.4 – Performance des méthodes de classification des nodules pulmonaires utilisées dans la littérature.

Référence	Année	Base de données	Modèle	Rappel	Spécificité	Accuracy
[156]	2016	JSRT	ResNet-50	92%	86%	-
[77]	2017	JSRT	AlexNet + Random Forest	69.3%	96.2%	-
[157]	2018	JSRT	DenseNet-121	74.68%	74.96%	74.43%
[158]	2019	JSRT	R-CNN	64.2%	-	-
[83]	2020	JSRT	CNN	79.6%	-	-
[81]	2020	JSRT	GoogleNet	97.1%	-	-

TABLE 2.5 – Performance des méthodes de classification du Covid-19 utilisées dans la littérature.

Référence	Année	Base de données	Modèle	Accuracy	Rappel	Spécificité	F1-score
[159]	2020	Covid-chestxray, RSNA, SIRM	MobileNet v2	96.78%	98.66%	96.46%	-
[160]	2020	Covid-chestxray, Unspecified Kaggle dataset, Open-i	Inception v3	100%	100%	100%	100%
[161]	2020	Covid-chestxray, COVIDx, RSNA Pneumonia	VGG-16, Faster R-CNN	97.36%	97.65%	95.48 %	98.46%
[162]	2020	Covid-chestxray	MobileNet	98.6%	87.8%	99.3%	87.8%
[163]	2020	Covid-chestxray, Chest X-ray Images (Pneumonia)	VGG-16	88.1%	97.62%	78.57 %	-
[164]	2020	Covid-chestxray, MC, Shenzhen, Chest X-ray Images (Pneumonia)	Truncated InceptionNet	99.92%	93%	100 %	96%
[165]	2020	Covid-chestxray, RSNA Pneumonia	CNN + Capsule network	98.3%	80%	98.6%	-
[57]	2020	Covid-chestxray, RSNA Pneumonia	Xception	99%	99.3%	98.6%	98.5%
[166]	2021	Eight publicly datasets	CNN	99.12%	95.3%	97.36%	96.88%
[167]	2022	Kaggle dataset	CNN	96.71%	-	-	91.89%
[62]	2023	Five publicly datasets	ResNet-34 + HRNet	99.99%	99.98%	100%	-
[129]	2023	Kaggle dataset	EfficientNet + GoogLeNet + XceptionNet	98.95%	-	-	-

2.8 Classification multi-classe des pathologies

La pandémie COVID-19 a souligné l'importance cruciale de disposer de techniques de classification précises et robustes pour les pathologies respiratoires. Cette nécessité est amplifiée par le fait que COVID-19 peut présenter des symptômes semblables à d'autres pathologies telles que la pneumonie ou le cancer du poumon, rendant un diagnostic rapide et précis essentiel pour déterminer le traitement adéquat et pour alléger la charge sur les systèmes de soins de santé. En outre, des étiologies de la tuberculose représentent également un défi pour les professionnels de la santé.

Face à ces défis, le développement des systèmes d'intelligence artificielle robustes, capables de distinguer efficacement entre différentes pathologies pulmonaires aux signes cliniques parfois similaires, devient crucial. Ces technologies promettent d'optimiser la prise de décision clinique, assurant ainsi une gestion thérapeutique précise et adaptée pour chaque patient. La recherche dans ce domaine a connu une attention accrue au cours des dernières années, se concentrant sur le développement des modèles capables de réaliser une classification multi-classe précise des pathologies pulmonaires. Le tableau 2.6 présente les méthodes et les résultats rapportés dans la littérature pour la classification multi-classe de différentes pathologies pulmonaires à partir des images CXR.

Plusieurs architectures CNN, notamment MobileNet, VGGNet, ResNet, DenseNet, XceptionNet et EfficientNet, ont été utilisées dans de nombreuses études pour effectuer une classification multi-classe des maladies pulmonaires. Ces modèles ont atteint des taux d'accuracy élevés allant de 87,02% à 99,3% (comme le montre la Table 2.6). Par exemple, Apostolopoulos et Mpesiana [159] ont utilisé le modèle MobileNet v2 pour la tâche de classification à trois classes : COVID-19, pneumonie et normal, atteignant une accuracy de 94,72%. Khan et al. [57] ont développé un modèle d'apprentissage profond pour la identifier le COVID-19 dans les images CXR, en classifiant les images en quatre classes distinctes : pneumonie bactérienne, pneumonie virale, COVID-19 et images normales. Leur modèle, basé sur l'architecture Xception pré-entraînée sur ImageNet, a atteint une accuracy de 89,6% pour la classification à quatre classes, qui s'est améliorée à 94,59% en regroupant les cas de pneumonie non COVID-19.

D'autres études ont utilisé des modèles d'ensemble intégrant des variations des architectures ResNet, XceptionNet, DenseNet, VGG, GoogleNet et EfficientNet. Par exemple, Mamalakis et al. [119] ont développé un modèle d'apprentissage par transfert profond appelé DenResCov-19 pour diagnostiquer le COVID-19, la pneumonie, la tuberculose et les

TABLE 2.6 – Performance des modèles de classification multi-classe des pathologies pulmonaires utilisés dans la littérature.

Référence	Année	Base de données	Modèle	Classification	Accuracy	Rappel	AUC
[159]	2020	COVID-19, Normal, Bacterial Pneumonia	MobileNet v2	3-class	94.72%	-	-
[168]	2020	COVID-19, Normal, Viral Pneumonia, Bacterial Pneumonia	VGG-16	4-class	95%	-	-
[121]	2020	COVID-19, Normal, Pneumonia	Xception + ResNet-50 v2	3-class	91.4%	-	-
[169]	2020	COVID-19, Normal, Pneumonia	DarkNet	3-class	87.02%	85.35%	-
[57]	2020	COVID-19, Normal, Viral Pneumonia, Bacterial Pneumonia	Xception	4-class 3-class	89.6% 94.59%	89.92% 96.9%	- -
[59]	2021	COVID-19, Normal, Bacterial Pneumonia	Inception-ResNet v2	3-class	92.18%	92.11%	-
[120]	2021	COVID-19, Normal, Viral Pneumonia, Bacterial Pneumonia	DarkNet-53	4-class 3-class	88.25% 97.11%	85.66% 96.4%	- -
[170]	2021	COVID-19, Normal, Pneumonia, Lung Cancer	VGG-19 + CNN	4-class	98.05%	-	99.66%
[87]	2021	COVID-19, Normal, Viral Pneumonia, Bacterial Pneumonia, Tuberculosis	ResNet-50	5-class	96.32%	-	-
[119]	2021	COVID-19, Normal, Pneumonia, Tuberculosis	ResNet-50 + DenseNet-121	4-class 3-class	- -	69.7% 89.38%	95% 96.51%
[166]	2021	COVID-19, Normal, Viral Pneumonia, Bacterial Pneumonia.	CNN	4-class 3-class	91.2% 94.2%	91.9% 92.5%	- -
[61]	2022	COVID-19, Normal, Pneumonia	Xception	3-class	99.3%	99%	-
[171]	2022	COVID-19, Normal, Viral Pneumonia, lung opacity	EfficientNet-B1	4-class	96.13%	-	-
[60]	2022	COVID-19, Normal, Pneumonia, Tuberculosis	CNN + DenseNet-169	4-class	95.10%	95.10%	-
[155]	2023	COVID-19, Normal, Pneumonia	VGG-16 + NN	3-class	95.4%	95.4%	98.8%
[88]	2023	COVID-19, Normal, Pneumonia, tuberculosis	Xception	4-class	95.1%	95.1%	-
[129]	2023	COVID-19, Normal, Pneumonia, tuberculosis	EfficientNet + GoogLeNet + XceptionNet	4-class	99.21%	-	-
[172]	2023	COVID-19, Normal, Pneumonia, tuberculosis, Lung Cancer, Lung Opacity	VGG-19 + CNN	6-class	96.48%	93.75%	99.82%

cas normaux à partir d'images CXR. Ce modèle est composé des réseaux DenseNet-121 et ResNet-50, atteignant une AUC de 96,51% pour la classification en 3 classes (COVID-19, normal et pneumonie) et 95% la classification en 4 classes (COVID-19, normal, pneumonie et tuberculose).

Dans une étude unique, Xu et al. [87] ont proposé le réseau multi-attention (MANet) pour classer les images CXR en cinq catégories distinctes : normal, COVID-19, tuberculose, pneumonie bactérienne et pneumonie virale. L'architecture MANet comprend deux étapes : la segmentation et la classification. Initialement, les régions pulmonaires des images CXR sont segmentées à l'aide d'un modèle U-Net qui utilise ResNet comme structure de base. Ensuite, un mécanisme d'attention est appliqué sur les masques prédits pour ajuster les caractéristiques du CNN à la deuxième étape, éliminant ainsi les caractéristiques dont les champs réceptifs se trouvent à l'extérieur des poumons. Quatre architectures CNN avec et sans MA, à savoir ResNet-34, ResNet-50, VGG-16 et Inception v3, sont employés pour classer les images CXR segmentées en cinq catégories. Les résultats expérimentaux indiquent que le modèle ResNet-50 avec MA atteint l'accuracy de test moyenne la plus élevée de 96,32%.

2.9 Classification multi-label des pathologies

L'imagerie par radiographie thoracique, riche en informations diagnostiques, révèle souvent la présence simultanée de plusieurs pathologies, posant ainsi un défi de classification multi-label. Ainsi, une image CXR peut être associée à plusieurs pathologies, telles que COVID-19, pneumonie, et cardiomégalie, simultanément. Ce type de problème se distingue de la classification multi-classe qui se limite à attribuer une seule étiquette de classe pour chaque entrée. En outre, il peut y avoir des corrélations potentielles entre les différentes étiquettes, qui peuvent offrir des indices diagnostiques importants. Par exemple, la présence d'une cardiomégalie sur une image CXR peut indiquer un risque plus élevé d'œdème pulmonaire [173]. L'exploration de ces dépendances ou corrélations entre les étiquettes peut contribuer à affiner le diagnostic. Face à ces défis, le développement des méthodes d'apprentissage qui tiennent compte de la nature multi-label de la tâche devient essentiel. Cette section présente les travaux existants pour la classification multi-label des pathologies pulmonaires à partir d'images de radiographie thoracique. Les méthodes et les résultats de chaque étude, utilisant les bases de données CheXpert et MIMIC-CXR, ainsi que la base de données ChestX-ray14, sont résumés dans les Tables

2.7 et 2.8 respectivement.

TABLE 2.7 – Performance des modèles de classification multi-label utilisés dans la littérature sur les bases de données MIMIC-CXR et CheXpert.

Référence	Année	Base de données	Modèle	AUC
[174]	2018	MIMIC-CXR	DenseNet-121	72.1%
[175]	2019	CheXpert	DenseNet-121	81.2%
[45]	2019	CheXpert (5 pathologies)	DenseNet-121	90.7%
[142]	2021	CheXpert (5 pathologies)	Set of six CNN architectures	94%
[124]	2021	CheXpert	DenseNet-121 + DenseNet-169	81.5%

Plusieurs architectures basées sur des modèles pré-entraînés ont été utilisées pour la classification multi-label des pathologies à partir des images CXR [47, 40, 176, 177, 178, 179, 123, 180, 181, 126, 182, 183, 184, 185, 82, 3, 173, 186, 187, 188, 189]. Souid et al. [186] ont proposé une approche utilisant le modèle MobileNet v2, suivi par des couches CNN pour la classification des 14 pathologies de la base de données ChestX-ray14, atteignant une AUC de 81%. D'autres études ont utilisé des modèles d'ensemble pour améliorer les performances des modèles [124, 190, 191, 192, 124, 193].

De nombreuses études ont intégré des mécanismes d'attention dans les modèles de classification, montrant que le fait de forcer le modèle à se focaliser sur des régions spécifiques des images CXR donne de meilleurs résultats, surpassant la classification basée sur les images complètes [178, 182, 183, 184, 3, 173, 187, 188]. Guan et al. ont développé un cadre d'apprentissage de l'attention résiduelle par catégorie (CRAL) pour la classification des pathologies thoraciques à partir de la base de données ChestX-ray14 [173]. Le cadre CRAL est composé de deux modules : un module d'intégration des caractéristiques pour l'extraction des caractéristiques de haut niveau à l'aide d'un réseau de neurones convolutif tel que DenseNet-121, et un module d'attention qui attribue des poids aux différentes positions spatiales des cartes de caractéristiques en fonction des scores d'attention basés sur les catégories. Leur méthode a obtenu un score AUC moyen de 81,6%. Chen et al. [188] ont proposé un réseau à double attention intégrant à la fois un module d'attention de canal et un module d'attention spatiale. Cette conception vise à améliorer sélectivement les caractéristiques fortement associées aux zones de lésion, optimisant ainsi la précision de la classification. Leur approche a atteint un score AUC moyen de 81,86% pour la classification des pathologies de la base de données ChestX-ray-14.

TABLE 2.8 – Performance des modèles de classification multi-label utilisés dans la littérature sur la base de données ChestX-ray14.

Référence	Année	Base de données	Modèle	AUC	Accuracy
[47]	2017	ChestX-ray14	ResNet-50	74.51%	-
[194]	2017	ChestX-ray14	DenseNet+LSTM	80.27%	-
[39]	2017	ChestX-ray14	DenseNet-121	84.2%	-
[40]	2018	ChestX-ray14	DenseNet-121	82.8%	-
[176]	2018	ChestX-ray14	DenseNet-161	79.5%	-
[177]	2018	ChestX-ray14	ResNet-152	78.1%	-
[178]	2018	ChestX-ray14	ResNet-50	80.27%	-
[179]	2018	ChestX-ray14	DenseNet-121	84.62%	-
[195]	2018	ChestX-ray14	DenseNet	83.02%	-
[123]	2018	ChestX-ray14	DenseNet-121	80.7%	-
[123]	2018	ChestX-ray14 + PLCO	DenseNet-121	87.4%	-
[180]	2018	ChestX-ray14	ResNet-50	80.64%	-
[196]	2018	ChestX-ray14	CapsuleNet on CNN	77.5%	-
[125]	2019	ChestX-ray14	CNN	-	89.77%
[175]	2019	ChestX-ray14	DenseNet-121	88.2%	-
[181]	2019	ChestX-ray14	DenseNet-121	84.3%	-
[126]	2019	ChestX-ray14	ResNet-38	80.6%	-
[182]	2019	ChestX-ray14	DenseNet-169	82.4%	-
[183]	2019	ChestX-ray14	ResNet-101	79.41%	-
[190]	2019	ChestX-ray14	DenseNet-121 + ResNet-50	82.3%	-
[184]	2019	ChestX-ray14	DenseNet-121	81.7%	-
[80]	2019	ChestX-ray14	DenseNet-121	81.5%	-
[191]	2020	ChestX-ray14	DenseNet-169 + ResNet-152	82%	-
[185]	2020	ChestX-ray14	DenseNet-121	84.32%	-
[82]	2020	ChestX-ray14	DenseNet-169	82.3%	-
[3]	2020	ChestX-ray14	DenseNet-121	87.1%	-
[173]	2020	ChestX-ray14	DenseNet-121	81.6%	-
[186]	2021	ChestX-ray14	MobileNet v2	81%	90.2%
[192]	2021	ChestX-ray14	Inception-ResNet-v2	80.1%	-
[187]	2021	ChestX-ray14	DenseNet-121	85.37%	-
[124]	2021	ChestX-ray14	DenseNet-121 + DenseNet-169	85%	-
[193]	2022	ChestX-ray14	EfficientNet-B4 + DenseNet-169 + ResNet-50	82.7%	-
[188]	2022	ChestX-ray14	DenseNet	81.86%	-
[7]	2023	ChestX-ray14	ConvNeXt + BioBert	82.6%	-
[189]	2023	ChestX-ray14	DenseNet-121	83.01%	-

D'autres études ont développé des fonctions de perte pour prendre en compte la dépendance entre les étiquettes et pour améliorer les performances des modèles [179, 181, 185, 124]. Ma et al. [184] ont développé un réseau d'attention croisée pour la classification des pathologies thoraciques, intégrant une fonction de perte d'attention pour encourager le modèle à se focaliser sur les régions pathologiques, et une perte d'équilibre multi-label pour surmonter le déséquilibre entre les classes. Le modèle a atteint un score AUC moyen de 81,7% pour la classification des pathologies de la base de données ChestX-ray-14.

Dans une étude unique, Baltruschat et al. ont développé une architecture qui intègre des données non liées à l'image, telles que l'âge, le genre et le type d'acquisition des images, dans le processus de classification à partir de la base de données ChestX-ray14 [126]. Ils ont examiné l'utilisation de l'apprentissage par transfert avec et sans réglage fin, ainsi que la formation d'un réseau à partir de zéro dédié spécifiquement aux CXRs. Ils ont également évalué plusieurs profondeurs d'architectures ResNet, notamment ResNet-38, ResNet-50 et ResNet-101. Les résultats de leur étude indiquent que l'utilisation d'une architecture ResNet-38 spécifiquement adaptée aux rayons X et l'intégration des données non images ont donné les meilleurs résultats globaux avec un score AUC moyen de 80,6%.

2.10 Conclusion

Dans ce chapitre, nous avons présenté les bases de données disponibles et les techniques de prétraitement, telles que l'amélioration de l'image, qui améliorent la performance de classification des différentes pathologies pulmonaires. Nous avons également donné un aperçu des algorithmes de classification utilisés pour l'identification des pathologies pulmonaires.

L'apprentissage profond, notamment les réseaux de neurones convolutifs et les nombreuses architectures, a démontré une efficacité élevée dans le diagnostic des pathologies pulmonaires à partir des images de radiographie thoracique. Ces résultats remarquables résultent de l'application des techniques telles que le prétraitement, l'augmentation des données, la segmentation ainsi que l'incorporation des mécanismes d'attention.

Ce chapitre a fait l'objet d'une publication dans la revue "*SN Computer Science*" [41].

Dans le chapitre suivant, nous présenterons notre première contribution à la classification des pathologies pulmonaires à partir des images CXR.

DEUXIÈME PARTIE

**Contributions : Classification des
maladies pulmonaires par des
modèles d'apprentissage
automatique**

AMÉLIORATION DU DIAGNOSTIC DES MALADIES PULMONAIRES : ANALYSE D'IMAGES ET EXPLOITATION DES DONNÉES CLINIQUES POUR UNE MEILLEURE CLASSIFICATION

3.1 Introduction

Pour la mise en œuvre efficace des algorithmes d'apprentissage automatique dans le domaine de la radiologie, il est crucial que les chercheurs s'appuient sur des données fiables. Cela inclut non seulement les images radiologiques elles-mêmes, mais aussi les métadonnées qui sont indispensables pour le processus de décision [197]. Les métadonnées associées aux images radiographiques, telles que la position de vue de la radiographie (c'est-à-dire si les images de rayons X ont été acquises en position postéro-antérieure (PA) ou antéro-postérieure (AP)), l'âge et le genre du patient, sont cruciales pour l'interprétation précise des radiographies par les médecins [198].

Ces informations supplémentaires fournissent un contexte précieux qui peut affecter de manière significative non seulement l'interprétation des images, mais aussi la prise de décision concernant la présence d'une pathologie. Par exemple, la position de vue d'une image peut révéler des détails anatomiques spécifiques qui sont cruciaux pour identifier correctement les anomalies ou les pathologies [199]. L'âge du patient influence les critères d'évaluation en raison des changements anatomiques et physiologiques chez les enfants par rapport aux adultes, car certaines pathologies sont plus prévalentes ou se manifestent différemment selon les groupes d'âge. De même, le genre du patient impacte l'évaluation

des images, compte tenu des différences anatomiques et physiologiques entre les hommes et les femmes, qui peuvent affecter à la fois l’apparence et l’interprétation des signes radiologiques des pathologies.

L’objectif de ce chapitre est d’aller au-delà de la simple analyse d’image conventionnelle en exploitant les données cliniques contenues dans la base de données. Prendre en compte ces métadonnées dans les systèmes diagnostiques automatisés permet d’ajuster les paramètres des algorithmes pour s’adapter à ces variables, permettant une évaluation d’image plus précise. Cela contribue à affiner le modèle d’apprentissage automatique et donc le diagnostic selon le profil du patient, menant à des décisions thérapeutiques plus précises et spécifiques au patient, améliorant ainsi l’efficacité du diagnostic.

La base de données ChestX-ray14 [47], largement utilisée pour développer des modèles d’apprentissage automatique pour le diagnostic des pathologies thoraciques, comprend des images de radiographie thoracique ainsi que des données cliniques telles que la position de vue de la radiographie, l’âge et le genre du patient. Dans cette base de données, 322 images sont étiquetées comme pneumonie et 60 361 images comme absence de pathologies (no finding). Pour remédier au déséquilibre des classes, nous avons sous-échantillonné uniquement les images de no finding pour obtenir 354 images. Par conséquent, nous avons mené nos expériences sur un total de 676 images et en exploitant leurs données cliniques correspondantes. La Figure 3.1 illustre le nombre d’images dans chaque classe ainsi que le nombre total d’images dans l’échantillon utilisé de la base de données.

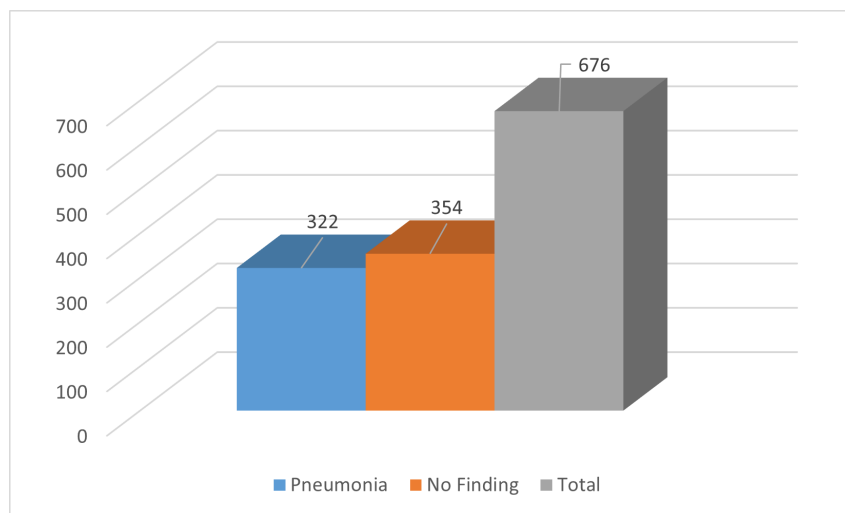


FIGURE 3.1 – Visualisation du nombre d’images dans chaque classe de l’échantillon de la base de données.

Un exemple d'images radiographiques de pneumonie et de no finding qui sont acquises à partir de la base de données ChestX-ray14 est illustré dans la Figure 3.2.

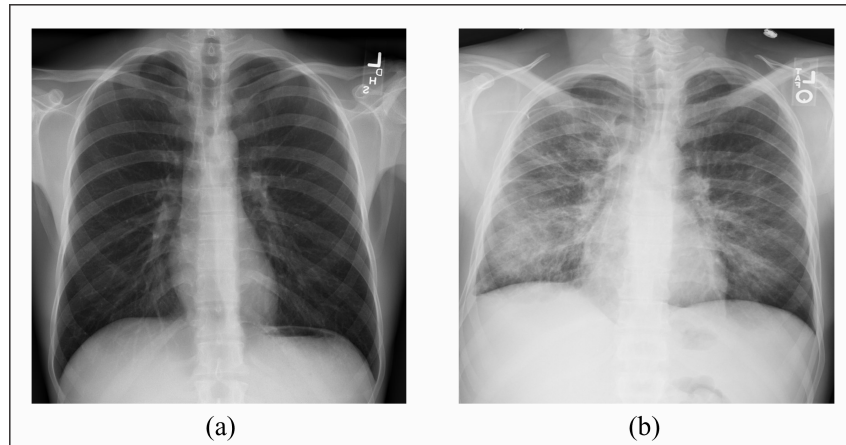


FIGURE 3.2 – Exemples d'images acquises à partir de la base de données ChestX-ray14 : (a) CXR normal et (b) Pneumonie. La pathologie en (b) peut être distinguée par la présence d'opacités anormales dans le lobe pulmonaire.

3.2 Approche proposée

La Figure 3.3 montre un aperçu de la méthodologie utilisée. Un total de 676 images CXR de pneumonie et no finding avec leurs caractéristiques cliniques correspondantes ont été utilisées. Notre approche implique l'utilisation du clustering non supervisé puis l'analyse des données cliniques dans chaque cluster pour identifier des clusters partageant des caractéristiques similaires. Le principal avantage de l'utilisation de l'analyse hiérarchique en clusters est qu'elle permet de diviser nos données en groupes sans présupposer le nombre de clusters, ainsi qu'elle permet une visualisation des sous-groupes et une meilleure compréhension des caractéristiques cliniques de la population étudiée. Ensuite, pour chaque cluster, les images ont été redimensionnées à 224×224 pixels, et normalisées dans un intervalle de $[0; 1]$. La méthode Masque Flou a été appliquée pour améliorer le contraste des images. Par la suite, 20% des données de chaque cluster ont été utilisées comme données de test, 10% comme données de validation et 70% ont été consacrées à l'entraînement des données (choisies aléatoirement). Un modèle d'apprentissage profond a ensuite été développé pour la classification, intégrant une couche d'attention pour se concentrer sur les régions pertinentes de l'image. Les images de l'ensemble de test ont

été utilisées pour évaluer la performance du modèle qui a été quantifiée par l'aire sous la courbe (AUC). Le langage de programmation utilisé est Python avec l'implémentation des bibliothèques suivantes : numpy, pandas, matplotlib, scipy, tensorflow, scikit-learn et scikit-image.

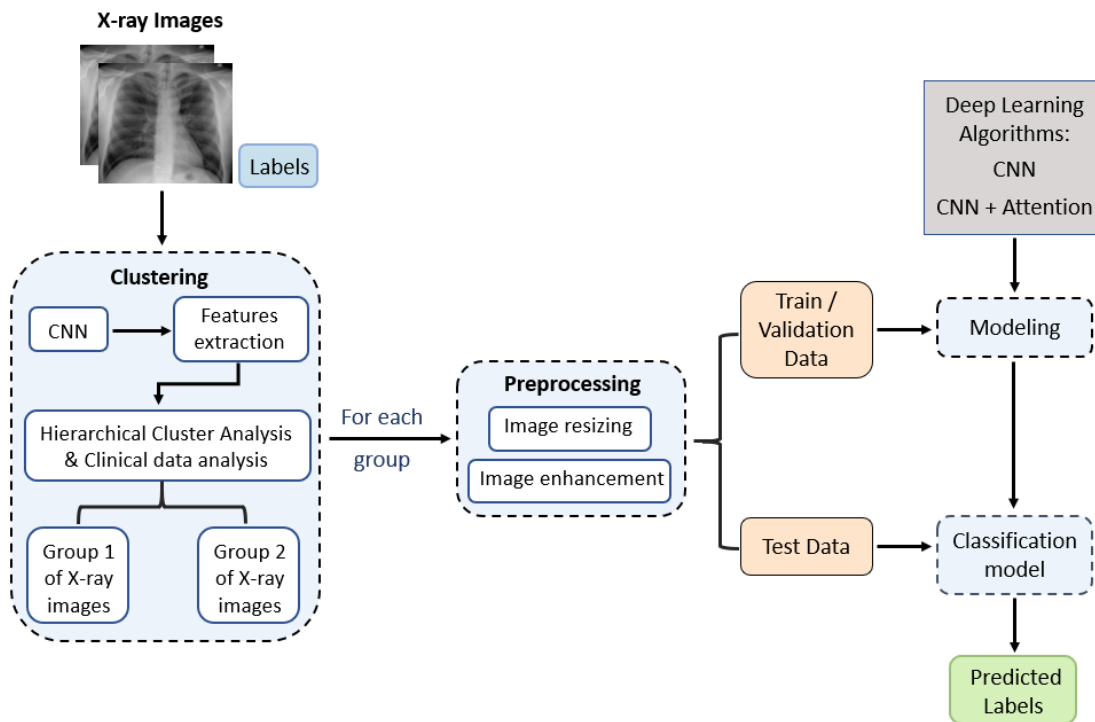


FIGURE 3.3 – Aperçu de la méthodologie proposée.

3.3 Identification des groupes d'images avec l'Analyse Hiérarchique en Clusters

Pour exploiter pleinement les informations de la base de données et ne pas se restreindre exclusivement à l'analyse des images, nous avons adopté une nouvelle approche, via l'Analyse hiérarchique en clusters, qui consiste à considérer les caractéristiques cliniques dans la classification, ce qui permet une analyse plus complète en examinant à la fois les images et les données cliniques associées.

L'Analyse Hiérarchique en Clusters ou *Hierarchical Cluster Analysis* (HCA) est une méthode de classification automatique non supervisée qui organise les données en groupes

basés sur leur similarité [200, 201]. Cette méthode est particulièrement utile pour comprendre la structure sous-jacente des données analysées. Elle utilise souvent la distance euclidienne pour mesurer la similarité entre les objets de données, bien que d'autres mesures puissent également être appliquées. Dans ce processus, le linkage définit comment la distance entre les clusters est calculée à partir des distances entre objets individuels, avec pour objectif de former des clusters dont les objets sont le plus similaires possible. Un type de linkage couramment utilisé est le linkage de Ward, qui vise à minimiser l'augmentation de la variance totale à l'intérieur des clusters après chaque fusion. Ainsi, le linkage de Ward cherche à fusionner les clusters de manière à ce que l'impact sur l'augmentation de la variance intra-cluster soit minimisé, préservant l'homogénéité des clusters au cours du processus de regroupement.

Le dendrogramme, une figure arborescente produite par le clustering hiérarchique, illustre les relations entre les groupes de données. Chaque point de donnée est représenté individuellement à la base du dendrogramme, et à mesure que l'on monte vers le sommet, les groupes de données deviennent de plus en plus larges, regroupant progressivement tous les points de données selon leur similarité basée sur une métrique de distance définie. La longueur des branches reflète la distance entre les groupes : les branches plus courtes indiquent une forte similarité, signifiant que les points de données dans ces clusters sont très proches, tandis que les branches plus longues suggèrent une moindre similarité.

Dans notre travail, nous avons effectué une HCA afin d'identifier les images des patients qui présentent des caractéristiques similaires, facilitant la compréhension des motifs et des relations au sein de la population étudiée. Cette connaissance peut être précieuse pour la prise de décision clinique, car elle permet de développer des stratégies de prévention ou de surveillance plus ciblées et plus adaptées à chaque sous-groupe. Cette approche améliore potentiellement la précision des soins de santé, contribuant ainsi à des résultats cliniques plus efficaces et bénéfiques pour les patients et à une efficacité accrue des traitements.

Tout d'abord, nous avons utilisé un CNN avec cinq couches de convolution pour extraire les caractéristiques des images de radiographie thoracique. Ces caractéristiques sont ensuite utilisées comme données d'entrée pour la HCA. Le dendrogramme produit par la HCA, illustré dans la Figure 3.4 (a), indique une partition des données en quatre groupes. Nous avons procédé à une analyse détaillée des caractéristiques cliniques associées aux images de chaque cluster pour comprendre les nuances des profils de patients et des images.

Les Figures 3.4 (b) et (c) montrent que les clusters contiennent majoritairement des patients âgés de plus de 49 ans, avec une prédominance de patients masculins par rapport aux patients féminins. Cependant, une variation significative est observée dans la position de vue des images entre les clusters, indiquant des préférences techniques distinctes qui influencent la composition des groupes. Par conséquent, nous avons regroupé les clusters en fonction de leurs caractéristiques dominantes de position de vue, formant ainsi deux nouveaux clusters : un majoritairement composé d'images en vue postéro-antérieure (PA) (91.5%) et l'autre en vue antéro-postérieure (AP) (60%). Cette distinction est cliniquement significative car la position de vue peut affecter la visibilité des structures anatomiques et la détection des anomalies, ce qui est crucial pour un diagnostic précis.

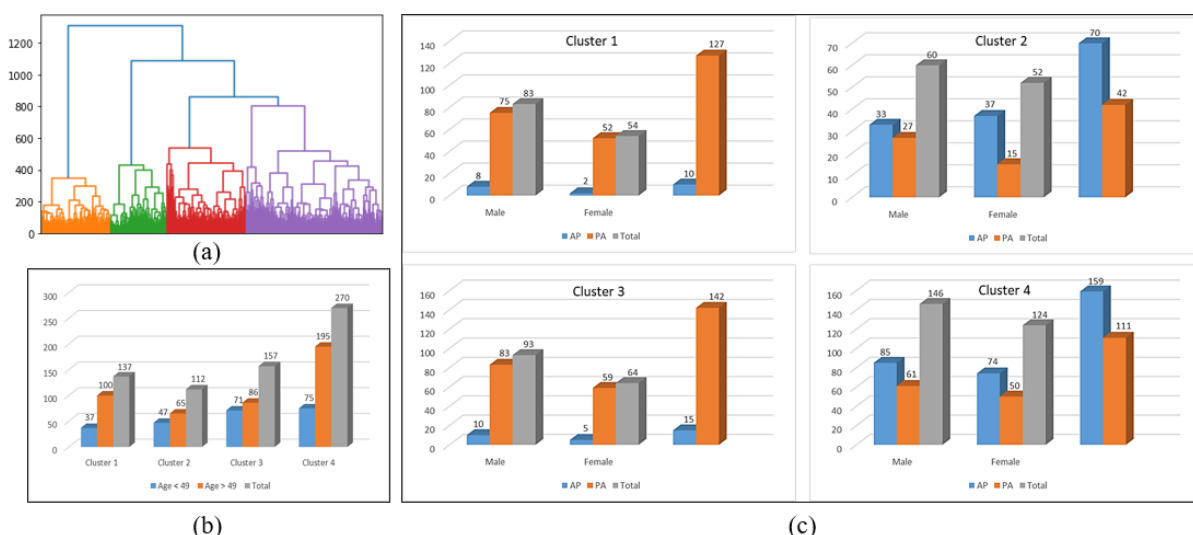


FIGURE 3.4 – (a) Dendrogramme produit par la HCA pour les images de pneumonie et de no finding, (b) Répartition de l'âge des patients dans chaque cluster, et (c) Répartitions du genre des patients et de la position de vue des images CXR dans chaque cluster.

Dans la suite de notre étude, nous avons appliqué une méthode de prétraitement adaptée et des modèles de classification spécifiques à chaque cluster, puis nous avons calculé l'aire sous la courbe (AUC) pondérée pour évaluer la performance des modèles de classification.

3.4 Prétraitement des images avec le Masque Flou

Après le clustering, nous avons amélioré le contraste des images en utilisant la méthode Masque Flou ou *Unsharp Masking* (UM). Le prétraitement des images est essentiel pour en améliorer la qualité et extraire les informations pertinentes, les rendant ainsi plus adaptées aux algorithmes d'apprentissage automatique. Cette technique UM affine l'image originale en augmentant son contraste, ce qui permet de mettre en évidence les textures et les détails. Le principe de base de la méthode UM consiste à soustraire l'image originale par une version floue de l'image elle-même, ce qui permet de conserver uniquement les contours accentués [63, 64]. La formule typique de l'UM est donnée par l'équation 3.1 :

$$Image_améliorée = Image_originale + (Image_originale - Image_floue) \times Montant \quad (3.1)$$

L'efficacité de la technique UM est influencée par le choix des paramètres de rayon (radius) et du montant (amount). Ces paramètres jouent un rôle crucial dans la détermination du degré d'accentuation appliqué à l'image. Le processus commence par l'application d'un flou, généralement réalisé avec un filtre gaussien. Le paramètre de rayon du filtre, aussi appelé sigma, contrôle le degré de flou appliqué à l'image originale, affectant directement la zone autour des contours qui sera mise en évidence par l'accentuation. L'effet d'amélioration est quant à lui déterminé par le paramètre d'intensité ou montant, qui spécifie la quantité de contraste ajoutée aux bords. En ajustant finement ces deux paramètres, on peut optimiser la clarté et le contraste des détails de l'image, améliorant ainsi la visibilité des éléments essentiels pour une analyse précise.

Dans notre étude, pour déterminer les paramètres optimaux pour la technique UM appliquée à chaque cluster, nous avons mené une étude de sensibilité sur les paramètres, testant diverses valeurs du rayon de 1 à 5 et du montant de 1 à 20. Cette analyse a révélé que les combinaisons de rayon et d'intensité de 4 et 6, ainsi que de 5 et 8, étaient les plus efficaces pour améliorer la performance de classification des images normales et de pneumonie dans le premier et le second cluster, respectivement. En conséquence, nous avons utilisé ces valeurs pour toutes les analyses ultérieures dans notre étude. La Figure 3.5 présente un exemple de l'amélioration obtenue avec le Masque Flou sur une image radiographique de chaque cluster en utilisant ces paramètres spécifiques.

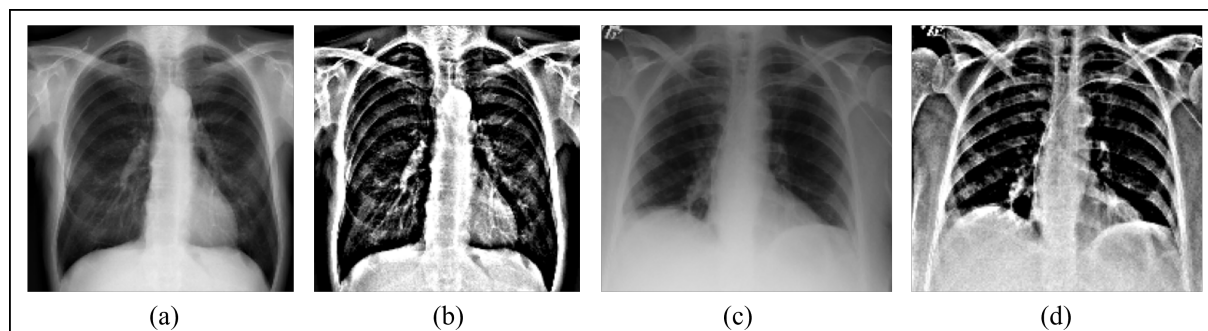


FIGURE 3.5 – Exemple d'images CXR de pneumonie : (a) et (c) images originales du premier et du second cluster respectivement, (b) et (d) images améliorées en utilisant le Masque Flou avec les paramètres optimaux déterminés pour chaque cluster.

3.5 Modèle de classification

3.5.1 Réseaux de neurones convolutifs

Les réseaux de neurones convolutifs (CNN) sont largement utilisés dans la classification d'images [105, 168, 202, 203]. Les CNN utilisent des filtres pour extraire des caractéristiques à différents niveaux de complexité, depuis les éléments les plus simples jusqu'aux motifs les plus complexes. Cette méthode de traitement hiérarchique des caractéristiques rend les CNN particulièrement adaptés pour des applications variées.

Dans notre étude, nous avons conçu un modèle CNN composé de cinq couches de convolution (Conv2D), ayant 64, 96, 128, 256 et 512 filtres respectivement, chaque filtre ayant une taille de 3×3 . Chaque couche de convolution est suivie d'une activation ReLU pour introduire de la non-linéarité. Pour réduire la dimensionnalité des cartes de caractéristiques, cinq couches de Max Pooling ont été intégrées, chacune utilisant un filtre 2×2 pour diviser par deux les dimensions spatiales des cartes de caractéristiques. Après les couches de convolution et de pooling, nous avons ajouté une couche de Flatten, suivie de deux couches denses, chacune contenant 500 neurones, et de deux couches de dropout pour réduire le risque de surajustement, avec un taux de dropout fixé à 0,5 pour ces couches denses. Enfin, une couche de sortie dense avec deux neurones utilisant la fonction d'activation sigmoïde classe les images en "pneumonie" ou "normale".

L'apprentissage du modèle a été réalisé par mini-lots de taille 16, en utilisant l'optimiseur ADAM [204] avec un taux d'apprentissage de 10^{-4} . Pour éviter le surapprentissage, nous avons intégré une stratégie d'arrêt précoce (*earlystopping*), fixée à une patience de

cinq époques. Cette méthode interrompt l'entraînement si aucune amélioration n'est observée sur la perte de validation durant le délai spécifié, ce qui optimise l'efficacité du processus d'entraînement. Pour la compilation du modèle, nous avons utilisé la fonction de perte 'binary_crossentropy', adaptée pour la classification binaire des résultats.

Pour une tâche de classification, nous extrayons les caractéristiques des images pour classer ces dernières en différentes catégories en fonction de la variation de ces caractéristiques. Cependant, toutes les caractéristiques n'ont pas le même poids d'importance, et seules celles qui sont pertinentes sont bénéfiques pour la classification. Pour mieux cibler ces caractéristiques significatives, nous intégrons un mécanisme d'attention dans notre modèle. Ce mécanisme permet de focaliser l'apprentissage sur les zones les plus informatives des images, augmentant ainsi l'efficacité de la classification. Nous explorerons ce mécanisme plus en détail dans la section suivante.

3.5.2 Mécanisme d'attention

Le mécanisme d'attention s'est avéré utile dans les tâches de vision par ordinateur, en particulier pour les applications complexes comme la classification des images médicales, où des détails fins peuvent être cruciaux pour un diagnostic précis. Le mécanisme d'attention permet de pondérer différemment les caractéristiques des images selon leur importance et leur pertinence [205], aidant les modèles à se focaliser sur des zones spécifiques plutôt qu'une analyse globale de l'image.

Face aux défis posés par les maladies pulmonaires, souvent localisées dans de petites régions spécifiques de l'image, nous avons développé un modèle CNN avec une couche d'attention spatiale. Cette combinaison capitalise sur l'efficacité des CNN pour capturer avec efficacité des motifs et détails des images, tandis que le mécanisme d'attention permet au modèle de se concentrer sur les caractéristiques et les régions les plus pertinentes pour la tâche de classification.

Comme illustré dans la Figure 3.6, notre modèle applique un mécanisme d'attention spatial sur les cartes de caractéristiques générées par la dernière couche de convolution du modèle CNN. Ce mécanisme d'attention évalue les cartes de caractéristiques et leur attribue des poids différents, en donnant des poids élevés aux régions jugées les plus pertinentes pour la classification. Ainsi, il permet au modèle de se concentrer sur les aspects les plus significatifs et informatives de l'image, négligeant les parties non pertinentes. Cette focalisation augmente considérablement l'efficacité du modèle et améliore la précision dans la classification des maladies pulmonaires.

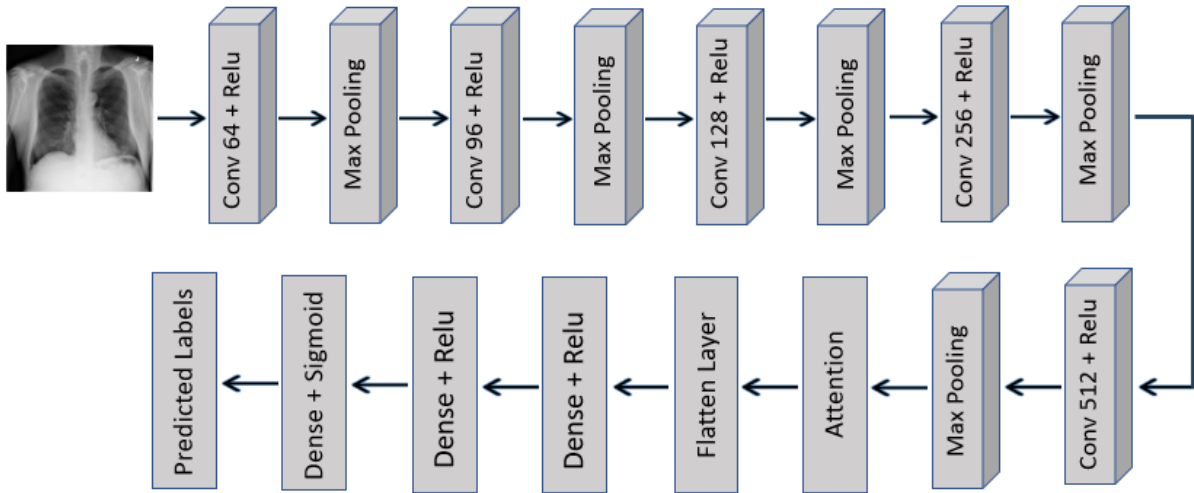


FIGURE 3.6 – Architecture de notre modèle CNN-Attention.

3.6 Résultats et discussion

L'objectif de cette étude est d'exploiter les données cliniques disponibles dans la base de données, plutôt que de se limiter uniquement à l'analyse des images, pour la classification des images normales et celles montrant une pneumonie. Nous avons adopté une approche qui identifie des groupes partageant des caractéristiques cliniques communes, ce qui permet d'ajuster le modèle en fonction des particularités de chaque sous-groupe, assurant ainsi une classification plus précise et adaptée.

Pour atteindre cet objectif, nous avons initialement appliqué la HCA aux caractéristiques extraites des images. Suite à l'analyse des caractéristiques cliniques des images dans chaque cluster, nous avons identifié deux clusters principaux : le premier composé majoritairement d'images prises en vue PA et le second principalement en vue AP. Une fois les groupes formés, les images ont été redimensionnées à une taille de 224×224 pixels et normalisées dans l'intervalle $[0; 1]$. Ensuite, nous avons procédé à l'amélioration du contraste des images de chaque cluster via la méthode UM, en appliquant des paramètres spécifiques à chaque groupe. L'efficacité de cette technique a été mesurée par la comparaison des performances de classification de la pneumonie sur les différents clusters, avec et sans l'application du UM. Pour garantir la stabilité des résultats, nous avons effectué dix partitions différentes des données en ensembles d'entraînement et de test et nous avons présenté les résultats en termes de moyenne et d'écart type. Les modèles ont été évalués sur les données de test pour déterminer leur performance.

Les résultats, présentés dans la Table 3.1, indiquent une amélioration significative des scores AUC lorsque l’UM est utilisé. Ce phénomène s’explique par le fait que le Masque Flou augmente le contraste, rendant ainsi les images plus nettes, ce qui aide à mieux distinguer les textures et les détails fins qui sont essentiels pour identifier les caractéristiques pathologiques dans les images de radiographie thoracique. En conséquence, nous avons adopté le Masque Flou comme méthode de prétraitement des images pour les analyses ultérieures de cette étude. Ensuite, nous avons procédé à la classification des images de chaque cluster en utilisant le modèle CNN et le modèle CNN intégrant un mécanisme d’attention. Les performances ont été évaluées par l’AUC pour chaque modèle et pour chaque cluster, ainsi que par une AUC pondérée sur les données de test. Sans l’utilisation de l’UM, les AUC pondérées obtenues étaient de 77,89% pour le modèle CNN et de 79,67% pour le modèle CNN avec attention. L’introduction de l’UM a permis d’améliorer ces résultats, avec une AUC de 80,11% pour le CNN et 81,76% pour le CNN avec attention. Le Tableau 3.1 démontre que l’utilisation de l’UM et du mécanisme d’attention avec le modèle CNN améliore les résultats de classification. Cette combinaison a permis d’atteindre la plus haute AUC pondérée de 81,76% pour la classification binaire des images normales et celles présentant une pneumonie, confirmant ainsi l’efficacité de notre approche.

TABLE 3.1 – AUC de chaque cluster et l’AUC pondérée avec et sans UM pour la classification binaire de la pneumonie.

AUC (%)	Without UM		With UM	
	CNN	CNN + Attention	CNN	CNN + Attention
Cluster 1	78.14 \pm 1.4	79.85 \pm 1.5	80.67 \pm 1.4	81.95 \pm 1.4
Cluster 2	77.56 \pm 1.5	79.43 \pm 1.6	79.39 \pm 1.2	81.52 \pm 1.5
Weighted AUC	77.89 \pm 1.4	79.67 \pm 1.5	80.11 \pm 1.3	81.76 \pm 1.4

Globalement, le modèle CNN-Attention représente une avancée significative pour aborder le défi posé par les maladies pulmonaires, souvent localisées dans des petites zones spécifiques à la maladie. Ce modèle hybride combine efficacement les capacités du CNN avec celles du mécanisme d’attention, optimisant ainsi l’analyse d’image. Au lieu de traiter toute l’image uniformément, le mécanisme d’attention ajuste les poids, mettant en évidence les zones les plus cruciales pour le diagnostic des maladies. Ainsi, le modèle est particulièrement efficace pour identifier et classifier les maladies pulmonaires, même lorsqu’elles se manifestent dans de petites zones de l’image radiographique. Cette approche ciblée rend le diagnostic à la fois plus précis et plus fiable.

En outre, l’application du modèle sur des ensembles de données spécifiquement regroupées par caractéristiques techniques, telles que la position de vue, a permis d’améliorer significativement les performances du modèle. Comme indiqué dans la Table 3.2, lorsqu’il est appliqué à toutes les images sans le clustering, le modèle CNN avec mécanisme d’attention a obtenu un score AUC de 79,41% pour la classification de la pneumonie. Cependant, après avoir regroupé les images en fonction de leur position de vue et réappliqué le modèle, l’AUC a augmenté pour atteindre 81,76%. Cette amélioration de 2,35% démontre clairement que la position de vue des images a un impact significatif sur la performance de classification, en optimisant la capacité du modèle à identifier les caractéristiques pertinentes des pathologies pulmonaires.

TABLE 3.2 – AUC de la classification binaire des trois maladies avec et sans l’étape de clustering.

Disease	Unsharp Masking + CNN-Attention	
	With Clustering	Without Clustering
Pneumonia	81.76% \pm 1.4%	79.41% \pm 1.5%
Consolidation	82.87% \pm 1.3%	80.35% \pm 1.5%
Infiltration	77.24% \pm 1.5%	75.68% \pm 1.4%

Il est donc essentiel de prendre en compte la position de vue lors du diagnostic des maladies pulmonaires à partir d’images radiographiques. Dans la pratique clinique, la position de vue des radiographies est cruciale car elle affecte directement la visibilité et l’interprétation des structures anatomiques et des pathologies potentielles. Les vues PA et AP peuvent modifier significativement l’apparence de la cavité thoracique, des organes et de toutes anomalies présentes. Les vues PA offrent une image plus claire et plus fidèle des poumons et du thorax, minimisant l’agrandissement du cœur et des autres structures du médiastin, ce qui facilite une évaluation plus précise des tissus pulmonaires. À l’inverse, les vues AP, souvent utilisées pour les patients qui ne peuvent pas être positionnés en vue PA notamment en raison de leur état de santé, peuvent modifier la perception des anomalies pulmonaires, masquant ou exagérant certaines pathologies pulmonaires, et peuvent compliquer la détection de détails subtils ou l’évaluation exacte de la taille et de la forme des lésions pulmonaires. Prendre en compte la position de vue lors de l’évaluation des images radiographiques est donc essentiel pour assurer une interprétation diagnostique précise et éviter les erreurs potentielles, ce qui optimise la qualité des diagnostics, facilite la prise des décisions thérapeutiques appropriées et améliore l’exactitude du traitement administré aux patients.

Nous avons adopté cette approche pour la classification binaire de deux autres pathologies pulmonaires de la base de données ChestX-ray14 : la consolidation et no finding (2 751 images), ainsi que l'infiltration et no finding (20 049 images). Pour cela, nous avons sélectionné les 1 310 images étiquetées comme présentant une consolidation et les 9 547 images comme présentant une infiltration. Pour équilibrer les classes, nous avons sous-échantillonné les images no finding à 1 441 pour la classification binaire de la consolidation et à 10 502 pour classification binaire de l'infiltration. Après extraction des caractéristiques à l'aide du CNN, nous avons utilisé la HCA pour regrouper les images en deux clusters basés aussi sur la position de vue : l'un majoritairement en vue PA et l'autre en vue AP. Nous avons ensuite amélioré le contraste des images de chaque cluster en utilisant la technique de Masque Flou, avec des paramètres ajustés spécifiquement pour chaque groupe. Ces images prétraitées ont été soumises à la classification à l'aide de deux modèles : le CNN classique et le CNN intégrant un mécanisme d'attention, afin de comparer leur efficacité dans la discrimination des pathologies mentionnées.

Ainsi, après avoir appliqué notre méthodologie développée initialement pour la classification de la pneumonie, nous avons également observé une amélioration significative des performances pour la classification binaire de la consolidation et de l'infiltration en utilisant le clustering, le Masque Flou, et le modèle CNN basé sur l'attention, atteignant un score AUC pondéré de 82,87% pour la consolidation et de 77,24% pour l'infiltration. Les résultats détaillés de la classification binaire de chaque pathologie sont présentés dans la Table 3.2, illustrant l'efficacité de notre approche. Sans le clustering, l'application du modèle CNN-Attention a produit une AUC de 80,35% pour la consolidation et de 75,68% pour l'infiltration. Cependant, après le clustering en fonction de la position de vue, les performances ont été améliorées, atteignant respectivement 82,87% et 77,24% pour la consolidation et l'infiltration.

Nous avons comparé notre approche avec d'autres travaux existants qui utilisent la même base de données. Comme présenté dans la Table 3.3, notre approche dépasse les scores AUC rapportés dans la littérature pour la classification binaire des trois pathologies pulmonaires étudiées. Spécifiquement, notre modèle a démontré une amélioration de 4,06%, 0,47% et 4,34% dans la classification de la pneumonie, de la consolidation, et de l'infiltration, respectivement. Cette amélioration est due à notre stratégie de regroupement des données en clusters selon des caractéristiques techniques dominantes telles que la position de vue, combinée à l'amélioration du contraste des images et à l'application du mécanisme d'attention pour concentrer le modèle sur les zones les plus informatives. Cette

stratégie améliore non seulement la précision du diagnostic, mais également la pertinence clinique des résultats, permettant des décisions thérapeutiques plus précises et adaptées.

TABLE 3.3 – Comparaison des résultats obtenus avec d’autres méthodes existantes dans la littérature pour la classification binaire des trois maladies pulmonaires de la base de données ChestX-ray14.

Reference	Year	Pneumonia	Consolidation	Infiltration
Wang et al. [47]	2017	66.4%	77%	63.6%
Yao et al. [194]	2017	68.4%	74.1%	70%
Rajpurkar et al. [39]	2017	76.8%	74.1%	70%
Li et al. [180]	2018	67%	80%	70%
Ma et al. [184]	2019	72.2%	75%	69.6%
Xu et al. [124]	2021	77.7%	81.7%	71.5%
Souid et al. [186]	2021	73.3%	79%	71.1%
Yang et al. [154]	2022	73.4%	75.6%	72.2%
Jin et al. [7]	2023	73.9%	72.5%	72.4%
Mann et al. [189]	2023	74.3%	82.4%	72.9%
Our proposed approach	2024	81.76% ±1.4%	82.87% ±1.3%	77.24% ±1.5%

La visualisation de l’attention offre une perspective précieuse tant pour les chercheurs que pour les médecins, permettant d’identifier les régions spécifiques de l’image sur lesquelles le modèle se concentre lors de la classification des pathologies pulmonaires. En superposant les cartes d’attention générées par le modèle aux images radiographiques originales, nous pouvons déterminer les zones jugées pertinentes par le modèle pour le diagnostic. La Figure 3.7 illustre des exemples où notre modèle a localisé des anomalies associées à des pathologies telles que la pneumonie, la consolidation et l’infiltration. Cette capacité de localisation et de visualisation améliore non seulement l’explicabilité et l’interprétabilité des décisions du modèle mais contribue également à la transparence et à la confiance des médecins dans l’application clinique des diagnostics assistés par IA.

Pour valider nos résultats, nous avons présenté ces images à un médecin qui a confirmé que les zones ciblées par notre modèle coïncidaient avec les zones cliniquement significatives affectées par les maladies. Cette validation clinique par un professionnel de la santé confirme l’efficacité et la fiabilité de notre approche, soulignant son potentiel en tant qu’outil de support diagnostique efficace en radiologie thoracique.

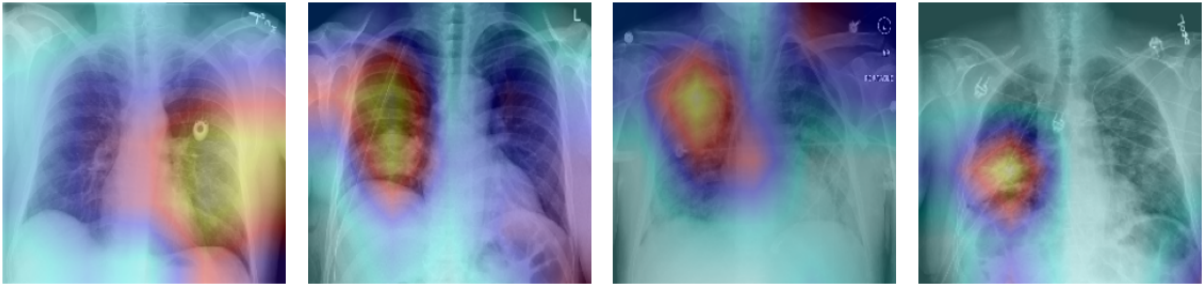


FIGURE 3.7 – Exemples de visualisation de la localisation des maladies pulmonaires sur les images de test de pneumonie, de consolidation et d’infiltration. Une réponse plus élevée est indiquée en jaune.

3.7 Conclusion

Dans ce chapitre, nous avons présenté une approche pour considérer les données cliniques lors de la classification des maladies pulmonaires en utilisant la base de données ChestX-ray14. Notre analyse a combiné l’examen des images radiographiques avec les données cliniques pour identifier des groupes homogènes basés sur la position de vue dominante. Nous avons appliqué la technique de Masque Flou pour améliorer le contraste des images dans chaque groupe. De plus, nous avons renforcé notre modèle CNN avec un mécanisme d’attention pour se focaliser sur les régions cruciales de l’image, afin d’assurer une classification précise des pathologies pulmonaires. L’efficacité de notre approche a été mesurée par l’AUC pondérée, qui a surpassé les performances rapportées dans les études antérieures de 4,06%, 0,47%, et 4,34% pour la classification binaire de la pneumonie, de la consolidation et de l’infiltration, respectivement.

L’application de notre modèle sur des groupes partageant des caractéristiques cliniques communes s’avère cruciale pour maximiser sa pertinence et son efficacité. Notre étude démontre que la prise en compte de la position de vue dans l’analyse radiographique des maladies pulmonaires améliore significativement la performance du diagnostic. Cette amélioration permet aux radiologues de fournir des diagnostics plus précis et de prendre des décisions thérapeutiques mieux adaptées pour les patients. En outre, l’intégration d’une couche d’attention après la dernière couche de convolution de notre modèle CNN améliore sa capacité à se concentrer sur les zones les plus pertinentes pour la classification des maladies pulmonaires. La visualisation de ces régions d’attention fournit une compréhension du fonctionnement interne du modèle, soulignant les zones qu’il juge cruciales pour la classification des maladies. Cette fonctionnalité explicative aide les professionnels de la santé à valider et à comprendre les prédictions du modèle, renforçant ainsi leur confiance

dans l'utilisation de cette technologie avancée pour le diagnostic médical.

Ce chapitre a fait l'objet d'un article soumis à la revue "*IEEE Transactions on Bio-medical Engineering*".

Cependant, les images radiographiques peuvent contenir des artefacts, comme des objets, des dispositifs médicaux, des fils et des électrodes, qui introduisent du bruit et peuvent obscurcir des structures anatomiques pertinentes, compliquant ainsi l'identification précise des pathologies et rendant le diagnostic plus difficile et moins précis. Dans le chapitre suivant, nous aborderons le problème de ces artefacts afin d'améliorer les résultats de classification des maladies pulmonaires.

AMÉLIORATION DE LA CLASSIFICATION DES MALADIES PULMONAIRES : UNE NOUVELLE APPROCHE DE PRÉTRAITEMENT BASÉE SUR CYCLEGAN POUR LA RÉDUCTION DES ARTEFACTS

4.1 Introduction

La radiographie thoracique est une technique d'imagerie médicale largement utilisée pour le diagnostic des maladies thoraciques. Cependant, comme illustré dans la Figure 4.1, ces images peuvent souvent contenir des artefacts tels que des objets étrangers, des dispositifs médicaux, des fils, et des électrodes, qui peuvent altérer la qualité de l'image. Ces éléments non pertinents introduisent du bruit qui peut compromettre la visibilité des structures anatomiques essentielles et rendre difficile la distinction des caractéristiques pathologiques pertinentes.

Le bruit introduit par ces artefacts peut fausser les interprétations radiologiques en masquant ou en simulant des anomalies, ce qui peut conduire à des erreurs de diagnostic. Par exemple, un fil d'électrode pourrait être mal interprété comme une anomalie dans le poumon, menant potentiellement à un diagnostic erroné de maladie pulmonaire. Ces erreurs de diagnostic peuvent mener à des traitements inappropriés ou au manque de traitement nécessaire, mettant en danger la santé du patient. Il est donc impératif d'adopter des techniques de prétraitement qui minimisent l'impact de ces artefacts sur les images

radiographiques, permettant ainsi une meilleure interprétation et analyse par les modèles d'apprentissage automatique. En réduisant ou en éliminant l'impact de ces éléments non pertinents, on peut non seulement améliorer la qualité des diagnostics mais également renforcer l'efficacité des décisions thérapeutiques, optimisant ainsi la prise en charge des patients.

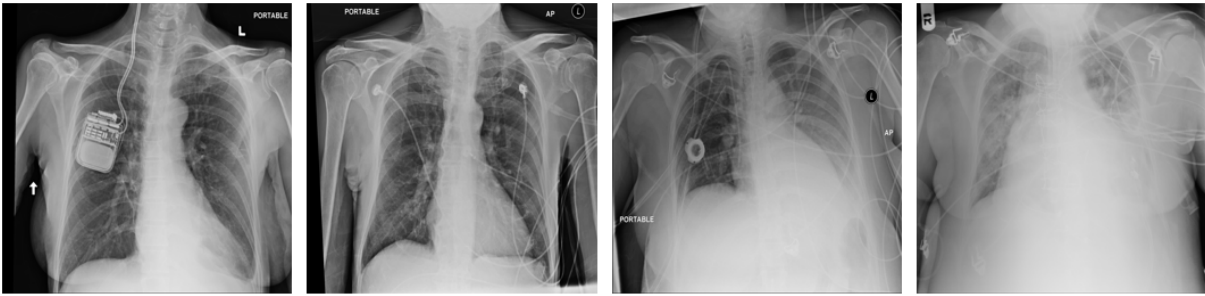


FIGURE 4.1 – Exemples d'images de radiographie thoracique présentant des artefacts dans la base de données ChestX-ray14.

Il est important de souligner que le problème de manque d'étiquettes pour les images contenant des composants électroniques peut se présenter dans toute base de données. Ceci met en évidence l'utilité du clustering (étiquetage non supervisé) pour le traitement efficace des données d'image, surtout dans le contexte médical.

L'objectif de ce chapitre est de traiter le problème posé par la présence d'artefacts dans les images de radiographie thoracique. Nous avons développé une nouvelle méthode de prétraitement destinée à distinguer les images nettes de celles affectées par des artefacts, et à minimiser l'effet du bruit produit par ces derniers en générant des images améliorées grâce à l'utilisation d'un modèle CycleGAN. Cette amélioration de la qualité des données conduit à des résultats plus précis et significatifs dans le processus de classification des images médicales, ce qui renforce la fiabilité et l'efficacité des analyses diagnostiques dans le domaine de la santé.

L'impact de notre approche est évalué en fonction de sa capacité à améliorer les performances de classification des maladies pulmonaires, ce qui pourrait potentiellement optimiser les diagnostics médicaux. Le modèle de classification que nous avons développé est basé sur DenseNet et inclut des modules d'attention spatiale et par canal, ce qui lui permet de maximiser l'utilisation des aspects pertinents de l'image. En pratique clinique, les radiologues ne s'appuient pas uniquement aux images pour poser un diagnostic mais utilisent également des informations cliniques complémentaires. C'est pourquoi nous

avons intégré les caractéristiques cliniques contenues dans la base de données dans notre modèle. Cette intégration enrichit la précision de la classification en combinant l'analyse des données visuelles et le contexte clinique des patients, ce qui est essentiel pour des diagnostics médicaux fiables et précis.

Notre étude utilise la base de données ChestX-ray14 [47], qui comprend 322 images étiquetées comme présentant une pneumonie et 60 361 images classées comme sans pathologie (no finding). Afin de gérer le déséquilibre significatif entre les classes, nous avons sous-échantillonné les images no finding pour en avoir 354 images. Cela nous a permis de constituer un ensemble de données plus équilibré de 676 images, sur lequel nous avons basé nos expériences, tout en prenant en compte les données cliniques correspondantes pour chaque image.

4.2 Approche proposée

Notre méthodologie, comme illustrée dans la Figure 4.2, comprend quatre étapes principales : le prétraitement, le clustering (étiquetage non supervisé), l'application du CycleGAN et la classification.

Dans la phase de prétraitement, nous avons utilisé la technique d'égalisation d'histogramme pour améliorer le contraste et la visibilité des détails des images. Ensuite, nous avons extrait des caractéristiques liées aux contours et à la texture des images pour réaliser le clustering en utilisant l'algorithme K-means, aboutissant à la création de deux clusters : le premier contenant des images avec des artefacts, et le second contenant des images nettes sans artefacts. Pour réduire l'effet de bruit causé par les artefacts présents dans le premier cluster, nous avons utilisé un modèle CycleGAN qui a transformé ces images en versions plus nettes et sans artefacts. Ces images générées ont ensuite été combinées avec celles du second cluster pour effectuer la classification. Les images ont été redimensionnées à 224×224 pixels et normalisées dans un intervalle $[0 ; 1]$. Pour maximiser l'utilisation des informations disponibles dans la base de données, nous avons ajouté les caractéristiques cliniques telles que l'âge et le genre du patient, ainsi que la position de vue de l'image CXR.

Dans notre étude, 20% des données ont été utilisées pour test, 10% pour la validation, et 70% ont servi à l'entraînement (choisies aléatoirement). La classification a été effectuée en utilisant le modèle DenseNet-121 enrichi par un Module d'Attention de Bloc Convolutionnel (CBAM) qui permet de maximiser l'utilisation des aspects pertinents de l'image

en renforçant sélectivement les caractéristiques dans les données d’entrée qui sont fortement liées aux zones pathologiques. L’efficacité du modèle a été évaluée à l’aide de l’AUC, une métrique prédominante dans la littérature pour évaluer la performance des modèles de classification des maladies pulmonaires, ce qui justifie son choix dans le contexte de notre étude. Nous avons également calculé d’autres métriques d’évaluation telles que la précision, le rappel et le F1-score pour valider la robustesse et la fiabilité de notre approche.

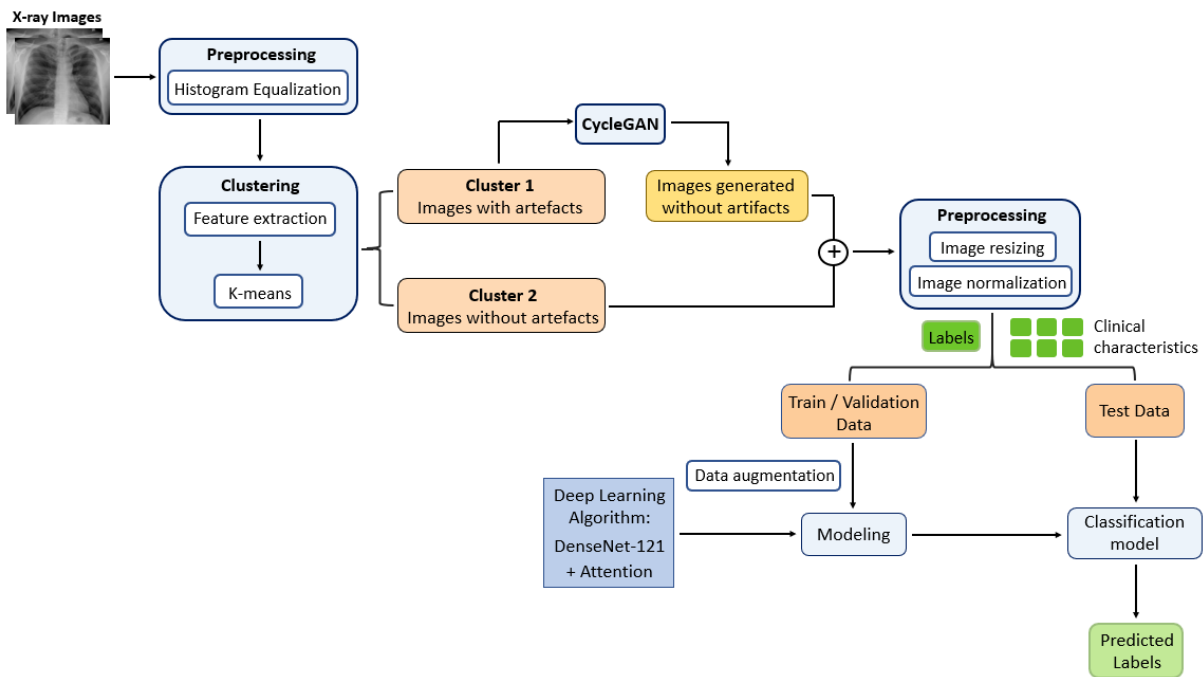


FIGURE 4.2 – Aperçu de la méthodologie proposée.

4.3 Prétraitement des images avec l’égalisation d’histogramme

Le prétraitement des images est une étape cruciale pour améliorer la qualité des images radiographiques, optimisant ainsi la détection des caractéristiques essentielles par les algorithmes d’apprentissage. Dans cette étude, nous avons utilisé l’égalisation d’histogramme pour améliorer le contraste de chaque image.

L'égalisation d'histogramme ou *Histogram equalization* (HE) est une technique de traitement d'image qui vise à répartir uniformément les niveaux de gris sur toute l'échelle disponible. Cela augmente le contraste global, particulièrement utile dans les zones sombres et peu contrastées de l'image. Par cette méthode, les nuances de gris sont redistribuées, ce qui améliore significativement la clarté et les détails des structures anatomiques dans l'image radiographique, affine les contours, et rehausse la qualité visuelle générale des images, facilitant ainsi une meilleure interprétation et analyse [69].

L'histogramme d'une image numérique, avec des niveaux d'intensité allant de 0 à $(L - 1)$ où L est le nombre total de niveaux de gris, est une fonction représentée comme suit :

$$h(r_k) = n_k \quad (4.1)$$

où r_k est le k -ième niveau d'intensité et n_k est le nombre de pixels dans l'image ayant ce niveau d'intensité r_k .

L'histogramme est généralement normalisé en divisant chaque pixel par le nombre total de pixels dans l'image. Pour une image de taille $M \times N$, la fonction d'histogramme normalisée est donnée par l'équation 4.2 :

$$p(r_k) = \frac{n_k}{M \times N} \quad (4.2)$$

Cette fonction $p(r_k)$ est la probabilité de l'occurrence d'un pixel avec le niveau d'intensité r_k . Il est clair que $\sum p(r_k) = 1$.

Ensuite, la fonction de densité cumulative est calculée à partir de l'histogramme [206]. Cette fonction, donnée par l'équation 4.3, représente la probabilité cumulative de chaque valeur d'intensité dans l'image.

$$c(r_k) = \sum_{i=0}^k p(r_i) \quad (4.3)$$

pour $k = 0, 1, \dots, L - 1$ et où par définition $c(r_{L-1}) = 1$.

Enfin, la fonction de densité cumulative est utilisée pour calculer de nouvelles valeurs d'intensité pour chaque pixel de l'image d'origine, afin de répartir uniformément les intensités. En effet, l'égalisation d'histogramme est un schéma qui mappe l'image d'entrée sur toute la plage dynamique, (X_0, X_{L-1}) , en utilisant la fonction de densité cumulative comme fonction de transformation. La fonction de transformation $f(x)$ basée sur la

fonction de densité cumulative est défini comme suit :

$$f(r_k) = r_0 + (r_{l-1} - r_0)c(r_k) \quad (4.4)$$

Alors, l'image de sortie produite par l'égalisation d'histogramme, $Y = Y(i, j)$, peut être exprimée comme suit :

$$Y = \{f(X(i, j)) \mid \forall X(i, j) \in X\} \quad (4.5)$$

Par conséquent, l'égalisation d'histogramme redistribue les valeurs d'intensité d'une image pour exploiter pleinement la plage dynamique disponible, entraînant ainsi une amélioration du contraste et de la qualité visuelle de l'image. La Figure 4.3 démontre clairement les améliorations apportées à des images de radiographie thoracique de la base de données ChestX-ray14 suite à l'application de l'égalisation d'histogramme.

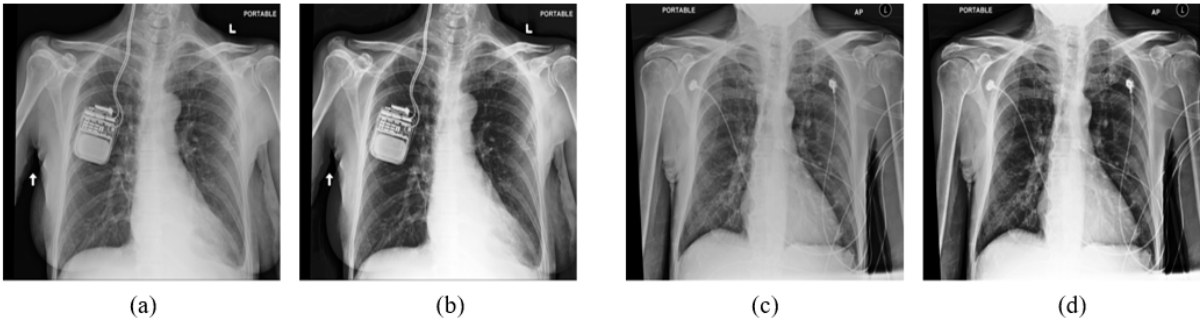


FIGURE 4.3 – Exemple d'images CXR : (a) et (c) images originales, (b) et (d) images améliorées à l'aide de la technique d'égalisation d'histogramme.

4.4 Séparation des images avec et sans artefacts à l'aide du clustering

Après l'amélioration du contraste des images, pour gérer la présence non étiquetée des composants électroniques qui peuvent introduire du bruit dans les résultats de la classification, nous avons développé une stratégie de prétraitement afin de traiter ce problème. Cette stratégie commence par séparer les images en deux clusters : un cluster pour les images contenant des artefacts et un autre pour celles sans artefacts. Pour cela, nous avons

extrait des caractéristiques pertinentes des images, liées à la texture et aux contours, et ensuite nous avons utilisé l'algorithme de clustering K-means pour séparer les images en deux clusters.

4.4.1 Extraction des caractéristiques

L'extraction de caractéristiques, en particulier celles liées aux contours et à la texture des images radiographiques, est cruciale pour interpréter et comprendre la topologie des informations contenues dans les images. En identifiant les détails subtils de contours et de la texture, cela permet à l'algorithme K-means de séparer efficacement les images nettes des images affectées par des artefacts. Dans le cadre de cette étude, nous avons extrait quatre caractéristiques principales, formant un vecteur de 580 648 caractéristiques par image, incluant :

1. **Caractéristique de la variance des gradients** : Cette caractéristique calcule les dérivées partielles horizontales et verticales de l'image à l'aide du filtre de Sobel. Ces dérivées partielles permettent de quantifier les variations d'intensité des pixels dans l'image et fournissent des informations sur la manière dont la luminosité change dans différentes directions de l'image. En combinant ces dérivées, on calcule la magnitude du gradient à chaque point de l'image. La variance de ces valeurs de magnitude à travers l'image permet ensuite de mesurer les variations de texture et évaluer l'uniformité de la texture dans l'image. Ainsi, les régions avec des textures variées ou complexes tendront à présenter une variance élevée des gradients, indiquant une grande variation de l'intensité des pixels, tandis que les zones plus uniformes présenteront une variance réduite. Ainsi, cette caractéristique fournit une indication sur les variations d'intensité et de texture dans une image, permettant d'identifier les images avec des différences de texture significatives.
2. **Caractéristique de la netteté pulmonaire** : Cette caractéristique quantifie la netteté des contours des poumons dans une image radiographique. Elle utilise le filtre de Sobel pour calculer les dérivées partielles horizontales et verticales, qui permettent de détecter les variations d'intensité des pixels. En combinant ces dérivées, on obtient la magnitude du gradient, qui mesure directement la netteté des contours : une magnitude élevée indique des contours nets et bien définis. Pour concentrer cette mesure sur les poumons, un masque est créé pour isoler cette région. Ce masque élimine les informations inutiles des autres zones de l'image à

l'extérieur des poumons, garantissant que seuls les contours de la zone pulmonaire sont pris en compte dans le calcul de la netteté. Ainsi, une valeur élevée de cette caractéristique indique des contours pulmonaires nets et bien définis, tandis qu'une valeur faible peut indiquer la présence d'artefacts comme un mauvais réglage de l'imagerie ou d'autres perturbations affectant la clarté de l'image.

3. **Caractéristique du ratio noir-blanc** : Cette caractéristique évalue la distribution des pixels noirs et blancs dans une image fournissant un indicateur quantitatif de la densité des structures. Elle est calculée en mesurant le pourcentage de pixels noirs qui représentent des zones de faible intensité, et de pixels blancs qui correspondent à des zones de haute intensité. Cette mesure est particulièrement utile dans l'analyse des radiographies thoraciques, car elle permet de distinguer les zones pulmonaires généralement plus sombres, des autres zones anatomiques souvent plus claires. Un ratio élevé de pixels noirs peut indiquer une densité tissulaire élevée, tandis qu'un ratio élevé de pixels blancs peut signaler la présence d'artefacts, une surexposition ou d'autres objets non pertinents qui affectent la répartition des intensités dans l'image.
4. **Caractéristique des Histogrammes de Gradients Orientés (HOG)** : Les caractéristiques HOG (*Histogram of Oriented Gradients*) sont conçues pour capturer les motifs de gradient dans une image, offrant une description robuste des formes et des structures. Le processus débute par la division de l'image en petites régions appelées cellules. Pour chaque cellule, les gradients sont calculés, ce qui permet de déterminer les variations de luminosité [207]. Ces informations sont ensuite compilées en histogrammes des gradients orientés pour chaque cellule. Le processus se poursuit par la normalisation de ces histogrammes dans des blocs plus grands. Enfin, les histogrammes de toutes les cellules sont concaténés pour former le descripteur HOG final de l'image. Ce descripteur intègre des informations détaillées sur la texture et l'orientation des contours, ce qui le rend particulièrement utile pour des applications telles que la détection d'objets et la reconnaissance de motifs [208, 209], particulièrement dans les images de radiographie thoracique [210].

4.4.2 K-means clustering

Après l'extraction des caractéristiques de contours et de la texture des images CXR, celles-ci ont été normalisées et utilisées comme données d'entrée pour l'algorithme d'ap-

prentissage automatique non supervisé, le K-means. L'objectif de cet algorithme est de grouper les données en maximisant la similarité des données dans chaque groupe (c.à.d. en minimisant la variance intra-cluster) tout en maximisant les différences entre les groupes existants (c.à.d. en maximisant la variance inter-cluster), sans nécessiter une supervision humaine directe [211, 212].

Le processus du K-means débute par la sélection aléatoire des points de données comme centres initiaux ou centroides des clusters. Les données sont ensuite attribuées au cluster dont le centroïde est le plus proche, basé sur la distance euclidienne, formant ainsi des groupes. Au cours des itérations successives, les centroides sont recalculés comme points moyens de leur classe et les données sont réaffectées selon ces nouveaux centroides. Ce processus se répète jusqu'à ce que la position des centroides se stabilise, signalant une convergence de l'algorithme. Finalement, ce processus itératif de K-means aboutit à des clusters bien définis où chaque cluster contient des images qui sont similaires entre elles mais significativement distinctes des images des autres clusters.

Dans notre contexte, cette méthode d'apprentissage non supervisée permet de révéler des structures cachées dans notre base de données, permettant ainsi la séparation efficace des images nettes de celles avec des artefacts. Dans notre étude, le K-means est appliqué à la fois aux images labélisées comme no finding et à celles diagnostiquées comme pneumonie, aboutissant à la formation de deux clusters pour chaque classe. En effet, le K-means a efficacement regroupé les images selon leurs caractéristiques extraites, formant ainsi deux clusters distincts : un cluster comprenant des images jugées nettes, tandis que l'autre cluster contenant des images avec des artefacts et des objets indésirables. La Table 4.1 présente la distribution des images appartenant aux classes de pneumonie et de no finding dans chaque cluster après l'application de l'algorithme de clustering K-means.

TABLE 4.1 – Nombre d'images des classes de pneumonie et de no finding dans chaque cluster obtenues à l'aide du clustering par K-means.

	Pneumonia	No finding	Total
Cluster 1 : Images with artifacts	152	133	285
Cluster 2 : Images without artifacts	170	221	391
Total number of images in both clusters	322	354	676

Pour évaluer notre approche de clustering non supervisé, nous avons procédé à l'éti-

quetage manuel des images de pneumonie et de no finding en les classant comme contenant ou non des artefacts. Ces étiquettes ont servi de référence pour évaluer l'efficacité de notre approche de clustering avec K-means et qui est basée sur les caractéristiques de texture extraites des images. Nous avons évalué les performances de notre approche en calculant la matrice de confusion pour chaque classe, illustrée dans la Figure 4.4. Les métriques de performance détaillées sont également présentées dans la Table 4.2.

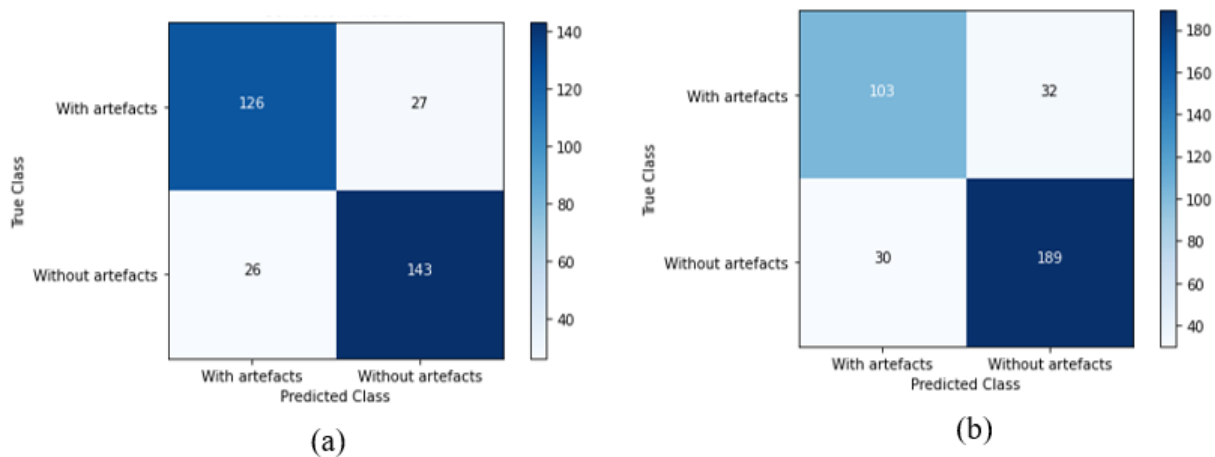


FIGURE 4.4 – Matrice de confusion de notre approche de clustering pour : (a) la classe pneumonie et (b) la classe no finding.

TABLE 4.2 – Métriques de performance de notre approche de clustering pour les classes pneumonie et no finding.

	Pneumonia	No finding
Accuracy	83.54%	82.49%
Recall	84.62%	86.3%
Precision	84.12%	85.52%
F1-score	84.37%	85.91%

D'après la Table 4.2, pour le clustering des cas de no finding, nous avons obtenu une accuracy de 82,49% et un F1-score de 85,91%. Concernant les cas de pneumonie, nous avons atteint une accuracy de 83,54% et un F1-score de 84,37%. L'analyse des matrices de confusion (a) et (b) de la Figure 4.4 a révélé que seulement 53 sur 322 images et 62 sur 354 images ont été incorrectement classées pour chacune des classes de pneumonie et de no finding, respectivement.

Ces résultats démontrent l'efficacité de notre approche de clustering (étiquetage non supervisé) pour distinguer les images nettes de celles contenant des artefacts. Ainsi notre

approche constitue une alternative plus rapide et moins laborieuse que l'étiquetage manuel traditionnel effectué par des professionnels de la santé, offrant la possibilité de traitement d'un grand volume d'images avec efficacité.

Il est essentiel de noter que les clusters obtenus ne peuvent pas être directement combinés en raison d'un biais potentiel qui pourrait affecter les résultats de classification ultérieurs. Pour résoudre ce problème, nous avons adopté une stratégie d'harmonisation des données. Pour cela, dans notre étude, nous avons utilisé le modèle CycleGAN, reconnu pour sa capacité à transférer des informations d'un centre de référence à un centre à harmoniser, c.à.d. entre différents domaines. Cette technique permet d'harmoniser les images des deux clusters, en adaptant les caractéristiques de l'un pour qu'elles ressemblent davantage à celles de l'autre, réduisant ainsi l'impact du bruit causé par les artefacts. L'application de cette méthode d'harmonisation, qui sera détaillée dans la section suivante, permet de minimiser les effets négatifs des artefacts sur les images, améliorant ainsi la qualité des données et, par conséquent, la fiabilité et la précision des résultats de classification.

4.5 Réduction des artefacts avec CycleGAN

Développé par Zhu et al. en 2017 [213], le réseau antagoniste cohérent de cycle ou *Cycle-Consistent Generative Adversarial Network*, connu sous le nom de "CycleGAN", est une variante spécifique des réseaux antagonistes génératifs (GANs). CycleGAN est spécifiquement conçu pour les tâches de conversion d'image à image non appariées où les paires de données d'entraînement correspondantes ne sont pas disponibles [214].

Comme le montre la Figure 4.5, CycleGAN se caractérise par l'utilisation de deux générateurs et de deux discriminateurs. Le générateur vise à produire des images synthétiques, tandis que le discriminateur évalue si ces images générées ont une apparence réaliste, poussant le générateur à améliorer la qualité des images générées jusqu'à ce qu'elles soient indistinguables des véritables échantillons. De plus, CycleGAN intègre le concept de "cohérence cyclique", ou *cycle consistency*, qui permet la conversion d'images entre deux domaines sans nécessiter de données appariées. Ce principe assure qu'une image transformée d'un domaine A à un autre domaine B puis retransformée au domaine A retrouve sa forme originale [215].

Par conséquent, le CycleGAN repose sur deux types principaux de fonctions de perte : la perte de cohérence cyclique et la perte adversaire.

Pour les générateurs G (qui convertit l'image du domaine A au domaine B) et F (du

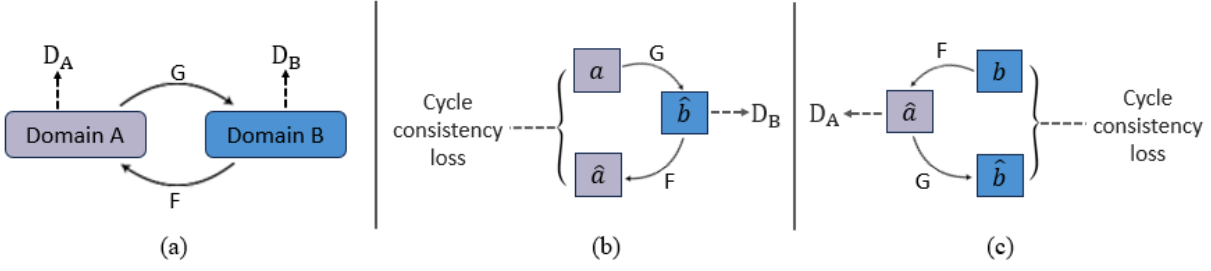


FIGURE 4.5 – (a) Le modèle CycleGAN contient deux fonctions de transformation, $G : A \rightarrow B$ et $F : B \rightarrow A$, avec leurs discriminateurs respectifs D_B et D_A . Le discriminateur D_B vise à encourager G à transformer les images du domaine A en images qui semblent naturellement appartenir au domaine B , et inversement pour F avec D_A . (b) Perte de cohérence cyclique directe : pour chaque image du domaine A , notée a , elle est transformée par G en \hat{b} , qui est ensuite retransformée par F en \hat{a} , avec pour objectif que \hat{a} soit très proche de a (i.e., $a \rightarrow G(a) = \hat{b} \rightarrow F(\hat{b}) = \hat{a} \approx a$). (c) Perte de cohérence cyclique inverse : pour chaque image du domaine B , notée b , elle est transformée par F en \hat{a} , qui est ensuite retransformée par G en \hat{b} , avec pour objectif que \hat{b} soit très proche de b (i.e., $b \rightarrow F(b) = \hat{a} \rightarrow G(\hat{a}) = \hat{b} \approx b$).

domaine B au domaine A), la perte de cohérence cyclique est définie par l'équation 4.6 :

$$\mathcal{L}_{\text{cycle}}(G, F) = E_{a \sim p_{\text{data}}(a)} [\|F(G(a)) - a\|_1] + E_{b \sim p_{\text{data}}(b)} [\|G(F(b)) - b\|_1] \quad (4.6)$$

où P_{data} est la distribution des données réelles.

La perte adversaire évalue la capacité des générateurs à tromper leurs discriminateurs respectifs en produisant des images qui ne peuvent pas être distinguées des images réelles du domaine cible. Elle utilise la fonction de perte des moindres carrés pour évaluer la différence entre les sorties attendues et prédites. Les pertes adversaires sont définies dans les équations 4.7 et 4.8 :

$$\mathcal{L}_{\text{GAN}}(G, D_B, A, B) = E_{b \sim P_{\text{data}}(b)} [\log D_B(b)] + E_{a \sim P_{\text{data}}(a)} [\log(1 - D_B(G(a)))] \quad (4.7)$$

$$\mathcal{L}_{\text{GAN}}(F, D_A, B, A) = E_{a \sim P_{\text{data}}(a)} [\log D_A(a)] + E_{b \sim P_{\text{data}}(b)} [\log(1 - D_A(F(b)))] \quad (4.8)$$

En plus, CycleGAN intègre le concept de perte d'identité. Ce concept repose sur le principe que si un générateur $G : A \rightarrow B$ est conçu pour traduire des images du domaine A au domaine B , alors, logiquement, lorsque ce générateur reçoit une image appartenant déjà au domaine B , il devrait renvoyer une sortie très similaire à l'image d'entrée, préservant ainsi son identité ($a \rightarrow F(a) \approx a$, and $b \rightarrow G(b) \approx b$). Cette perte d'identité est calculée en mesurant l'écart entre l'image d'entrée et l'image générée lorsqu'une image du

domaine cible est passée à travers son propre générateur, renforçant la capacité du modèle à maintenir les caractéristiques originales des images dans des situations où aucune transformation entre les domaines n'est nécessaire. La perte d'identité est définie dans l'équation 4.9 :

$$\mathcal{L}_{\text{id}}(G, F) = E_{a \sim p_{\text{data}}(a)} [\|F(a) - a\|_1] + E_{b \sim p_{\text{data}}(b)} [\|G(b) - b\|_1] \quad (4.9)$$

Par conséquent, la fonction de perte totale de CycleGAN combine toutes ces fonctions de perte et est formulée comme suit :

$$\mathcal{L}_{\text{total}} = \mathcal{L}(G, F, D_A, D_B) = \mathcal{L}_{\text{GAN}}(G, D_B, A, B) + \mathcal{L}_{\text{GAN}}(F, D_A, B, A) + \lambda \mathcal{L}_{\text{cycle}}(G, F) + \sigma \mathcal{L}_{\text{id}}(G, F) \quad (4.10)$$

où λ et σ sont des paramètres de pondération qui contrôlent l'importance relative de la perte de cohérence cyclique et de la perte d'identité dans l'optimisation du modèle. Ces pondérations aident à équilibrer les différents aspects du modèle pour une performance optimale.

En conséquence, l'objectif pour l'optimisation globale du modèle est de résoudre l'équation 4.11 :

$$G^*, F^* = \arg \min_{G, F} \max_{D_A, D_B} \mathcal{L}(G, F, D_A, D_B) \quad (4.11)$$

L'architecture du réseau CycleGAN implique une configuration encodeur-décodeur pour les générateurs, et une configuration d'encodeur pour les discriminateurs. Les générateurs utilisent une architecture ResNet avec neuf blocs résiduels, appelée ResNet-9, qui permet une transformation profonde des caractéristiques tout en conservant les informations contextuelles grâce à des connexions résiduelles. D'autre part, les discriminateurs adoptent une architecture PatchGAN de 70×70 qui vise à classifier des patches individuels de l'image, de taille 70×70 , déterminant si chaque patch est réel ou généré artificiellement. Le modèle est entraîné avec des mini-batches de taille 1, en utilisant l'algorithme Adam avec un taux d'apprentissage de 0,0002, et avec $\lambda = 10$ et $\sigma = 5$ comme valeurs pour les pondérations de la fonction de perte.

La Figure 4.6 illustre le processus de CycleGAN pour transformer des images présentant des artefacts en images sans artefacts. Initialement, une image de radiographie thoracique issue de l'ensemble A , qui comprend de nombreuses images avec artefacts, est traitée par le générateur G_{AB} . Ce générateur a pour objectif de transformer l'image en éliminant les artefacts apparents pour produire une image qui semble appartenir à l'en-

semble B , constitué d'images sans artefacts. Le discriminateur B examine ensuite cette image générée pour évaluer son authenticité, se basant sur ses connaissances et sa familiarité avec les images sans artefacts de l'ensemble B . Il produit alors une décision binaire, 0 ou 1, indiquant si l'image semble ou non à une image sans artefacts. Cette image générée, traitée pour ressembler à une image de l'ensemble B , est ensuite réintroduite dans le générateur G_{BA} . Ce dernier tente de réintroduire des artefacts pour reconstruire l'image dans son état original, faisant partie de l'ensemble A . Ce processus de conversion inverse s'assure que l'image finale générée, connue sous le nom de cycle A , ressemble beaucoup à l'image d'entrée réelle originale. La qualité de cette reconstruction est évaluée par la perte de cohérence cyclique, une métrique qui mesure la similitude entre l'image originale de l'ensemble A et sa version reconstruite après avoir été transformée vers le domaine B puis reconvertie au domaine A . Une faible différence entre ces deux images indique une reconstruction de haute qualité, soulignant l'efficacité du CycleGAN à préserver l'intégrité des images tout au long du cycle de transformation.

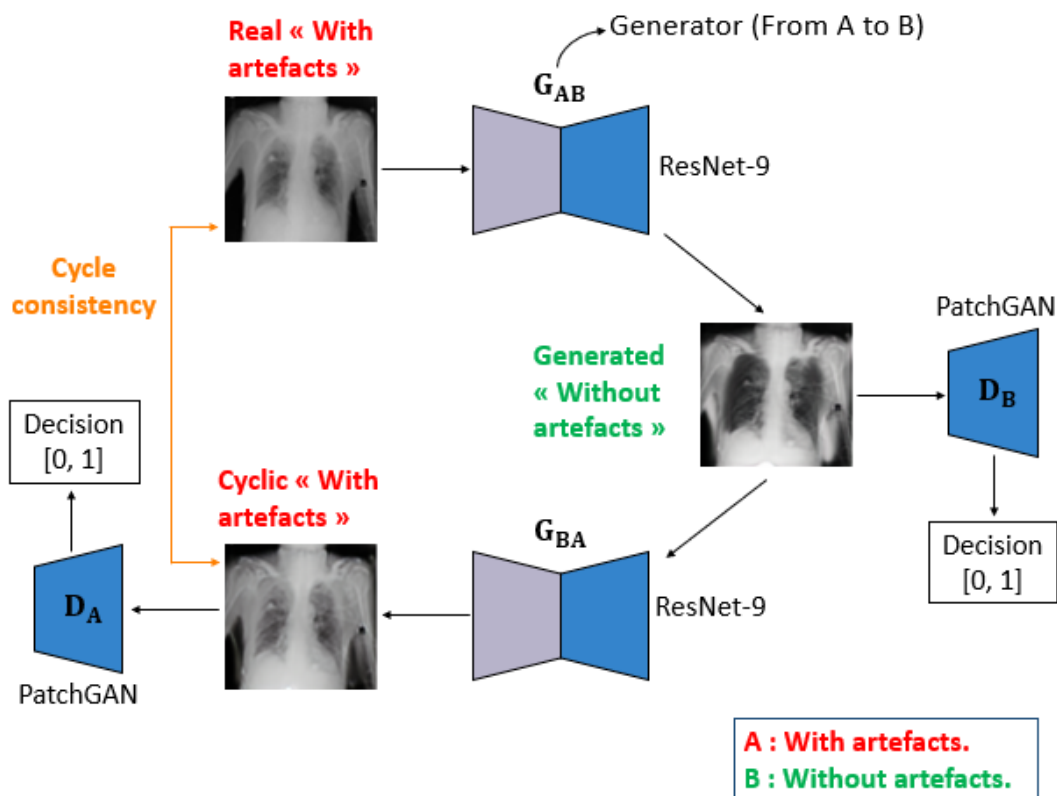


FIGURE 4.6 – L'architecture CycleGAN adaptée dans notre étude pour le scénario des images avec artefacts contre les images sans artefacts.

Nous avons appliqué ce modèle CycleGAN sur les 285 images contenant des artefacts de l'ensemble de données afin de générer des images plus nettes et de réduire l'effet du bruit des artefacts dans les images. La Figure 4.7 présente des exemples de transformation de ces images par le modèle CycleGAN, où les images générées montrent une réduction significative de l'impact des artefacts et des éléments électroniques qui deviennent presque invisibles. De plus, les images reconstruites démontrent également la capacité du modèle à restaurer les images à leur état original. Cela démontre l'efficacité du modèle CycleGAN non seulement à générer des images sans artefacts, mais aussi à restaurer avec précision les images originales, soulignant la précision du modèle CycleGAN dans le traitement des images médicales.

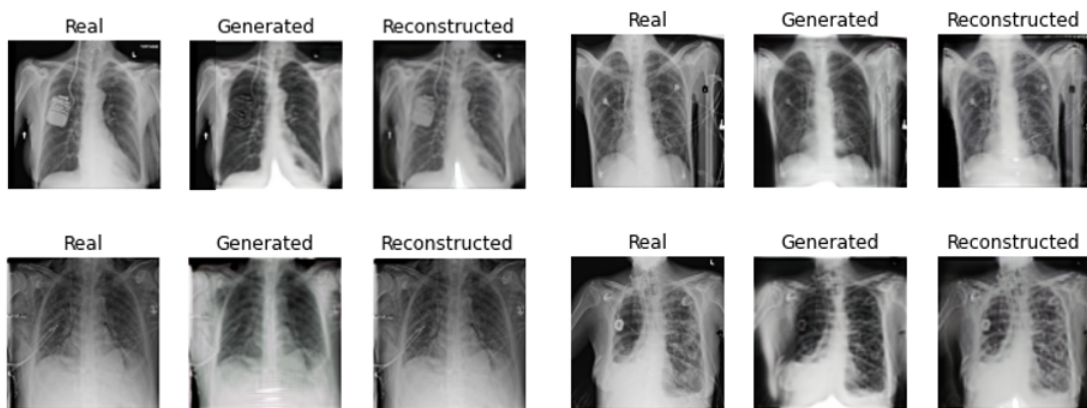


FIGURE 4.7 – Images avec et sans artefacts obtenues par le modèle CycleGAN.

Pour la suite de notre étude, les images sans artefacts générées par le modèle CycleGAN sont combinées avec celles du second cluster, qui contient des images sans artefacts, afin de procéder à la classification.

4.6 Modèle de classification

Pour la classification des images de pneumonie et de cas normaux, nous avons utilisé le modèle DenseNet-121. Après avoir redimensionné les images à 224×224 pixels et les avoir normalisées, elles ont été introduites dans le DenseNet-121, initialisé avec des poids pré-entraînés sur ImageNet. Nous avons gelé les poids de toutes les couches convolutionnelles inférieures, et nous avons ajouté deux couches denses de 1024 et 512 neurones, suivies d'une couche de dropout avec un taux de 0.5. La dernière couche de classification a été

ajustée pour produire une sortie binaire avec une fonction d'activation sigmoïde, indiquant si une image représente la pneumonie ou un cas normal.

Le modèle a été entraîné en utilisant des mini-batches de taille 16, avec l'optimiseur ADAM et un taux d'apprentissage de 10^{-4} . Pour prévenir le surapprentissage, une stratégie d'arrêt précoce (*earlystopping*) a été mise en place, interrompant l'entraînement si aucune amélioration n'était observée après cinq époques. La fonction de perte 'binary_crossentropy' a été utilisée pour la compilation du modèle.

Nous avons appliqué des techniques d'augmentation des données pour augmenter la taille de l'ensemble de données et ainsi améliorer la robustesse du modèle et sa capacité de généralisation. Ces techniques incluaient une rotation jusqu'à 10 degrés, un retournement horizontal, des décalages en hauteur et en largeur de 0, 1, un zoom de 0, 1, et l'utilisation du mode de remplissage nearest.

La performance du modèle a été évaluée avec l'aire sous la courbe (AUC) qui mesure l'efficacité du modèle à différencier avec précision entre les classes.

4.6.1 Caractéristiques cliniques

Dans notre étude utilisant la base de données ChestX-ray14, nous avons pris en compte non seulement les images de radiographie thoracique mais également des informations supplémentaires telles que le genre et l'âge du patient, ainsi que la position de vue de la radiographie (PA ou AP). Dans la pratique clinique, ces données cliniques sont utilisés par les médecins pour renforcer l'interprétation des images et affiner les diagnostics.

Ainsi, pour mieux exploiter les données disponibles, nous avons intégré ces informations cliniques et techniques dans notre modèle de classification. Pour ce faire, nous avons concaténé le vecteur de caractéristiques qui est la sortie de la dernière couche du DenseNet-121 avec le nouveau vecteur qui comprend les données cliniques. Ce vecteur additionnel, de dimension 3×1 , contient des informations telles que le genre et la position de vue, codés en binaire $\{0, 1\}$, ainsi que l'âge qui est normalisé dans l'intervalle $[0, 1]$, pour assurer une uniformité dans l'échelle des caractéristiques, évitant ainsi des biais potentiels liés à la variation dans l'amplitude des valeurs.

4.6.2 Convolutional Block Attention Module (CBAM)

Au cours des dernières années, le mécanisme d'attention a suscité un intérêt considérable et s'est avéré efficace dans les tâches de vision par ordinateur [205]. Il a également

démonstré sa supériorité dans une variété de tâches d'analyse d'images médicales [182, 216].

En général, les maladies sont caractérisées par des zones lésionnelles qui contiennent des informations cruciales pour le diagnostic. En effet, les systèmes de diagnostic assisté par ordinateur (CAD), intégrant des mécanismes d'attention, se focalisent sur ces caractéristiques lésionnelles spécifiques dans les images de radiographie thoracique. Au lieu de traiter uniformément toutes les parties de l'image, le mécanisme d'attention permet au modèle de prioriser et d'accorder plus d'importance aux régions contenant des caractéristiques sémantiques pertinentes liées aux lésions. Cette attention sélective aide à filtrer les informations non pertinentes, optimisant ainsi l'efficacité du modèle dans le diagnostic des maladies.

Le Convolutional Block Attention Module (CBAM) est un type de mécanisme d'attention qui peut être intégré dans les réseaux de neurones convolutifs pour en améliorer les performances [217]. Le CBAM introduit une approche de double attention, agissant à la fois sur les dimensions des canaux et spatiales. Comme le montre la Figure 4.8, la structure du CBAM comprend deux composants principaux : un module d'attention des canaux et un module d'attention spatiale.

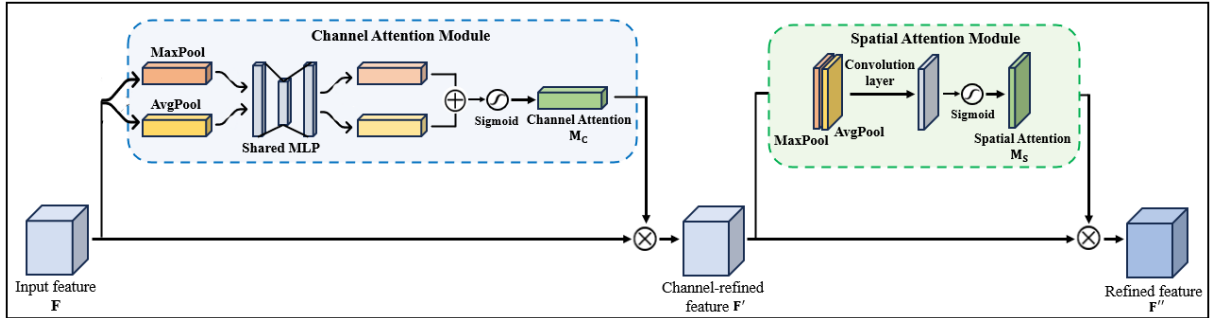


FIGURE 4.8 – Convolutional Block Attention Module (CBAM).

Une carte de caractéristiques donnée $F \in R^{C \times H \times W}$ (où C représente le nombre de canaux, H la hauteur, et W la largeur, de la carte de caractéristiques) est traitée par le module d'attention des canaux et le module d'attention spatiale de la manière suivante :

$$\begin{aligned} F' &= M_c(F) \otimes F \\ F'' &= M_s(F') \otimes F' \end{aligned} \quad (4.12)$$

où \otimes dénote la multiplication élément par élément, F' représente le résultat de la multiplication de la carte de caractéristiques par la carte d'attention des canaux, et F'' représente

le résultat de la multiplication de la carte d’attention spatiale par F' ou la sortie finale raffinée.

- **Module d’attention des canaux :** Chaque canal dans une carte de caractéristiques est considéré comme un détecteur de caractéristiques spécifique. L’attention des canaux met l’accent sur les informations essentielles et significatives, permettant d’extraire les canaux les plus informatifs.

Le module d’attention des canaux commence par compresser les carte de caractéristiques dans la dimension spatiale grâce à une opération de moyenne globale et de maximum global, comme illustré dans la partie bleue de la Figure 4.8, générant ainsi deux descripteurs de contexte spatial différents : F_c^{avg} et F_c^{max} . La couche de moyenne globale fournit les informations générales, tandis que la couche de maximum global met en évidence les caractéristiques les plus marquantes. Les deux descripteurs sont ensuite transmis à un réseau partagé pour produire la carte d’attention de canal $M_c \in R^{C \times 1 \times 1}$. Le réseau partagé est composé d’un perceptron multicouche (MLP) avec une seule couche cachée. Pour réduire la surcharge des paramètres, la taille de l’activation cachée est fixée à $R^{C/r \times 1 \times 1}$, où r est le ratio de réduction. Après traitement par le MLP, les vecteurs de caractéristiques issus de chaque descripteur sont combinés par addition élément par élément. La fonction sigmoïde est utilisée pour calculer la carte d’attention des canaux $M_c(F) \in R^{C \times 1 \times 1}$ qui est donnée par l’équation suivante :

$$\begin{aligned} M_c(F) &= \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \\ &= \sigma(W_1(W_0(F_{\text{avg}}^C)) + W_1(W_0(F_{\text{max}}^C))) \end{aligned} \quad (4.13)$$

où σ désigne la fonction sigmoïde, et W_0, W_1 sont les poids du MLP. Il est important de noter que les poids du MLP, W_0 et W_1 , sont partagés pour les deux entrées et que la fonction d’activation ReLU est appliquée après W_0 .

Ainsi, le module d’attention des canaux se concentre sur les informations pertinentes dans chaque canal de la carte de caractéristiques. Ce processus est accompli en apprenant à pondérer différemment les canaux en fonction de leur importance, en prenant en considération les relations inter-canaux des caractéristiques.

- **Module d’attention spatiale :** L’attention spatiale se concentre sur les parties informatives de la carte de caractéristiques. Comme illustré dans la partie verte de la Figure 4.8, le module d’attention spatiale commence par appliquer des opérations de pooling moyenne et de pooling maximum le long de l’axe des canaux, produisant

deux cartes de caractéristiques 2D : F_{avg}^S et F_{max}^S . L'application des opérations de pooling le long de l'axe des canaux s'avère efficace pour mettre en évidence les régions informatives [217]. Ces cartes sont ensuite concaténées et traitées par une opération de convolution avec un filtre de 7×7 . Après la convolution, la fonction sigmoïde calcule la carte d'attention spatiale 2D, $M_S(F) \in R^{H \times W}$, accentuant ainsi les zones significatives. L'attention spatiale est alors calculée par l'équation suivante :

$$\begin{aligned} M_S(F) &= \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \\ &= \sigma(f^{7 \times 7}([F_{avg}^S; F_{max}^S])) \end{aligned} \quad (4.14)$$

où σ désigne la fonction sigmoïde et $f^{7 \times 7}$ désigne l'opération de convolution avec une taille de filtre de 7×7 .

Ainsi, le module d'attention spatiale permet de mettre en évidence les zones les plus pertinentes, et minimiser l'impact des régions moins importantes. Cette focalisation sélective permet au réseau d'accorder une attention particulière aux endroits importants de la carte de caractéristiques, accentuant ainsi les éléments clés nécessaires à la réalisation de tâches spécifiques de classification ou de détection.

Par conséquent, le CBAM, combinant l'attention des canaux et l'attention spatiale, ajuste de manière adaptative son attention en traitant à la fois les dimensions des canaux et spatiales des cartes de caractéristiques. Cette double focalisation améliore substantiellement la capacité du réseau à identifier et à extraire les caractéristiques et les motifs les plus pertinents des données d'entrée.

Dans le cadre de notre étude, le CBAM a été intégré à la dernière couche de convolution du modèle DenseNet-121, optimisant ainsi les performances du modèle de classification en identifiant avec plus de précision les anomalies pulmonaires dans les images de radiographie thoracique.

4.7 Résultats et discussion

Dans cette étude, nous avons visé à réduire l'effet du bruit causé par les artefacts dans les images CXR pour améliorer la classification de la pneumonie. En effet, l'utilisation des images contenant des artefacts peut introduire du bruit pendant la classification. Étant donné que ces images ne sont pas étiquetées dans la base de données, nous avons développé un module de prétraitement pour distinguer les images claires de celles contenant des artefacts. Pour ce faire, diverses caractéristiques liées aux contours et à la texture ont

été extraites et utilisées comme entrées pour l'algorithme de clustering K-means. Comme indiqué dans la Table 4.1, cela a permis de créer deux clusters pour chaque classe : un cluster contenant des images sans artefacts et l'autre cluster contenant des images avec artefacts.

Nous avons ensuite réalisé deux expériences en utilisant le modèle DenseNet-121 : la première sans clustering préalable des images, et la seconde après avoir séparé les images en utilisant notre approche de clustering, en retenant uniquement le cluster d'images nettes pour la classification et éliminant le cluster contenant les images avec artefacts. L'AUC a été calculée pour chaque cas. Afin de garantir la robustesse et la stabilité des résultats, nous avons effectué dix répartitions aléatoires des données en ensembles d'entraînement, de validation et de test, et nous avons présenté les résultats sous forme de moyenne et d'écart type. Selon les résultats dans la Table 4.3, le modèle DenseNet-121 avec l'approche sans clustering a obtenu un score AUC de 77,96% pour la classification entre les cas de pneumonie et no finding. Cependant, avec notre méthode de clustering et l'utilisation du cluster d'images nettes seulement, l'AUC est amélioré à 79,58%. En comparaison, l'application du modèle DenseNet-121 sur les images sans artefacts identifiées manuellement a donné une AUC de 81,47%.

Bien que la différence d'AUC entre les approches d'étiquetage non supervisée et l'étiquetage manuel ne soit pas très élevée, elle illustre l'efficacité de notre méthode de clustering (étiquetage non supervisé) pour identifier précisément les images claires parmi celles contenant des artefacts. Cette technique pourrait potentiellement réduire la charge de travail des professionnels de santé en séparant de manière plus rapide et moins laborieuse les images avec et sans artefacts, et pourrait être appliquée à un plus grand nombre d'images.

Pour améliorer l'utilisation des images contenant des artefacts dans la classification, notre approche a consisté à appliquer CycleGAN aux images avec artefacts issues du clustering non supervisé. Cette technique a permis de transformer les images bruitées en équivalents plus nets, réduisant l'effet de bruit causé par la présence d'artefacts dans les images. Après la transformation par CycleGAN, ces images générées ont été combinées avec le cluster d'images déjà classées comme sans d'artefacts pour faire la classification. Le processus du clustering (étiquetage non supervisé) suivi de l'application de CycleGAN a amélioré la performance de classification, avec une augmentation de l'AUC de 3,98% par rapport à l'utilisation initiale du modèle DenseNet-121 de base (sans clustering). Selon les résultats présentés dans la Table 4.3, la classification de la pneumonie a atteint un score AUC de 81,94% après l'application de ces approches (clustering et CycleGAN).

TABLE 4.3 – AUC de la classification binaire des images de pneumonie et no finding en utilisant les différentes approches.

		Binary Classification (Pneumonia/No Finding)
Without clustering	DenseNet-121	77.96% \pm 1.5%
With clustering (unsupervised labeling)	DenseNet-121	79.58% \pm 1.6%
With clustering (manual labeling)	DenseNet-121	81.47% \pm 1.8%
With clustering (unsupervised labeling) and CycleGAN	DenseNet-121	81.94% \pm 1.4%
	DenseNet-121 + clinical variables	82.66% \pm 1.5%
	DenseNet-121 + CBAM + clinical variables	83.87% \pm1.4%

En plus, pour exploiter pleinement les informations disponibles dans la base de données, nous avons enrichi la représentation des images en ajoutant des données cliniques telles que l'âge, le genre et la position de vue. Ces caractéristiques supplémentaires ont été concaténées au vecteur de caractéristiques de la dernière couche du modèle DenseNet-121. Bien que l'intégration de ces caractéristiques n'ait pas significativement augmenté la performance du modèle, une légère amélioration a été observée. Cela pourrait s'expliquer par le fait que les caractéristiques des images intègrent implicitement certaines informations cliniques. En effet, l'intégration de ces variables cliniques a légèrement amélioré l'AUC, atteignant 82,66%.

Dans une autre expérience, après avoir appliqué le clustering (étiquetage non supervisé) et CycleGAN, nous avons introduit un mécanisme d'attention, à savoir le CBAM, dans notre modèle DenseNet-121. Nous avons également intégré les variables cliniques au modèle. Cette configuration a permis d'atteindre un score AUC de 83,87% pour la classification binaire entre les images de pneumonie et celles de no finding, comme indiqué dans la Table 4.3. Les résultats détaillés de notre approche pour la classification de la pneumonie sont présentés dans la Table 4.4.

Il est donc essentiel de prendre en compte la présence d'artefacts dans les images radiographiques lors du diagnostic des maladies pulmonaires. Les artefacts, qu'ils résultent de défauts techniques, d'interférences indésirables ou d'autres facteurs, peuvent considérablement altérer la qualité des images médicales et, par conséquent, la précision des

TABLE 4.4 – Performances du modèle DenseNet-121 pour les deux classes, pneumonie et no finding, après l’application du clustering, CycleGAN, et l’intégration du CBAM et des variables cliniques.

	Precision	Recall	F1-score	Accuracy
No finding	83 % \pm 1%	78 % \pm 1.3%	80 % \pm 1.1%	80.48 % \pm 1.5%
Pneumonia	76 % \pm 1.9%	81 % \pm 1.6%	78 % \pm 1.7%	

diagnostics. Ces perturbations peuvent fausser l’interprétation et la détection des anomalies pulmonaires, entraînant des diagnostics inexacts et des décisions thérapeutiques potentiellement inappropriées. En effectuant le clustering (étiquetage non supervisé) et en réduisant l’impact du bruit des artefacts en utilisant CycleGAN, notre approche améliore la performance et la fiabilité de classification du modèle d’apprentissage profond. En réduisant le bruit indésirable, notre approche permet aux modèles de se concentrer plus précisément sur les caractéristiques diagnostiques pertinentes. L’intégration du module CBAM renforce cette capacité en ajustant de manière adaptative l’attention du modèle, à la fois au niveau des canaux et spatial, pour se focaliser sur les caractéristiques les plus significatives et les zones spécifiques à la maladie dans l’image. Cela améliore non seulement la qualité des diagnostics en excluant les interférences inutiles mais aussi contribue à une classification plus précise et augmente l’efficacité globale du modèle en identifiant plus précisément les régions pertinentes pour la maladie. En conséquence, les radiologues peuvent prendre des décisions de traitement plus judicieuses et offrir des soins de meilleure qualité. Cette capacité à fournir des diagnostics plus précis et fiables a un impact direct et significatif sur les résultats cliniques pour les patients souffrant de maladies pulmonaires.

Pour généraliser notre approche, nous l’avons appliqué pour la classification binaire de la consolidation et no finding (2 751 images) de la base de données ChestX-ray14. Nous avons sélectionné toutes les images étiquetées comme consolidation, représentant 1 310 images, et avons sous-échantillonné les images no finding pour obtenir 1 441 images, équilibrant ainsi les deux classes. Ensuite, nous avons appliqué notre méthodologie, initialement développée pour la classification de la pneumonie, pour la classification de la consolidation. Après avoir extrait les caractéristiques liées aux contours et à la texture, nous avons utilisé l’algorithme K-means pour créer deux clusters pour chaque classe : l’un comprenant des images sans artefacts et l’autre des images contenant des artefacts, donnant un total de 1 328 images nettes et 1 423 images avec artefacts. Nous avons utilisé le CycleGAN pour traiter les 1 423 images contenant des artefacts, afin de les convertir en images sans artefacts. Les images générées ont été ensuite combinées avec le second clus-

ter d'images sans artefacts pour effectuer la classification. Les résultats, présentés dans la Table 4.5, indiquent une amélioration significative des performances de classification grâce à notre approche utilisant le clustering non supervisé et le CycleGAN, avec un modèle DenseNet-121 qui atteint un score AUC de 82,89%, comparativement à 78,52% avec le modèle DenseNet-121 de base sans clustering. Cette approche a mené à une amélioration de l'AUC de 4,37% comparativement au modèle de base. En intégrant le CBAM et les variables cliniques au modèle, cette configuration a abouti à un score AUC de 84,96% pour la classification de la consolidation et no finding, validant ainsi l'efficacité de notre approche comprenant le clustering non supervisé, la réduction des artefacts par CycleGAN, et l'ajout des modules d'attention et des données cliniques. Les résultats détaillés de notre approche pour la classification de la consolidation sont présentés dans la Table 4.6.

TABLE 4.5 – AUC de la classification binaire des images de consolidation et no finding en utilisant les différentes approches.

		Binary Classification (Consolidation/No Finding)
Without clustering	DenseNet-121	78.52% \pm 1.4%
With clustering (unsupervised labeling)	DenseNet-121	80.26% \pm 1.3%
With clustering (unsupervised labeling) and CycleGAN	DenseNet-121	82.89% \pm 1.6%
	DenseNet-121 + clinical variables	83.65% \pm 1.4%
	DenseNet-121 + CBAM + clinical variables	84.96% \pm1.5%

TABLE 4.6 – Performances du modèle DenseNet-121 pour les deux classes, consolidation et no finding, après l'application du clustering, CycleGAN, et l'intégration du CBAM et des variables cliniques.

	Precision	Recall	F1-score	Accuracy
No finding	83 % \pm 1.5%	82 % \pm 1.4%	82 % \pm 1.5%	81.67 % \pm 1.7%
Consolidation	79 % \pm 1.8%	81 % \pm 2.1%	80 % \pm 1.9%	

Nous avons comparé notre approche avec d'autres travaux rapportées dans la littérature utilisant la même base de données. Selon les résultats présentés dans la Table 4.7, notre approche dépasse les AUC rapportées dans la littérature pour la classification de

la pneumonie et de la consolidation. Ces résultats démontrent clairement que notre approche améliore significativement l’AUC du modèle de 5,91% pour la classification de la pneumonie et de 6,44% pour la classification de la consolidation, surpassant les études précédentes de 6,17% et 2,56% respectivement.

TABLE 4.7 – Comparaison des résultats obtenus avec d’autres méthodes existantes dans la littérature pour la classification des deux maladies pulmonaires de la base de données ChestX-ray14.

Reference	Year	Pneumonia	Consolidation
Wang et al. [47]	2017	66.4%	77%
Rajpurkar et al. [39]	2017	76.8%	74.1%
Ma et al. [184]	2019	72.2%	75%
Baltruschat et al. [126]	2019	76.7%	80%
Xu et al. [124]	2021	77.7%	81.7%
Soud et al. [186]	2021	73.3%	79%
Zhao et al. [187]	2021	76.92%	81.81%
Mann et al. [189]	2023	74.3%	82.4%
Our proposed approach	2024	83.87%±1.4%	84.96%±1.5%

La visualisation de l’attention offre un aperçu précieux de la logique interne des modèles d’apprentissage automatique, permettant aux chercheurs et aux médecins de mieux comprendre les bases des décisions prises par le modèle. Cette technique montre les régions de l’image sur lesquelles le modèle se concentre pour classer les maladies pulmonaires. En superposant la sortie de la couche d’attention à l’image originale, il devient possible de visualiser les zones que le modèle juge cruciales pour la classification. La Figure 4.9 illustre des exemples des zones sur lesquelles le modèle concentre son attention lors de la classification des cas de pneumonie et de consolidation. Cette visualisation est cruciale pour renforcer l’explicabilité et l’interprétabilité du modèle. En plus, visualiser les régions spécifiques qu’un modèle d’IA identifie comme pertinentes sur une radiographie thoracique peut renforcer la confiance des médecins dans l’usage de ces technologies d’IA en diagnostic médical.

Pour confirmer la pertinence de nos résultats, nous avons présenté ces images à un médecin expérimenté pour validation. Le médecin a confirmé que les régions identifiées par le modèle comme significatives correspondaient effectivement aux régions affectées par les maladies. Cette validation clinique par un expert de la santé renforce la fiabilité et la robustesse de notre approche.

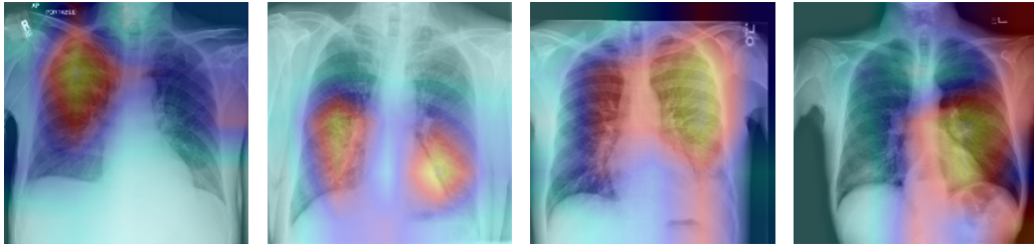


FIGURE 4.9 – Exemples de visualisation de la localisation des maladies sur les images de test de la pneumonie et de la consolidation. Une réponse plus élevée est indiquée en jaune.

4.8 Conclusion

Dans ce chapitre, nous avons présenté une nouvelle approche pour réduire l'effet du bruit causé par la présence d'artefacts tels que des fils ou des objets étrangers dans les images CXR, pour une meilleure classification des maladies pulmonaires. Nous avons d'abord effectué un clustering non supervisé pour distinguer les images claires de celles affectées par des artefacts. Pour atténuer l'impact du bruit des artefacts, nous avons généré des images plus nettes en utilisant le modèle CycleGAN. Par la suite, nous avons effectué la classification avec le modèle DenseNet-121, en y intégrant le CBAM et en ajoutant des variables cliniques. Les résultats expérimentaux ont montré que notre approche était efficace, améliorant l'AUC de 5,91% pour la classification de la pneumonie et de 6,44% pour la classification de la consolidation, par rapport au modèle de base.

L'identification précise des images contenant des composants électroniques et la réduction de l'impact du bruit électronique sont cruciales. Cela permet au modèle de se concentrer sur les caractéristiques diagnostiques pertinentes, améliorant ainsi la précision de la classification des maladies pulmonaires. De plus, la combinaison de l'apprentissage profond avec un mécanisme d'attention, tel que le CBAM, optimise la performance du modèle en ciblant les caractéristiques les plus pertinentes des images, tant au niveau des canaux que spatial. Enfin, la validation par un expert médical, qui a confirmé que les régions identifiées par le modèle correspondent aux zones affectées par la maladie, renforce la fiabilité et l'efficacité de notre approche.

Ce chapitre a fait l'objet d'un article soumis à la revue "*Computer Methods and Programs in Biomedicine*" et d'une présentation lors de la conférence internationale "*ICBES 2024*".

Il est courant que les images de radiographie thoracique révèlent la présence simultanée

de plusieurs pathologies. Ainsi, construire un seul modèle capable de classifier simultanément plusieurs pathologies présentes dans une seule image de radiographie thoracique peut significativement alléger la charge de travail des spécialistes, qui autrement devraient utiliser plusieurs systèmes distincts pour diagnostiquer différentes maladies. En conséquence, cela peut considérablement réduire le temps nécessaire pour prendre des décisions cliniques, améliorant ainsi l'efficacité du diagnostic.

Dans le chapitre suivant, nous aborderons la classification multi-label des maladies pulmonaires en passant par une méthode de segmentation pour générer des masques précis contenant les poumons ainsi que les régions les plus indicatives des pathologies.

NOUVELLE APPROCHE DE SEGMENTATION D'IMAGES DE RADIOGRAPHIE THORACIQUE BASÉE SUR CYCLEGAN POUR UNE MEILLEURE CLASSIFICATION MULTI-LABEL DES MALADIES THORACIQUES

5.1 Introduction

En pratique clinique, les radiologues utilisent la reconnaissance de motifs dans les images médicales pour la prise de décisions diagnostiques. Les compétences et les connaissances des radiologues peuvent être valorisées grâce à la radiomique, qui a prouvé l'efficacité des biomarqueurs basés sur l'imagerie, notamment pour le diagnostic et le pronostic du cancer [218]. Ce domaine, en pleine expansion dans l'analyse quantitative d'images, relie l'exploration approfondie des données d'image aux résultats cliniques et biologiques [219].

La radiomique permet l'extraction d'informations pertinentes à partir des images, offrant une analyse détaillée et une évaluation exhaustive de la région d'intérêt (ROI) [220]. Pourtant, une analyse radiomique efficace nécessite une segmentation précise et une génération de masques délimitant clairement les ROIs. Les approches traditionnelles comme U-Net, bien que largement utilisées pour la segmentation des images médicales, échouent souvent à inclure les régions spécifiquement affectées par les pathologies dans les masques, ce qui est crucial pour une analyse radiomique exacte.

L’objectif de ce chapitre est de résoudre le problème de l’exclusion des zones affectées par la maladie dans les masques utilisés pour l’extraction des radiomiques, une lacune qui peut compromettre la précision du diagnostic.

Nous avons développé une nouvelle approche de segmentation pour la génération de masques, conçue pour inclure les zones affectées par les pathologies dans la région d’intérêt. Cette approche assure que les caractéristiques radiomiques extraites proviennent directement des zones les plus indicatives des pathologies, améliorant ainsi les performances de classification. En plus, nous avons développé une nouvelle approche de sélection de caractéristiques pour identifier a priori les caractéristiques les plus pertinentes avant de faire la classification, concentrant ainsi notre analyse sur les caractéristiques les plus significatives pour accroître l’efficacité du modèle de classification.

Ces avancées en segmentation et en sélection des caractéristiques représentent un progrès significatif vers l’amélioration des modèles d’apprentissage automatique pour une classification des pathologies plus précise et efficace.

5.2 Base de données

Dans notre étude, nous avons utilisé la base de données ChestX-ray14 [47] qui comprend 112 120 images de radiographie thoracique en vue frontale provenant de 30 805 patients uniques. Chaque image CXR est annotée avec une ou plusieurs étiquettes d’un ensemble de 14 pathologies thoraciques.

Comme illustré dans la Figure 5.1, parmi ces 112 120 images CXR, 60 361 images sont classées comme "No finding" indiquant l’absence de pathologies. Les 51 759 images CXR restantes sont étiquetées avec une ou plusieurs des 14 pathologies thoraciques. Parmi celles-ci, 30 963 images CXR présentent une seule pathologie, offrant des aperçus sur les caractéristiques radiographiques distinctes des conditions thoraciques individuelles. Les 20 796 images CXR sont étiquetées avec deux maladies ou plus, ce qui est essentiel pour étudier les cooccurrences des différentes conditions thoraciques.

La Table 5.1 détaille le nombre d’images étiquetées pour chaque type de maladie dans la base de données ChestX-ray14. La première ligne indique le nombre total d’images associées à chaque maladie, fournissant un aperçu de la prévalence de chaque maladie dans la base de données. La troisième ligne présente le nombre d’images qui sont étiquetées avec une maladie spécifique. La deuxième ligne montre le nombre d’images étiquetées avec deux maladies ou plus, impliquant 820 combinaisons différentes de conditions patho-

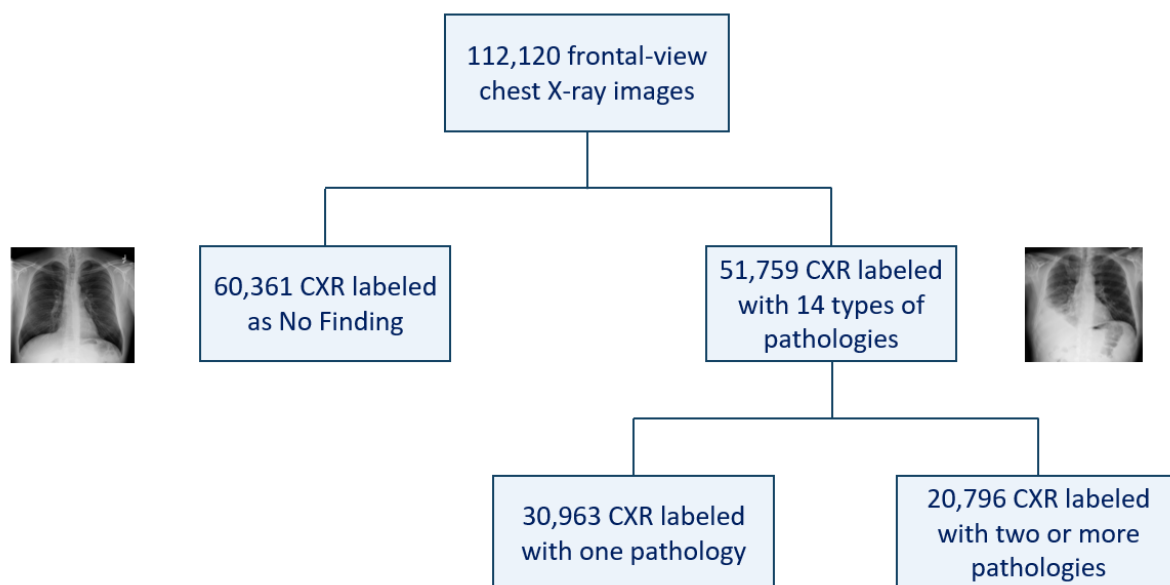


FIGURE 5.1 – Aperçu de la répartition des images de la base de données ChestX-ray14.

logiques dans la base de données. Le nombre d’images pour chaque type de maladies varie considérablement, allant de plusieurs milliers à seulement quelques-unes. Par exemple, 1 604 images sont étiquetées à la fois par infiltration et effusion ; 1 350 images sont étiquetées à la fois par infiltration et atelectasis ; 737 images sont étiquetées par infiltration, atelectasis et effusion ; 441 images sont étiquetées à la fois par infiltration et consolidation ; 81 images sont étiquetées par infiltration, atelectasis, consolidation et effusion. Cette variation illustre non seulement la prévalence de certaines conditions mais met également en évidence la rareté et la complexité des autres.

Pour aborder la tâche de classification multi-label, nous avons initialement sélectionné les maladies d’effusion et d’infiltration, principalement en raison de la grande disponibilité d’images CXR dans la base de données ChestX-ray14 où ces deux conditions apparaissent à la fois dans l’image, permettant ainsi de tester efficacement la capacité du modèle à identifier des pathologies multiples simultanément. Nous avons alors sélectionné les 1 604 images CXR de la base de données ChestX-ray14 qui sont simultanément étiquetées pour ces deux conditions. Afin d’équilibrer notre ensemble de données, nous avons également inclus un nombre identique d’images classées uniquement soit comme effusion soit comme infiltration. Nous avons donc mené nos expériences sur un total de 4 812 images CXR.

TABLE 5.1 – Nombre d’images catégorisées par type de pathologie dans la base de données ChestX-ray14

Pathology	Infiltration	Effusion	Atelectasis	Nodule	Mass	Pneumothorax	Consolidation
Image number	19 894	13 317	11 559	6 331	5 782	5 302	4 667
Multi-pathology image number	10 347	9 362	7 344	3 626	3 643	3 108	3 357
Single pathology image number	9 547	3 955	4 215	2 705	2 139	2 194	1 310

Pathology	Pleural Thickening	Cardiomegaly	Emphysema	Edema	Fibrosis	Pneumonia	Hernia
Image number	3 385	2 776	2 516	2 303	1 686	1 431	227
Multi-pathology image number	2 259	1 683	1 624	1 675	959	1 109	117
Single pathology image number	1 126	1 093	892	628	727	322	110

5.3 Approche proposée

La Figure 5.2 montre un aperçu de la méthodologie utilisée. Pour générer des masques précis incluant les zones affectées par la maladie, nécessaires pour l’extraction des caractéristiques radiomiques, nous avons développé une nouvelle approche de segmentation. Cette approche commence par l’application du modèle CycleGAN aux images radiographiques afin de générer une version restaurée des poumons, considérée comme pseudo-saine. L’étape suivante consiste à effectuer une soustraction d’image entre la version restaurée et l’image originale présentant des pathologies. Ce processus crée une carte de pathologie mettant en évidence avec précision les régions affectées par les anomalies dans l’image initiale. Enfin, cette carte est convertie en format binaire, produisant ainsi un masque détaillé des pathologies.

Ensuite, les images ont été redimensionnées à 224×224 pixels, converties en niveaux de gris, et les caractéristiques radiomiques ont été extraites. Pour sélectionner les caractéristiques les plus pertinentes, nous avons adopté une nouvelle approche de sélection des caractéristiques qui intègre initialement l’Analyse Hiérarchique en Clusters (HCA), suivie d’une évaluation rigoureuse en utilisant un test statistique.

Après le processus de sélection des caractéristiques, un algorithme d’apprentissage automatique a été utilisé pour la classification des images en fonction des caractéristiques sélectionnées. Les données ont été divisées en ensembles d’entraînement et de test, avec 80% alloués à l’entraînement et les 20% restants utilisées comme données de test (choisis aléatoirement). L’algorithme d’apprentissage automatique est entraîné en utilisant les caractéristiques des images de l’ensemble d’entraînement. Pour évaluer les performances

de nos modèles d'apprentissage automatique sur l'ensemble de test, plusieurs métriques ont été utilisées, y compris la matrice de confusion et ses paramètres métriques associés tels que l'accuracy, la précision, le rappel, le F1-Score, et l'aire sous la courbe (AUC).

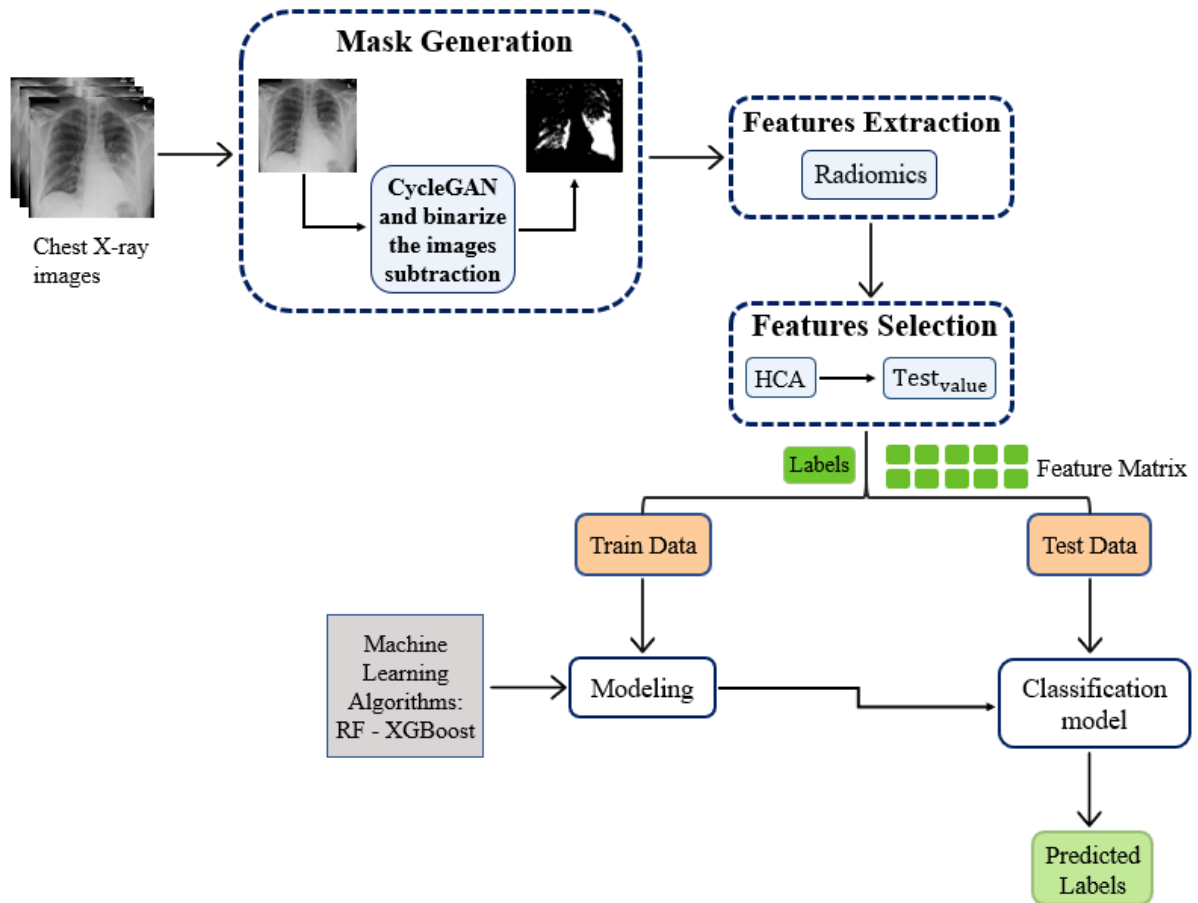


FIGURE 5.2 – Aperçu de la méthodologie proposée.

5.4 Génération des masques

Notre premier objectif est de développer une méthode pour créer des masques précis, permettant d'isoler les poumons ainsi que les régions les plus indicatives des pathologies. Dans cette section, nous commencerons par l'application du modèle U-Net pour générer des masques, puis nous introduirons notre approche pour affiner la génération des masques.

5.4.1 U-Net pour la segmentation et génération des masques

U-Net est un réseau de neurones convolutif conçu pour la segmentation des images médicales, avec une architecture en forme de U divisée en un encodeur et un décodeur [110]. L'encodeur extrait les caractéristiques essentielles tandis que le décodeur reconstruit l'image segmentée. Les connexions de "skip" entre l'encodeur et le décodeur améliorent la précision de la segmentation en transférant directement les informations contextuelles et les caractéristiques de bas niveau à travers le réseau. La structure de notre architecture de segmentation est illustrée dans la Figure 5.3.

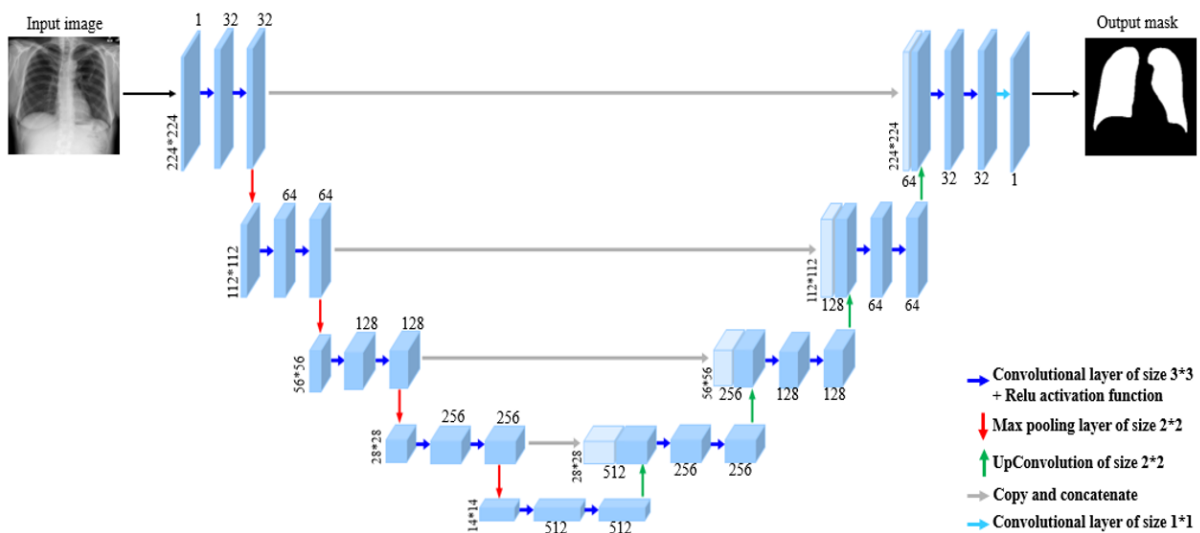


FIGURE 5.3 – L'architecture du réseau de segmentation U-Net.

Pour entraîner, valider et tester le modèle U-Net, nous avons utilisé l'ensemble de données open-source de Kaggle "Chest Xray Masks and Labels" [221]. Cet ensemble de données contient 704 images de radiographies pulmonaires et leurs masques correspondants. Une fois le segmenteur de poumons entraîné, nous l'avons utilisé pour générer des masques des images pulmonaires de notre ensemble de données.

Les images de sortie produites par le décodeur ont des valeurs de pixel binaires de 0 ou 1, où 0 correspond au noir et 1 au blanc. Cela permet de créer des masques en noir et blanc qui servent à isoler et mettre en évidence les zones pulmonaires pour une analyse ultérieure. La Figure 5.4 illustre les masques générés avec U-Net.

La Figure 5.4 (a) et (b) illustre la capacité de U-Net à générer des masques précis des

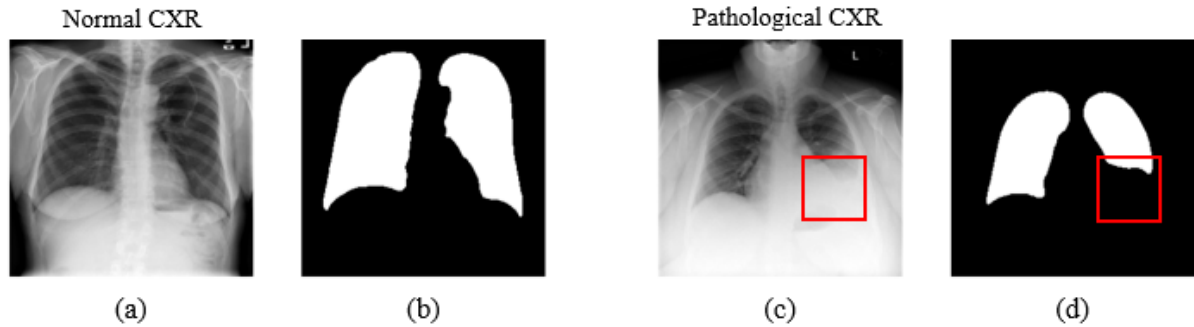


FIGURE 5.4 – Un exemple d’images de radiographie thoracique et de leurs masques obtenus correspondants : (a) et (c) CXR originaux, (b) et (d) masques obtenus à l’aide de U-Net. Le carré rouge dans les images indique la zone qui n’a pas été incluse dans le masque généré par U-Net, malgré son appartenance à la région pulmonaire affectée par la maladie.

poumons pour une radiographie thoracique sans pathologie. Cependant, lorsque la radiographie thoracique présente une pathologie, comme le montre la Figure 5.4 (c), indiquant une anomalie au bas du poumon gauche, indiquée par le carré rouge, cette région pathologique est exclue du masque (Figure 5.4 (d)). Par conséquent, le masque produit ne couvre que les parties saines des poumons. Cette limitation dans les performances de U-Net est probablement due à son entraînement sur des images ne comportant pas de maladies, ce qui limite sa capacité à segmenter précisément les zones affectées par la maladie. Ce problème est critique car cela signifie une exclusion des zones pathologiques pertinentes de l’analyse et de l’évaluation des caractéristiques pulmonaires.

Pour résoudre ce problème, nous avons développé une nouvelle approche pour générer des masques qui incluent les zones pathologiques. Cette approche est présentée et discutée dans la section suivante.

5.4.2 Génération avancée de masques de pathologie

Pour obtenir des masques de pathologie précis à partir des images de radiographie thoracique, nous avons développé une nouvelle approche de segmentation. Comme illustré dans la Figure 5.5, notre approche commence par l’application du modèle CycleGAN aux images de radiographie thoracique afin de générer une version restaurée des poumons considérée comme "pseudo-saine". Ensuite, une soustraction d’images est effectuée entre l’image restaurée et l’image originale présentant des pathologies. Cette opération produit une carte de pathologie qui délimite avec précision les zones affectées par des anomalies

dans l’image initiale. Enfin, cette carte est convertie en un format binaire, produisant ainsi un masque très précis des pathologies identifiées.

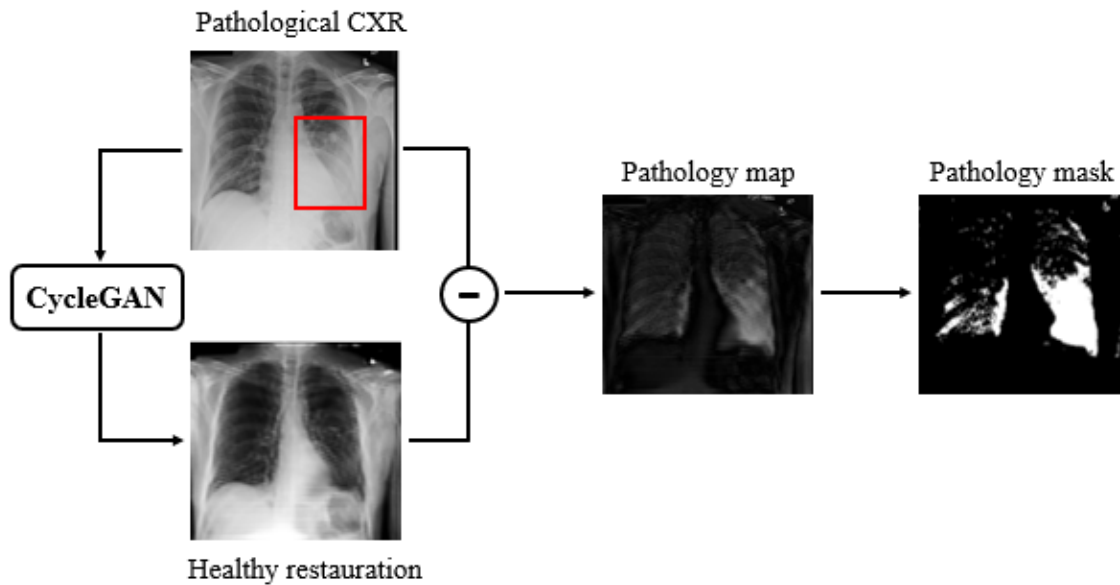


FIGURE 5.5 – Notre approche proposée pour la génération de masques de pathologie dans les images de radiographie thoracique.

1. Génération d’images pseudo-saine avec CycleGAN

CycleGAN est spécifiquement conçu pour les tâches de conversion d’image à image non appariées où les paires de données d’entraînement correspondantes ne sont pas disponibles. La Figure 5.6 illustre le processus de CycleGAN pour convertir des images présentant des pathologies en images sans pathologies, c’est-à-dire en une restauration saine des poumons.

Une image de radiographie thoracique issue de l’ensemble A , qui comprend de nombreuses images avec pathologies, est traitée par le générateur G_{AB} pour produire une image sans pathologies semblant appartenir à l’ensemble B , constitué d’images sans pathologies. Le discriminateur B évalue ensuite cette image générée pour déterminer son authenticité, se basant sur ses connaissances et sa familiarité avec les images sans pathologies de l’ensemble B . Il produit alors une décision binaire, 0 ou 1, indiquant si l’image semble ou non à une image sans pathologies. Cette image générée, traitée pour ressembler à une image de l’ensemble B , est

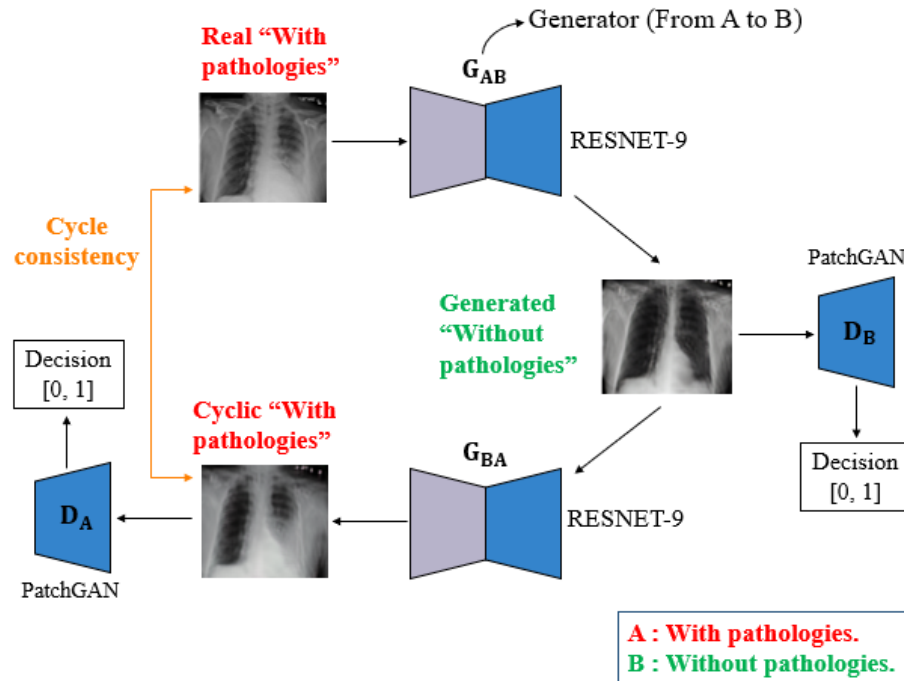


FIGURE 5.6 – L’architecture CycleGAN adaptée dans notre étude pour le scénario des images avec pathologies contre les images sans pathologies.

ensuite réintroduite dans le générateur G_{BA} . Ce dernier tente de réintroduire les pathologies dans l’image, restaurant ainsi son état pathologique d’origine. Ce processus de conversion inverse s’assure que l’image finale générée, connue sous le nom de cycle A , ressemble beaucoup à l’image d’entrée réelle originale. La qualité de cette reconstruction est évaluée par la perte de cohérence cyclique, une métrique qui mesure la similitude entre l’image originale de l’ensemble A et sa version reconstruite après avoir été transformée vers le domaine B puis reconvertie au domaine A . Une faible différence entre ces deux images indique une reconstruction de haute qualité, soulignant l’efficacité du CycleGAN à préserver l’intégrité des images tout au long du cycle de transformation.

Le modèle CycleGAN est appliqué aux images présentant des pathologies, dans le but de produire des images qui montrent les poumons dans un état sain. La Figure 5.7 présente des exemples de transformation des images par le modèle CycleGAN, où les images générées montrent que les poumons sont en bonne santé et sans pathologies. De plus, les images reconstruites démontrent la capacité du modèle à restaurer les poumons dans leur état d’origine, avec les pathologies pré-

sentes. Cela démontre l’efficacité du modèle non seulement à générer des images présentant un aspect normal et sain des poumons, mais aussi à restaurer avec précision les conditions pathologiques d’origine, soulignant la précision du modèle CycleGAN dans le traitement des images médicales.

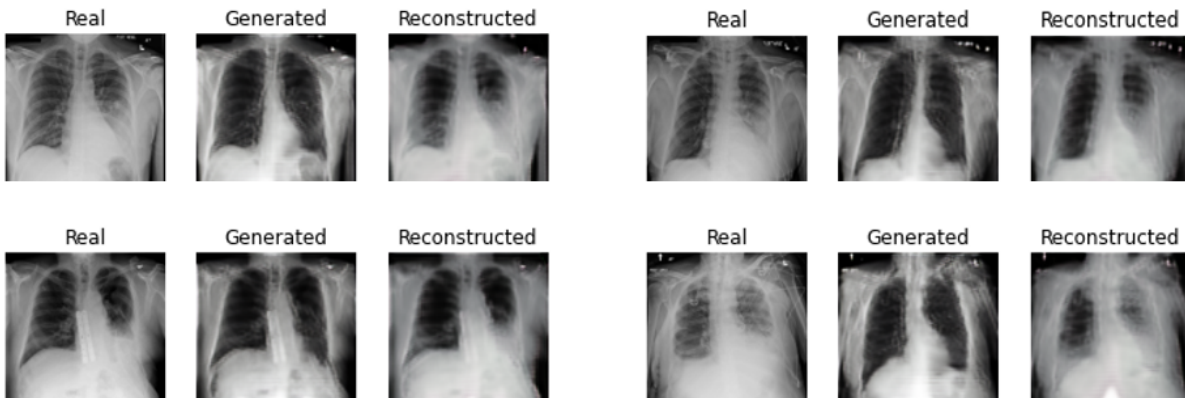


FIGURE 5.7 – Images avec et sans pathologies obtenues par le modèle CycleGAN.

2. Création des masques de pathologie par soustraction d’images et binarisation

Après avoir généré des images représentant une restauration saine des poumons, nous procédons à une soustraction d’image entre cette version restaurée, considérée comme pseudo-saine, et l’image originale présentant des pathologies. Cette opération produit une carte de pathologie qui délimite avec précision les zones affectées par les pathologies dans l’image d’origine. Pour convertir cette carte en un format binaire, nous avons utilisé la méthode d’Otsu, une technique de seuillage automatique conçue pour trouver le seuil optimal pour binariser une image en niveaux de gris [222]. La méthode d’Otsu identifie le seuil optimal qui maximise la variance entre les classes, ce qui équivaut à minimiser la variance dans chaque classe. Cette technique sépare efficacement l’image en régions distinctes en fonction de leurs niveaux d’intensité [223]. Ce processus conduit au développement d’un masque extrêmement précis, ciblant spécifiquement les régions affectées par les pathologies dans l’image.

La Figure 5.8 présente des exemples de génération de masques utilisant notre approche. Le masque obtenu permet une visualisation et une analyse détaillées des zones pathologiques, améliorant ainsi les capacités de diagnostic et permettant

ainsi de concentrer l'analyse sur les aspects pathologiques des images.

Ainsi, pour le reste de notre étude, nous avons utilisé les masques produits par notre approche afin d'extraire des caractéristiques des images.

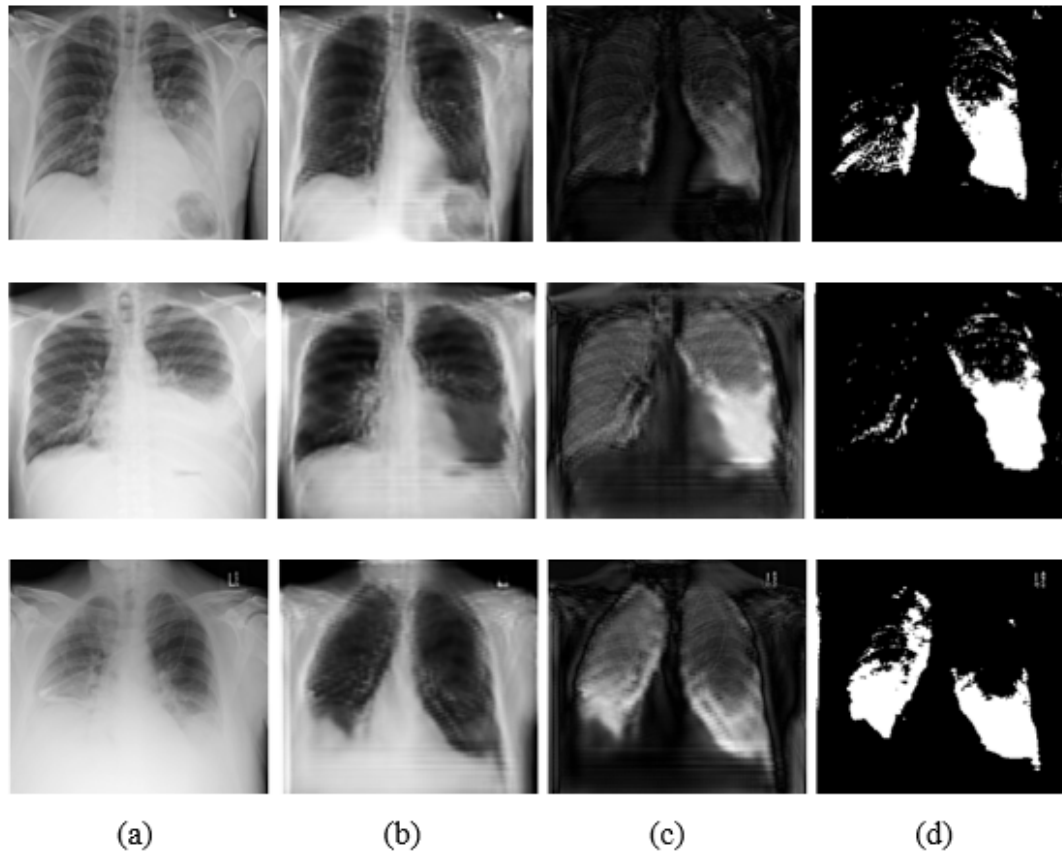


FIGURE 5.8 – Exemples de génération de masques utilisant notre approche proposée. La colonne (a) montre l'image originale avec pathologie, la colonne (b) présente l'image de poumon sain générée à l'aide du modèle CycleGAN, la colonne (c) illustre la cartographie de la pathologie obtenue en soustrayant les images originale et générée, et la colonne (d) présente le masque obtenu après binarisation.

5.5 Extraction des caractéristiques

En pratique clinique, l'expertise des radiologues repose sur la reconnaissance de motifs dans les images médicales, ce qui guide leurs décisions diagnostiques. Les compétences et les connaissances des radiologues peuvent être valorisées grâce à la radiomique. Ce domaine, en pleine expansion dans l'analyse quantitative d'images, établit un lien entre l'exploration exhaustive des données d'image et les résultats cliniques [219].

Le concept principal de la radiomique repose sur la reconnaissance que les images médicales contiennent une richesse d’informations bien au-delà de ce qui peut être perçu par l’œil humain. Les caractéristiques quantitatives de l’imagerie, également connues sous le nom de « caractéristiques radiomiques », offrent une perspective approfondie des phénotypes tumoraux, incluant des aspects tels que l’intensité, la forme, la taille ou le volume, et la texture [224]. Ces caractéristiques radiomiques, extraites à partir de différentes modalités d’imagerie, fournissent une compréhension complète du comportement et des propriétés de la pathologie. En exploitant ces caractéristiques radiomiques détaillées, les médecins peuvent obtenir une compréhension plus fine de la signification biologique et clinique des tumeurs, permettant des diagnostics et des stratégies de traitement plus précis.

En extrayant les caractéristiques radiomiques, il est possible de quantifier de manière précise les particularités propres à une pathologie et d’évaluer sa progression à partir d’une série d’images médicales. Cette approche repose sur des algorithmes avancés d’analyse d’image, qui transforment les images en données de haute dimension exploitables [225].

En utilisant les masques générés qui définissent avec précision les régions d’intérêt (ROIs) affectées par les pathologies dans les images originales, nous avons extrait des caractéristiques radiomiques à l’aide de l’outil PyRadiomics [226]. Ainsi, notre approche garantit que les caractéristiques sont extraites des zones définies par les masques, permettant ainsi de concentrer l’analyse sur les aspects pathologiques des images. Les caractéristiques radiomiques englobent un large éventail de caractéristiques, divisées en classes décrites ci-dessous [226]. Les caractéristiques calculées pour chaque groupe sont présentées dans la Table 5.2.

— **Caractéristiques statistiques du premier ordre**

Les statistiques du premier ordre décrivent la distribution des valeurs de pixels dans la région de l’image délimitée par le masque sans considérer les relations spatiales. Ces statistiques fournissent des informations sur les métriques d’intensité fondamentales, détaillant la variabilité et la distribution des intensités de pixels. Ainsi, elles offrent un aperçu complet des propriétés de base de la région de l’image examinée.

— **Caractéristiques de Gray Level Co-occurrence Matrix (GLCM)**

La GLCM décrit la fonction de probabilité conjointe du second ordre d’une région de l’image contrainte par le masque, en tenant compte des relations spatiales entre les pixels. Elle permet de caractériser la texture en analysant la fréquence d’occur-

rence des paires de pixels avec des valeurs d'intensité spécifiques et une relation spatiale donnée. La matrice de co-occurrence est une fonction de deux paramètres : la distance (d) mesurée en nombre de pixels et leur orientation (θ). L'orientation θ prend les valeurs 0° , 45° , 90° et 135° , qui représentent les quatre directions : horizontale, diagonale, verticale et anti-diagonale, respectivement. L'occurrence d'un motif de niveau de gris peut être représentée par une matrice de fréquence relative qui décrit la fréquence d'apparition de deux pixels de niveaux de gris dans la fenêtre qui sont séparés par une distance d dans la direction θ .

— **Caractéristiques de Gray Level Size Zone Matrix (GLSZM)**

La GLSZM quantifie les zones de niveaux de gris à l'intérieur de la région d'intérêt d'une image. Elle définit une zone de niveaux de gris comme étant le nombre de pixels connectés partageant la même intensité de niveaux de gris. La GLSZM capture la distribution de ces zones en fonction de leur niveau de gris.

— **Caractéristiques de Gray Level Run Length Matrix (GLRLM)**

La GLRLM quantifie les longueurs des séquences de pixels consécutifs qui partagent la même valeur de niveaux de gris.

— **Caractéristiques de Neighbouring Gray Tone Difference Matrix (NGTDM)**

La NGTDM quantifie la variation d'intensité entre la valeur de gris d'un pixel et la valeur de gris moyenne de ses pixels voisins dans une distance spécifiée. Pour chaque niveau de gris, la NGTDM calcule et stocke la somme des différences absolues entre la valeur de gris d'un pixel et la valeur de gris moyenne de ses pixels voisins.

— **Caractéristiques de Gray Level Dependence Matrix (GLDM)**

La GLDM quantifie les dépendances des niveaux de gris dans une image. Une dépendance de niveau de gris fait référence au nombre de pixels connectés dans une distance spécifiée qui dépendent du pixel central.

Après avoir extrait ces caractéristiques radiométriques, nous avons remarqué que certaines d'entre elles présentaient des valeurs identiques pour toutes les classes des images

étudiées. Ces caractéristiques, ne contribuant pas à une distinction significative entre les classes d’images, ont donc été éliminées. En conséquence, nous avons obtenu un ensemble de 62 caractéristiques distinctives qui fournissent des informations discriminantes sur les images.

5.6 Sélection des caractéristiques

Dans notre étude, nous avons mis en œuvre une nouvelle approche de sélection des caractéristiques pour identifier les caractéristiques les plus pertinentes pour chaque classe, avant même la phase de classification. En effet, la sélection des caractéristiques est essentielle pour optimiser la performance des modèles d’apprentissage automatique parce qu’elle permet d’identifier et de conserver uniquement les caractéristiques les plus pertinentes et significatives tout en éliminant les caractéristiques redondantes ou non essentielles. Cette approche permet de concentrer l’analyse sur les caractéristiques clés qui distinguent les différentes classes, améliorant ainsi la précision et l’efficacité des modèles.

De plus, notre approche de sélection des caractéristiques est particulièrement utile dans les scénarios multi-label. Elle permet d’identifier les caractéristiques pertinentes pour chaque classe, même lorsque plusieurs maladies sont présentes dans la même image. Cela contribue à un diagnostic médical précis et à une planification de traitement adaptée, notamment dans les cas complexes où l’identification précise des maladies est essentielle.

Notre approche commence par une analyse de clustering hiérarchique (HCA) pour regrouper les caractéristiques extraites en fonction de leurs similarités. Ensuite, ces clusters sont soumis à une évaluation rigoureuse à l’aide d’un test statistique ($Test_{value}$) spécialement conçue pour sélectionner les caractéristiques qui sont statistiquement les plus significatives et les plus pertinentes. Cette étape de sélection garantit que seules les caractéristiques les plus informatives sont retenues, optimisant ainsi la performance du modèle d’apprentissage automatique.

5.6.1 Clustering avec l’Analyse Hiérarchique en Clusters

Pour diviser nos données en groupes sans présupposer le nombre de clusters, nous avons utilisé l’analyse hiérarchique en clusters (HCA), une méthode statistique qui identifie des clusters en analysant les similarités entre les échantillons [201]. Dans cette étude, nous avons utilisé le critère de liaison de Ward, qui minimise la variance intra-cluster totale au

TABLE 5.2 – Caractéristiques extraites pour chaque groupe de caractéristiques radio-miques.

Feature group	Features
First-order statistics	Mean, Median, Variance, Standard Deviation, Minimum, Maximum, 10th percentile, 90th percentile, Energy, Total Energy, Entropy, Range, Interquartile Range, Skewness, Kurtosis, Uniformity, Mean Absolute Deviation, Robust Mean Absolute Deviation, Root Mean Squared
Gray Level Co-occurrence Matrix (GLCM)	Contrast, Dissimilarity, Homogeneity, Energy, Angular Second Moment, Correlation
Gray Level Size Zone Matrix (GLSZM)	Small Area Emphasis, Large Area Emphasis, Gray Level Non-Uniformity, Gray Level Non-Uniformity Normalized, Size-Zone Non-Uniformity, Size-Zone Non-Uniformity Normalized, Zone Percentage, Gray Level Variance, Zone Variance, Zone Entropy, Low Gray Level Zone Emphasis, High Gray Level Zone Emphasis, Small Area Low Gray Level Emphasis, Small Area High Gray Level Emphasis, Large Area Low Gray Level Emphasis, Large Area High Gray Level Emphasis
Gray Level Run Length Matrix (GLRLM)	Short Run Emphasis, Long Run Emphasis, Gray Level Non-Uniformity, Gray Level Non-Uniformity Normalized, Run Length Non-Uniformity, Run Length Non-Uniformity Normalized, Run Percentage, Gray Level Variance, Run Variance, Run Entropy, Low Gray Level Run Emphasis, High Gray Level Run Emphasis, Short Run Low Gray Level Emphasis, Short Run High Gray Level Emphasis, Long Run Low Gray Level Emphasis, Long Run High Gray Level Emphasis
Neighbouring Gray Tone Difference Matrix (NGTDM)	Coarseness, Contrast, Busyness, Complexity, Strength
Gray Level Dependence Matrix (GLDM)	Small Dependence Emphasis, Large Dependence Emphasis, Gray Level Non-Uniformity, Gray Level Non-Uniformity Normalized, Dependence Non-Uniformity, Dependence Non-Uniformity Normalized, Gray Level Variance, Dependence Variance, Dependence Entropy, Dependence Percentage, Low Gray Level Emphasis, High Gray Level Emphasis, Small Dependence Low Gray Level Emphasis, Small Dependence High Gray Level Emphasis, Large Dependence Low Gray Level Emphasis, Large Dependence High Gray Level Emphasis

cours de la procédure de clustering. Le résultat de cette analyse est généralement illustré par un dendrogramme, où des distances plus courtes entre les branches indiquent de plus grandes similarités entre les échantillons.

En appliquant la HCA aux caractéristiques extraites, nous avons pu les regrouper en fonction de leurs similarités, organisant ainsi les données en clusters. Ensuite, les caractéristiques de chaque cluster ont été évaluées à l’aide d’un test statistique ($Test_{\text{value}}$) qui sera détaillé dans la section suivante.

5.6.2 Sélection des caractéristiques statistiquement significatives

Pour chaque cluster identifié, nous avons calculé le test statistique ($Test_{\text{value}}$) ainsi que la p – *value* pour identifier les caractéristiques les plus statistiquement significatives. Le test défini est un test statistique conçu pour comparer, pour chaque caractéristique dans un cluster, la moyenne de cette caractéristique à la moyenne globale de la caractéristique dans les données spécifiques à une classe de conditions pathologiques. Le test utilisé pour une caractéristique dans un cluster c est défini comme suit :

$$Test_{\text{value}} = \frac{\bar{x}_c - \bar{x}}{\sqrt{\frac{s^2}{n_c} \left(\frac{n-n_c}{n-1} \right)}} \quad (5.1)$$

où \bar{x}_c représente la moyenne d’une caractéristique dans un cluster spécifique c , \bar{x} est la moyenne globale de la caractéristique dans les données spécifiques à la condition pathologique, s^2 est la variance globale de la caractéristique dans les données de la condition pathologique, n_c est le nombre d’observations dans le cluster c , et n est le nombre total d’observations de la condition pathologique.

La p – *value* est calculée pour évaluer la significativité du $Test_{\text{value}}$. Une p – *value* faible indique que le test est statistiquement significatif, et donc que la caractéristique distingue le cluster. En d’autres termes, cela signifie que cette caractéristique est particulièrement informative pour ce cluster spécifique. Toutes les p – *value* obtenues étaient significatives (inférieures à 0,05). Notre attention se concentre sur les caractéristiques présentant un $Test_{\text{value}}$ positif, ce qui indique que la moyenne du cluster dépasse la moyenne globale. Ce critère montre l’importance de telles caractéristiques, soulignant leur rôle essentiel dans la distinction entre les différentes classes de conditions pathologiques dans la base de données.

En combinant les caractéristiques identifiées comme importantes pour chaque clus-

ter de la même condition pathologique, nous obtenons un ensemble de caractéristiques distinctives pour cette condition. Ces caractéristiques sélectionnées permettent alors de différencier efficacement entre les différentes classes dans notre étude.

Pour identifier les caractéristiques les plus significatives, nous avons appliqué l'approche suivante : Tout d'abord, nous avons classé les valeurs du $Test_{\text{value}}$ des caractéristiques par ordre décroissant. Ce classement nous permet de prioriser les caractéristiques en fonction de leur importance statistique. Ensuite, nous avons calculé les différences entre les valeurs du $Test_{\text{value}}$ consécutives, c'est-à-dire $e_i = V_i - V_{i+1}$. Cette étape aide à identifier les sauts significatifs entre les valeurs du $Test_{\text{value}}$. Enfin, nous avons déterminé l'indice K où la différence e_i est maximale, indiquant les points où la significativité des caractéristiques change de manière remarquable, c'est-à-dire $K = \text{indice de } \max(e_i)$. En sélectionnant les K premières caractéristiques, nous nous assurons de retenir celles qui présentent la plus grande significativité selon le $Test_{\text{value}}$.

Par conséquent, ce processus de sélection garantit que les caractéristiques sélectionnées sont très pertinentes et contribuent de manière significative à la différenciation des conditions pathologiques spécifiques de l'ensemble de données.

Pour la classification multi-label des deux maladies, l'effusion et l'infiltration, nous avons défini trois ensembles de caractéristiques : Le premier ensemble comprend des caractéristiques provenant d'images ne contenant que de l'effusion. Le deuxième ensemble comprend des caractéristiques provenant d'images présentant uniquement de l'infiltration. Le troisième ensemble est composé de caractéristiques provenant d'images contenant à la fois de l'effusion et de l'infiltration.

Ensuite, nous avons considéré deux scénarios pour la sélection des caractéristiques :

- Scénario 1 : Nous avons considéré toutes les caractéristiques identifiées comme les plus significatives pour les trois ensembles afin de réaliser la classification multi-label des maladies.
- Scénario 2 : Nous avons sélectionné les caractéristiques les plus significatives pour les trois ensembles en excluant les caractéristiques communes entre eux pour la classification multi-label des maladies.

En appliquant cette approche de sélection de caractéristiques, nous avons identifié 47 caractéristiques comme les plus significatives dans le premier scénario, et 19 caractéristiques dans le deuxième scénario, pour la classification multi-label des deux maladies : l'effusion et l'infiltration. La Table 5.3 présente les caractéristiques sélectionnées.

Grâce à notre approche de sélection de caractéristiques, nous avons pu identifier les

TABLE 5.3 – Caractéristiques sélectionnées pour les différentes conditions pathologiques.

Condition	Selected features
Effusion	Small Area Emphasis, Small Dependence Emphasis, Zone Percentage
Infiltration	Kurtosis, asm3, Contrast2, energy4, asm4, Contrast1, asm1, energy3, energy2, Contrast3, asm2, Contrast4, energy1
Effusion+Infiltration	Mean, Median, Root Mean Squared
Effusion Infiltration	Large Area Emphasis, Zone Variance, Large Dependence Emphasis
Effusion Effusion+Infiltration	Minimum, 10Percentile
Infiltration Effusion+Infiltration	Total Energy, homogeneity2, dissimilarity3, dissimilarity2, Gray Level Non Uniformity3, homogeneity3, dissimilarity4, dissimilarity1, homogeneity1, homogeneity4
Effusion Infiltration Effusion+Infiltration	Dependence Variance, Run Variance, Run Percentage, Run Entropy, Zone Entropy, Size Zone Non Uniformity, Gray Level Non Uniformity2, Short Run Emphasis, Long Run Emphasis, Dependence Entropy, Dependence Non Uniformity, Run Length Non Uniformity, Gray Level Non Uniformity

caractéristiques les plus significatives pour chaque condition pathologique de manière spécifique. Par exemple, les caractéristiques de la première ligne de la Table 5.3 se sont avérées particulièrement pertinentes pour la classe correspondant à l'effusion (c'est-à-dire, les images montrant la maladie de l'effusion), tandis qu'elles n'étaient pas pertinentes pour les images présentant la maladie de l'infiltration. De même, les caractéristiques de la deuxième ligne étaient plus significatives pour la maladie de l'infiltration. Les caractéristiques telles que "Mean", "Median" et "Root Mean Squared" étaient spécifiquement les plus importantes pour les images présentant les deux maladies à la fois. En outre, les caractéristiques communes aux images présentant l'effusion et à celles présentant l'infiltration ont été identifiées, comme indiqué dans la quatrième ligne de la Table 5.3. Les caractéristiques "Minimum" et "10Percentile" ont été identifiées comme les plus significatives pour les images présentant l'effusion tout en conservant leur importance pour les images présentant les deux maladies à la fois. Les caractéristiques mentionnées dans la sixième ligne de la Table 5.3 se sont révélées les plus significatives pour les images présentant l'infiltration et sont restées très importantes pour les images contenant les deux maladies simultanément. Enfin, les caractéristiques de la dernière ligne de la Table 5.3 ont été considérées comme les plus significatives pour les images présentant l'effusion, les images présentant l'infiltration, ainsi que pour celles présentant les deux maladies simultanément.

Ainsi, avec notre approche de sélection de caractéristiques, nous avons pu identifier les caractéristiques les plus significatives spécifiques à chaque maladie tout en distinguant celles qui restent pertinentes lorsque les deux maladies sont présentes simultanément. Cela renforce la capacité des modèles de classification à distinguer avec précision les différentes conditions pathologiques, contribuant ainsi à des diagnostics plus fiables et à une meilleure compréhension des maladies et des interactions entre les différentes manifestations pathologiques.

5.7 Modèles de classification

Les caractéristiques sélectionnées ont été intégrées dans des algorithmes d'apprentissage automatique pour effectuer la tâche de classification multi-label de l'effusion et l'infiltration. Les données ont été divisées en deux ensembles, 80% étant utilisées pour l'entraînement et les 20% restants pour les tests (choisies aléatoirement). Deux algorithmes d'apprentissage automatique ont été utilisés : la forêt aléatoire ou *Random Forest* (RF) et *Extreme Gradient Boosting* (XGBoost). Ces algorithmes ont été choisis pour leur

capacité à traiter efficacement les données pour des tâches de classification.

5.7.1 Forêt aléatoire (Random Forest)

Un classificateur de forêt aléatoire ou *Random Forest* (RF) est un algorithme d'apprentissage automatique supervisé, reconnu comme une technique de classification d'ensemble. RF utilise un "assemblage parallèle" qui construit plusieurs classificateurs d'arbres de décision en parallèle sur différents sous-échantillons de l'ensemble de données, puis prend leur vote majoritaire pour déterminer la classe finale. Chaque arbre dans une forêt aléatoire échantillonne aléatoirement des sous-ensembles des données d'entraînement dans un processus appelé agrégation par bootstrap (bagging). Le modèle est ajusté à ces ensembles de données plus petits et les prédictions sont agrégées. Cela permet d'améliorer la précision de la prédiction et de minimiser le problème de surajustement par rapport à un seul arbre de décision. En effet, les modèles d'arbre de décision sont très sensibles aux données d'entraînement, et de petits changements peuvent conduire à des structures d'arbres très différents. Par conséquent, le modèle d'apprentissage RF avec plusieurs arbres de décision est généralement plus précis et plus robuste qu'un modèle basé sur un seul arbre de décision [227].

Dans notre étude, nous avons utilisé la technique "grid search" pour identifier les hyperparamètres optimaux du modèle d'apprentissage automatique. Cette technique consiste à définir un ensemble de valeurs possibles pour chaque hyperparamètre, puis à entraîner et évaluer le modèle pour chaque combinaison possible de ces valeurs. Son objectif est de trouver la meilleure combinaison d'hyperparamètres qui maximise les performances du modèle [228]. Différentes valeurs des hyperparamètres du modèle RF ont été testés pour déterminer celles qui offraient les meilleures performances avec les données expérimentales. Nous avons évalué le nombre d'arbres dans la forêt (*n_estimators*) avec les valeurs [50, 100, 200, 300] et le critère de fractionnement (*criterion*) avec les options [*gini*, *entropy*]. Le critère *gini* calcule l'impureté des noeuds et mesure la probabilité qu'un élément soit classé incorrectement lorsqu'il est choisi de manière aléatoire. Si tous les éléments sont associés à une seule classe, on peut alors dire qu'ils sont purs. L'*entropy* mesure l'impureté ou l'aléatoire dans les données. Un niveau élevé de désordre signifie un faible niveau d'impureté. Les meilleures valeurs d'hyperparamètres pour le modèle RF obtenues étaient 200 arbres avec un critère de *gini*.

5.7.2 Extreme Gradient Boosting (XGBoost)

Gradient boosting est un algorithme d'apprentissage ensembliste qui produit un modèle final à partir d'une série de modèles individuels, généralement des arbres de décision. Le gradient permet de minimiser la fonction de perte, de la même manière que les réseaux neuronaux optimisent les poids en utilisant la descente de gradient. L'algorithme est basé sur le concept de "boosting", qui combine les prédictions d'un ensemble de "faibles" apprenants pour développer un apprenant "fort" à travers un entraînement additif séquentiel. Cette approche entraîne séquentiellement un ensemble d'arbres de décision peu profonds, où chaque arbre successif est entraîné sur les résidus d'erreur du modèle précédent pour ajuster le modèle suivant. À tout instant t , les résultats du modèle sont pondérés en fonction des résultats du précédent instant $t - 1$. Les résultats prédits correctement reçoivent un poids plus faible et ceux mal classés sont pondérés plus fortement. La prédiction finale du modèle est une somme pondérée de toutes les prédictions des arbres.

Le boosting de gradient extrême (XGBoost) est un type de boosting de gradient qui minimise la perte en calculant les gradients du second ordre de la fonction de perte, ce qui améliore les capacités de généralisation et les performances du modèle et réduit le surajustement [227]. XGBoost est rapide à intégrer et peut gérer de grands ensembles de données. Le processus d'apprentissage additif dans XGBoost est expliqué ci-dessous. Le premier apprenant est ajusté à l'ensemble de données d'entrée, puis un deuxième modèle est ajusté à ces résidus pour surmonter les défauts d'un apprenant faible. Ce processus d'ajustement est répété plusieurs fois. La prédiction finale du modèle est une somme pondérée de toutes les prédictions des arbres. L'équation générale pour la prédiction à l'étape t est présentée comme suit :

$$f_i^{(t)} = \sum_{k=1}^t f_k(x_i) = f_i^{(t-1)} + f_t(x_i) \quad (5.2)$$

où $f_t(x_i)$ est l'apprenant à l'étape t , $f_i^{(t)}$ et $f_i^{(t-1)}$ sont les prédictions aux étapes t et $(t - 1)$ respectivement, et x_i est la variable d'entrée.

Les hyperparamètres par défaut pour l'algorithme XGBoost sont fournis par l'implémentation de la bibliothèque *xgboost*. Nous avons sélectionné le modèle basé sur les arbres, connu sous le nom de "gbtree", comme type de modèle à exécuter à chaque itération. Le modèle comporte 100 arbres avec une profondeur maximale de 6 niveaux pour chaque arbre. La profondeur maximale d'un arbre détermine le nombre de divisions successives qu'un arbre peut effectuer.

5.8 Résultats et discussion

Dans cette étude, nous avons visé à inclure les zones affectées par les pathologies dans la région d’intérêt des masques. Ainsi, nous avons développé une nouvelle approche de segmentation pour la génération des masques de pathologies.

Après avoir introduit des images présentant soit l’effusion, soit l’infiltration, soit les deux maladies simultanément, nous avons évalué l’impact des méthodes de génération de masques sur la classification multi-label des maladies pulmonaires. Nous avons comparé les performances des modèles d’apprentissage automatique utilisant les caractéristiques extraites à partir des masques générés par notre approche, qui est basée sur l’application de CycleGAN suivie de la soustraction d’images et de la normalisation, avec celles obtenues à partir des masques produits par U-Net.

Afin de garantir la robustesse et la stabilité des résultats, nous avons effectué dix ré-partitions aléatoires des données en ensembles d’entraînement et de test, et nous avons présenté les résultats sous forme de moyenne et d’écart type. Comme le montre la Table 5.4, les modèles RF et XGBoost ont atteint de meilleures performances en termes d’accuracy et de scores AUC en utilisant notre approche proposée pour la génération de masques pathologiques. Avec les masques générés par U-Net, les modèles RF et XGBoost ont obtenu, sur 10 divisions, une accuracy moyenne de 79,15% et 79,56%, et un score AUC moyen de 74,47% et 75,63%, respectivement. Cependant, en utilisant notre approche pour la génération de masques pathologiques, les modèles RF et XGBoost ont atteint, sur dix divisions, une accuracy moyenne de 88,26% et 89,52%, ainsi qu’un score AUC moyen de 85,54% et 86,55%, respectivement.

TABLE 5.4 – Impact de la méthode de génération de masques sur les résultats de classification multi-label de l’effusion et de l’infiltration. Les valeurs importantes sont en gras.

Mask generation approach	Accuracy (%)		AUC (%)	
	RF	XGBoost	RF	XGBoost
With U-Net	79.15±2.5	79.56±2.3	74.47±2.4	75.63±2.5
With our proposed method using CycleGAN, image subtraction and binarization	88.26±1.8	89.52±1.7	85.54±1.9	86.55±1.8

En effet, l’utilisation des masques générés par U-Net s’est avérée moins efficace, principalement parce que ces masques sont limités aux parties saines des poumons et négligent les zones pathologiques essentielles pour une caractérisation adéquate des maladies. Cette

limitation est probablement due au fait que U-Net est entraîné sur des données saines, ce qui limite sa capacité à détecter et segmenter les zones affectées par la maladie. Notre approche se distingue en garantissant que les masques incluent les régions d'intérêt affectées par des pathologies, permettant ainsi l'extraction des caractéristiques de ces zones critiques. Cette approche améliore la qualité et la pertinence des caractéristiques extraites, conduisant à une classification multi-label plus précise des maladies d'effusion et d'infiltration. Cela souligne l'avantage significatif de notre approche par rapport aux méthodes de segmentation conventionnelles. Par conséquent, pour la suite de notre étude, notre approche sera utilisée pour la génération de masques.

En utilisant les masques générés qui définissent avec précision les ROIs affectées par les pathologies et les images originales transformées en niveaux de gris, nous avons extrait des caractéristiques radiomiques, notamment les statistiques de premier ordre, GLCM, GLSZM, GLRLM, NGTDM et GLDM. Afin de sélectionner les caractéristiques les plus significatives pour chaque maladie, ainsi que dans les cas où les deux maladies sont présentes simultanément, nous avons appliqué notre approche de sélection de caractéristiques proposée, débutant par une HCA, suivie d'une évaluation rigoureuse à l'aide du T_{est_value} . Pour évaluer l'efficacité de la sélection de caractéristiques dans le processus de classification, nous avons analysé les performances des modèles avec et sans l'utilisation de notre approche de sélection de caractéristiques proposée avec les scénarios 1 et 2. Les résultats présentés dans la Table 5.5 indiquent clairement que l'application de la sélection de caractéristiques (scénario 2) en excluant les caractéristiques communes entre les ensembles conduit à des améliorations significatives des résultats du modèle de classification, soulignant ainsi l'importance de cette étape dans l'amélioration des performances de classification.

TABLE 5.5 – Comparaison des résultats de la classification multi-label de l'effusion et de l'infiltration en utilisant notre approche de génération de masque avec et sans sélection de caractéristiques. Les valeurs importantes sont en gras.

Our proposed mask generation approach	Accuracy (%)		AUC (%)	
	RF	XGBoost	RF	XGBoost
Without feature selection	88.26±1.8	89.52±1.7	85.54±1.9	86.55±1.8
With feature selection (Scenario 1)	89.32±1.6	90.59±1.5	87.21±1.7	87.83±1.5
With feature selection (Scenario 2)	91.16±1.4	92.05±1.3	88.92±1.5	89.48±1.4

Bien que les modèles RF et XGBoost aient initialement démontré de bonnes performances sans la sélection de caractéristiques, l'intégration de notre approche de sélection

de caractéristiques a significativement amélioré à la fois l’accuracy et le score AUC de ces modèles. En effet, sans sélection de caractéristiques, les modèles RF et XGBoost ont atteint une accuracy moyenne de 88,26% et 89,52%, et des scores AUC de 85,54% et 86,55%, respectivement. Avec l’application de la sélection de caractéristiques (scénario 1), l’accuracy des modèles s’est améliorée pour atteindre 89,32% pour RF et 90,59% pour XGBoost, tandis que les scores AUC ont augmenté à 87,21% pour RF et 87,83% pour XGBoost. De même, avec l’utilisation de la sélection de caractéristiques (scénario 2), les modèles RF et XGBoost ont atteint une accuracy de 91,16% et 92,05%, et un score AUC de 88,92% et 89,48%, respectivement.

Ces augmentations sont non seulement statistiquement significatives mais elles sont également cliniquement pertinentes. En effet, même de légères améliorations de l’accuracy et du score AUC peuvent avoir un impact considérable sur la fiabilité de la classification des maladies pulmonaires. Les résultats démontrent clairement que la sélection de caractéristiques (scénario 2), associée à notre approche de génération de masques, améliore significativement les performances des modèles dans la classification multi-label des maladies pulmonaires de 2,93%. Cela souligne l’importance cruciale d’une sélection rigoureuse des caractéristiques, mettant en évidence l’avantage de notre approche sur les résultats de classification.

Globalement, les résultats de cette étude démontrent l’efficacité des modèles d’apprentissage automatique, en particulier du modèle XGBoost suivi du modèle RF, dans la classification des pathologies pulmonaires. Bien que les deux modèles affichent des performances élevées, des marges d’amélioration existent. Le modèle XGBoost s’est révélé être le plus performant des deux grâce à sa méthode de boosting avancée. Contrairement à Random Forest, qui génère un grand nombre d’arbres de décision indépendants et s’appuie sur le vote majoritaire de ces arbres pour la classification finale, XGBoost adopte une approche itérative. Fondé sur le concept de boosting, XGBoost vise à minimiser la fonction de perte en intégrant séquentiellement des apprenants faibles, améliorés grâce à la descente de gradient. Cette stratégie permet à XGBoost de corriger continuellement ses prédictions en accordant une attention particulière aux instances mal classées lors des itérations précédentes, ce qui améliore sa capacité à identifier avec précision les pathologies pulmonaires sur les images de radiographie thoracique. En conséquence, le modèle XGBoost est fortement recommandé pour la classification des maladies pulmonaires à partir d’images de radiographie thoracique. Son utilisation représente une avancée significative

dans le domaine du diagnostic médical assisté par ordinateur, offrant aux professionnels de la santé un outil robuste pour une classification précise des maladies pulmonaires.

Les résultats présentés dans la Table 5.6 montrent que le modèle XGBoost a atteint des performances élevées dans la classification multi-label de l'effusion et de l'infiltration. Cette capacité à identifier de manière précise les maladies est cruciale dans le domaine médical, car elle joue un rôle déterminant dans l'élaboration des protocoles de traitement, contribuant à améliorer la qualité de soins des patients. Ces résultats soulignent l'importance de notre approche, combinant une méthode avancée de génération de masques pathologiques avec une sélection rigoureuse des caractéristiques. Cette combinaison permet non seulement d'affiner les performances de classification, mais aussi de renforcer la fiabilité des décisions cliniques, offrant ainsi un soutien précieux aux professionnels de la santé dans la gestion des maladies pulmonaires.

TABLE 5.6 – Résultats détaillés du modèle XGBoost utilisant notre approche de génération de masques avec sélection de caractéristiques (scénario 2), pour la classification multi-label de l'effusion et de l'infiltration.

Label	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)	AUC (%)
Effusion	90.68±1.4	96.38±1.2	93.45±1.3	91.06±1.4	88.53±1.4
Infiltration	92.95±1.3	97.29±1.1	95.07±1.2	93.04±1.2	90.42±1.3

Nous avons évalué l'efficacité de notre approche de sélection de caractéristiques en appliquant la méthode SHAP (*SHapley Additive exPlanations*) au modèle XGBoost entraîné avec l'ensemble complet des caractéristiques. La méthode SHAP est utilisée pour expliquer la sortie d'un modèle d'apprentissage automatique en calculant la contribution de chaque caractéristique à la prédiction. Par conséquent, elle permet d'évaluer comment la contribution de chaque caractéristique affecte le modèle [229]. L'importance des caractéristiques SHAP est calculée comme la moyenne des valeurs absolues de Shapley. L'idée de l'importance des caractéristiques SHAP est que les caractéristiques importantes sont celles ayant de grandes valeurs absolues de Shapley. La Figure 5.9 (a) et (b) illustre les caractéristiques identifiées par SHAP comme étant les plus influentes dans la classification des maladies d'effusion et d'infiltration, respectivement. Ces caractéristiques sont classées par ordre décroissant d'importance selon leur impact sur la classification. Ces visualisations offrent une compréhension approfondie des caractéristiques guidant les décisions de classification du modèle XGBoost pour chaque maladie. Ainsi, du côté médical, les spécialistes et les médecins peuvent interpréter les variables et savoir quelles caractéristiques

sont les plus importantes pour identifier et classer les maladies d’effusion et d’infiltration.

Cette analyse SHAP révèle non seulement les caractéristiques les plus importantes, mais valide également l’efficacité de notre approche de sélection de caractéristiques en démontrant une concordance entre les caractéristiques sélectionnées a priori par notre approche et celles identifiées par SHAP comme ayant le plus grand impact sur les sorties du modèle de classification. Ainsi, notre approche confirme sa capacité à sélectionner les caractéristiques les plus pertinentes pour la classification des maladies pulmonaires, améliorant la précision et la fiabilité des diagnostics, et contribuant à des traitements mieux ciblés.

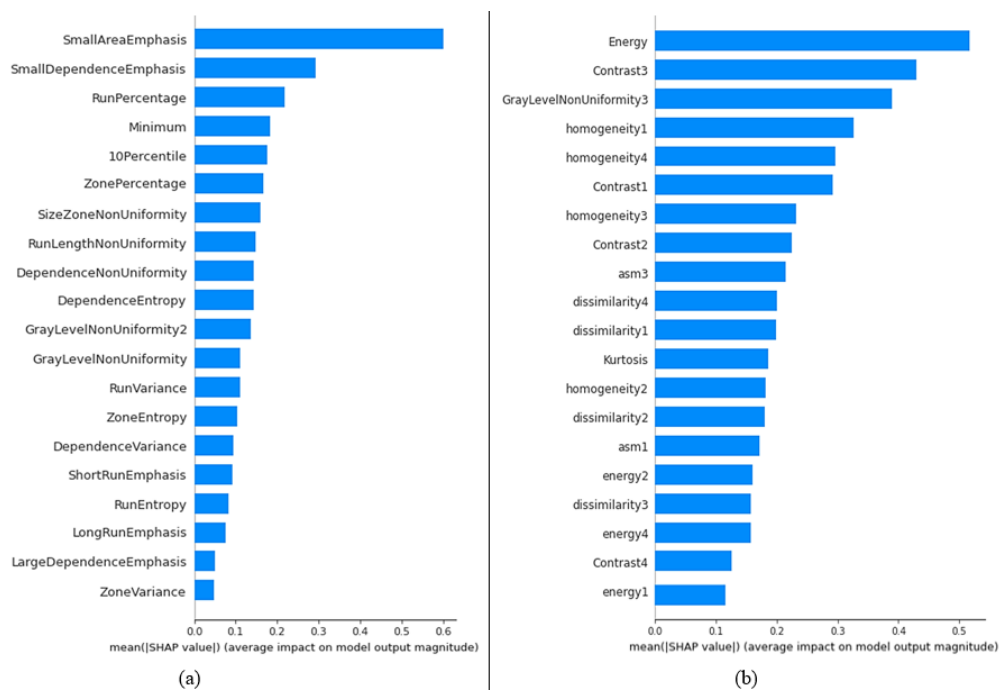


FIGURE 5.9 – Résultats du modèle SHAP montrant les caractéristiques les plus importantes pour la classification multi-label des deux maladies : (a) Effusion, (b) Infiltration.

Pour généraliser notre approche, nous l’avons appliqué pour la classification multi-label des 14 pathologies de la base de données ChestX-ray14. Notre approche de génération de masques a été appliquée pour inclure les zones pathologiques dans les masques afin d’extraire des caractéristiques plus précises et pertinentes.

Un aspect essentiel de notre approche a été la sélection rigoureuse des caractéristiques les plus significatives pour chacune des 14 maladies. Dans le cas des images montrant

des combinaisons de pathologies, c'est-à-dire présentant deux ou plusieurs maladies simultanément, nous nous sommes limités uniquement aux combinaisons représentées par plus de 100 images. Cette restriction nous a permis de garantir une taille d'échantillon suffisamment grande pour assurer la fiabilité et la robustesse de notre analyse.

Par conséquent, notre approche nous a permis de sélectionner 58 caractéristiques jugées les plus significatives pour l'analyse en considérant le premier scénario, et 15 caractéristiques pour le scénario 2 qui exclut les caractéristiques communes entre les classes. Cependant, l'utilisation du scénario 2 n'a pas permis de différencier efficacement les différentes maladies. En effet, il n'y avait pas de caractéristiques spécifiques uniques aux différentes classes de maladies, car elles partageaient de nombreuses caractéristiques communes. Par conséquent, pour la classification des 14 maladies, nous avons utilisé notre approche de sélection des caractéristiques selon le scénario 1. Cette sélection ciblée des caractéristiques est cruciale pour une classification précise et fiable des maladies thoraciques, nous permettant d'optimiser la performance de notre modèle de classification.

La Figure 5.10 illustre les courbes ROC de notre méthode proposée appliquée aux 14 maladies de la base de données ChestX-ray14. Ces courbes fournissent une évaluation quantitative de la performance du modèle, en mettant en évidence sa capacité à discriminer précisément entre la présence et l'absence de pathologies spécifiques. Cette capacité est essentielle pour assurer des diagnostics précis et orienter de manière efficace les choix de traitement pour les maladies thoraciques.

Les résultats de classification obtenus par notre méthode pour les 14 maladies sont comparés à ceux des approches les plus récentes de la littérature dans la Table 5.7. Notre méthode a atteint un score AUC moyen de 83,12%, surpassant les méthodes existantes. Notre méthode a obtenu les scores AUC les plus élevés pour 7 des 14 maladies (indiqués en gras), tandis que les scores pour les 7 autres maladies restent compétitifs avec les meilleurs résultats actuels. Plus précisément, notre méthode surpasse les études précédentes avec des améliorations significatives, telles que pour l'atelectasis (3.66%), consolidation (8.25%), effusion (1.92%), infiltration (13.91%), nodule (2.93%), pleural_thickening (6.41%) et pneumonia (4,25%). Ces améliorations soulignent l'efficacité de notre approche dans le raffinement de la classification des maladies thoraciques, offrant ainsi des perspectives prometteuses pour améliorer la précision du diagnostic et optimiser la gestion clinique des maladies thoraciques, ce qui pourrait avoir une influence significative sur les protocoles de traitement et les résultats pour les patients.

Ainsi, notre étude, qui combine une technique avancée de génération de masques et

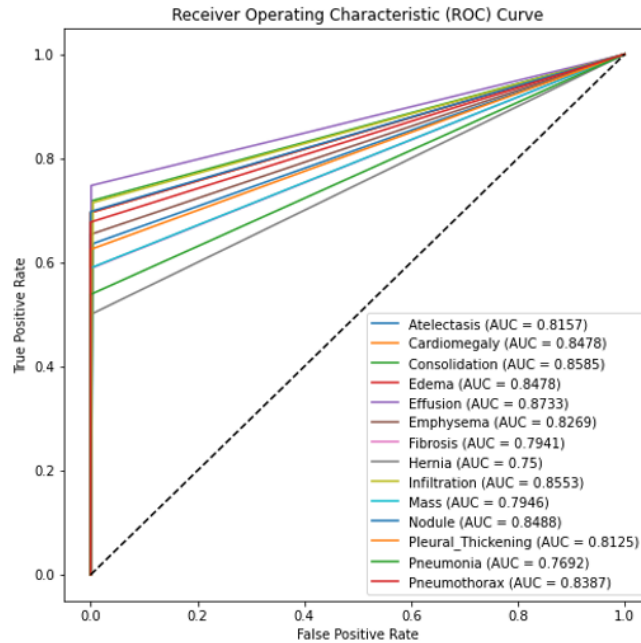


FIGURE 5.10 – Courbes ROC et scores AUC de toutes les maladies sur l’ensemble de test de la base de données ChestX-ray14 utilisant le modèle XGBoost avec notre approche de génération de masques et de sélection des caractéristiques.

TABLE 5.7 – Comparaison des résultats AUC de notre méthode avec d’autres méthodes existantes dans la littérature pour la classification multi-label des maladies thoraciques de la base de données ChestX-ray14. Les valeurs importantes sont en gras.

Abnormality	Guendel et al. (2018) [123]	Ma et al. (2019) [184]	Guan et al. (2020) [173]	Gündel et al. (2021) [230]	Chen et al. (2022) [188]	Yang et al. (2022) [193]	Our approach
Atelectasis	76.7%	77.7%	78.1%	78.5%	75.9%	78.8%	82.46%±1.5%
Cardiomegaly	88.3%	89.4%	88%	89.2%	89.83%	87.5%	85.13%±2.2%
Consolidation	74.5%	75%	75.4%	74.7%	75.12%	75.6%	83.85%±1.7%
Edema	83.5%	84.6%	85%	83.7%	86.36%	85.4%	84.41%±2.4%
Effusion	82.8%	82.9%	82.9%	83.6%	82.66%	83.7%	85.62%±1.4%
Emphysema	89.5%	90.8%	90.8%	92.5%	88.98%	93.4%	84.44%±2.3%
Fibrosis	81.8%	82.7%	83%	83.8%	83.59%	84.9%	82.07%±2.1%
Hernia	89.6%	93.4%	91.7%	90.5%	94.04%	94.4%	76.93%±2.9%
Infiltration	70.9%	69.6%	70.2%	71%	70.99%	72.2%	86.11%±1.3%
Mass	82.1%	83.8%	83.4%	82.6%	83.2%	84.5%	82.86%±1.6%
Nodule	75.8%	77.1%	77.3%	75.5%	75.33%	79.6%	82.53%±1.5%
Pleural_Thickening	76.1%	77.9%	77.8%	78.5%	78.94%	78.9%	85.35%±1.7%
Pneumonia	73.1%	72.2%	72.9%	73.5%	73.56%	73.4%	77.81%±1.4%
Pneumothorax	84.6%	86.2%	85.7%	84.7%	87.6%	86.2%	84.16%±1.8%
Average AUC	80.7%	81.7%	81.6%	81.59%	81.86%	82.7%	83.12%±1.8%

une approche robuste de sélection de caractéristiques, améliore efficacement la précision de la classification des pathologies thoraciques à partir d’images de radiographie thoracique. Cela contribue de manière significative à l’amélioration des techniques de diagnostic assisté par ordinateur, avec un impact potentiellement important sur les stratégies de traitement et la gestion des patients souffrant de maladies thoraciques.

5.9 Conclusion

Dans ce chapitre, nous avons présenté deux contributions visant à améliorer la performance de la classification multi-label des pathologies thoraciques.

Notre première contribution réside dans le développement d’une approche de génération de masques, conçue pour inclure les zones affectées par la pathologie dans la région d’intérêt. Notre approche commence par l’application du modèle CycleGAN aux images de radiographie thoracique pour générer une version restaurée et saine des poumons. Ensuite, une soustraction d’images est effectuée entre l’image restaurée et l’image originale, ce qui permet de produire une carte des zones affectées par des anomalies dans l’image. Cette carte est ensuite convertie en un format binaire, créant ainsi un masque très précis des pathologies identifiées. Cette méthode optimise l’extraction des caractéristiques radiomiques à partir des régions explicitement liées aux pathologies, améliorant ainsi de manière significative la performance de la classification des maladies. En effet, notre méthode a amélioré l’AUC du modèle XGBoost de 10,92% par rapport à l’utilisation du modèle U-Net pour la génération des masques.

La deuxième contribution de ce chapitre repose sur une approche de sélection de caractéristiques visant à identifier a priori les caractéristiques les plus significatives et pertinentes pour chaque condition pathologique. Notre approche commence par un clustering non supervisé pour regrouper les caractéristiques extraites en fonction de leurs similarités. Ensuite, les clusters obtenus sont soumis à une évaluation rigoureuse à l’aide d’un test statistique ($Test_{value}$), permettant de sélectionner les caractéristiques statistiquement les plus significatives. Cette approche améliore l’efficacité du modèle de classification en ciblant avec précision les caractéristiques les plus pertinentes des pathologies, ce qui optimise la performance globale du modèle et renforce la fiabilité des diagnostics.

En intégrant ces deux contributions, notre pipeline offre des résultats prometteurs sur la base de données ChestX-ray14, surpassant les résultats des études précédentes avec des améliorations significatives. Notre approche non seulement optimise les performances du

modèle de classification, mais elle renforce également la précision du diagnostic médical. Les médecins peuvent ainsi prendre des décisions diagnostiques et thérapeutiques plus précises, améliorant l'efficacité des traitements et potentiellement les résultats pour les patients.

Ce chapitre a fait l'objet d'un article soumis à la revue "*Artificial Intelligence In Medicine*" et d'une présentation acceptée à la conférence internationale "*IEEE IPAS 2025*".

CONCLUSION ET PERSPECTIVES

L'objectif principal de cette thèse est de développer des modèles d'apprentissage automatique pour améliorer la classification des maladies pulmonaires à partir des images de radiographie thoracique. Pour ce faire, plusieurs approches de prétraitement sont utilisées afin d'améliorer la qualité des images, permettant ainsi une classification automatique des pathologies avec plus de précision.

Pour la mise en œuvre efficace des algorithmes d'apprentissage automatique dans le domaine de la radiologie, il est important d'inclure non seulement les images radiologiques elles-mêmes, mais aussi des métadonnées telles que la position de vue de la radiographie, l'âge et le genre du patient. Ces informations supplémentaires offrent un contexte précieux qui peut considérablement influencer l'interprétation des images et la prise de décision concernant la présence d'une pathologie. En intégrant ces données contextuelles, les modèles d'apprentissage automatique peuvent être affinés selon le profil du patient, ce qui mène à des décisions thérapeutiques plus précises et spécifiques, améliorant ainsi l'efficacité diagnostique et la qualité des soins des patients. Cette première contribution, présentée dans le chapitre 3, débute par l'utilisation du clustering non supervisé pour identifier les images des patients présentant des caractéristiques similaires, suivie de l'analyse des données cliniques de chaque cluster. Une variation significative a été observée dans la position de vue des images, formant ainsi deux clusters : l'un majoritairement composé d'images en vue postéro-antérieure et l'autre en vue antéro-postérieure. Cette distinction est cliniquement significative car la position de vue peut affecter la visibilité des structures anatomiques et la détection des anomalies, ce qui est crucial pour un diagnostic précis. Pour améliorer le contraste des images dans chaque groupe, nous avons appliqué la technique du Masque Flou. De plus, nous avons renforcé notre modèle CNN avec un mécanisme d'attention afin de se concentrer sur les régions critiques de l'image, garantissant ainsi une classification précise des pathologies pulmonaires. Notre approche a amélioré l'AUC de la classification binaire de la pneumonie de 2,35% par rapport au modèle appliqué sur toutes les images. Par conséquent, l'application du modèle à des groupes partageant des caractéristiques cliniques communes est cruciale pour maximiser sa pertinence et son efficacité. Cette approche est essentielle pour la prise de décision clinique car elle permet

de développer des stratégies de prévention plus adaptées à chaque sous-groupe, ce qui conduit à des diagnostics plus précis et fiables, améliorant ainsi la qualité des soins aux patients.

Ensuite, les images de radiographie thoracique peuvent contenir des artefacts, tels que des objets, des dispositifs médicaux, des fils et des électrodes. Ces éléments introduisent du bruit dans les images et peuvent obscurcir des structures anatomiques essentielles, compliquant ainsi l'identification précise des pathologies et menant à des erreurs de diagnostic et des traitements inappropriés pour les patients. Pour remédier à ce problème, nous avons développé, dans le chapitre 4, une méthode de prétraitement visant à minimiser l'effet du bruit produit par ces artefacts sur les images de radiographie thoracique. Étant donné que les images contenant des artefacts ne sont pas étiquetées dans la base de données, nous avons extrait des caractéristiques liées aux contours et à la texture des images afin de réaliser un clustering non supervisé. Cela a permis de créer deux clusters : l'un comprenant des images avec artefacts et l'autre des images nettes sans artefacts. Pour atténuer l'impact des artefacts dans les images du premier cluster, nous avons généré des images plus nettes en utilisant le modèle CycleGAN. Ces images générées ont ensuite été combinées avec celles du second cluster pour effectuer la classification à l'aide du modèle DenseNet-121, en intégrant le CBAM (Convolutional Block Attention Module) et en ajoutant les variables cliniques. Les résultats expérimentaux ont démontré l'efficacité de notre approche, avec une amélioration de l'AUC de 5,91% pour la classification binaire de la pneumonie et de 6,44% pour la classification binaire de la consolidation, par rapport au modèle de base appliqué sur toutes les images originales. Par conséquent, l'identification précise des images contenant des artefacts et la réduction de leur impact permettent au modèle de se concentrer sur les caractéristiques diagnostiques pertinentes, améliorant ainsi la précision de la classification des maladies pulmonaires. En conséquence, les radiologues peuvent prendre des décisions thérapeutiques plus efficaces et offrir des soins de meilleure qualité aux patients souffrant de maladies pulmonaires.

Il est fréquent que les images de radiographie thoracique révèlent la présence simultanée de plusieurs pathologies. Par conséquent, la construction d'un seul modèle capable de classer plusieurs pathologies présentes dans une seule image de radiographie thoracique peut considérablement soulager la charge de travail des spécialistes. Sans cela, ils devraient utiliser plusieurs systèmes distincts pour diagnostiquer différentes maladies. Cette approche peut significativement réduire le temps nécessaire à la prise de décisions cliniques, améliorant ainsi l'efficacité du diagnostic. Ainsi, dans le chapitre 5, nous avons

abordé la classification multi-label des maladies pulmonaires en utilisant une approche d'ingénierie des caractéristiques. Les caractéristiques radiomiques extraites à partir des images quantifient de manière précise les particularités propres aux pathologies. En exploitant ces caractéristiques radiomiques, les médecins peuvent obtenir une compréhension plus fine de la signification biologique et clinique des tumeurs, permettant des diagnostics et des stratégies de traitement plus précis. Cependant, une analyse radiomique efficace nécessite une segmentation précise et une génération de masques délimitant précisément les régions d'intérêt (ROIs). Les approches traditionnelles telles que U-Net, bien que largement utilisées pour la segmentation des images médicales, échouent à inclure les régions spécifiquement affectées par les pathologies dans les masques, ce qui est crucial pour une analyse radiomique exacte. Afin de résoudre ce problème, pour le prétraitement, nous avons développé une approche de génération de masques qui inclue les zones affectées par la pathologie dans la région d'intérêt. Nous avons d'abord appliqué le modèle CycleGAN aux images de radiographie thoracique pour générer une version restaurée et saine des poumons. Ensuite, nous avons soustrait l'image originale de l'image restaurée, produisant ainsi une carte des zones affectées par des anomalies dans l'image. Cette carte est convertie en un format binaire, créant ainsi un masque très précis des pathologies identifiées. Notre approche garantit que les caractéristiques sont extraites des zones les plus indicatives des pathologies définies par les masques, permettant ainsi de concentrer l'analyse sur les aspects pathologiques des images. Par conséquent, notre approche améliore de manière significative la performance de la classification des maladies pulmonaires par rapport à l'utilisation du modèle U-Net pour la génération des masques. De plus, nous avons développé une méthode de sélection des caractéristiques pour identifier a priori les caractéristiques les plus pertinentes et améliorer la performance des modèles d'apprentissage automatique. Cette approche, grâce à sa capacité à identifier les caractéristiques les plus importantes pour la classification multi-label des maladies pulmonaires, apporte une contribution notable à la pratique médicale. Elle souligne l'importance d'une sélection rigoureuse des caractéristiques pour distinguer différentes conditions pulmonaires et permettre une classification plus précise des maladies pulmonaires. En conséquence, les médecins peuvent prendre des décisions diagnostiques et thérapeutiques plus précises et efficaces, améliorant ainsi l'efficacité des traitements et la qualité des soins pour les patients.

Par ces contributions relatives à l'analyse d'imagerie de radiographie thoracique, nous espérons apporter un soutien essentiel aux médecins dans leur processus décisionnel et de

renforcer l'efficacité des interventions médicales. Nous visons à permettre une classification plus précise et efficace des pathologies thoraciques, ce qui contribuera à optimiser la qualité des diagnostics et à améliorer la prise en charge des patients.

Notre recherche ouvert de nombreuses perspectives pour des travaux futurs. Une direction potentielle consiste à explorer des architectures alternatives, telles que les réseaux de neurones sur graphes ou *Graph Neural Networks* (GNN), pour la classification multi-label des images de radiographie thoracique. Ces modèles offrent la possibilité d'exploiter les corrélations et les dépendances entre les différentes étiquettes de pathologie, ce qui pourrait considérablement améliorer la capacité du modèle à effectuer des classifications précises pour les images présentant plusieurs pathologies simultanément.

En plus, il est crucial d'aborder le problème du déséquilibre des données, comme c'est le cas dans la base de données ChestX-ray14. Les modèles de diffusion (*diffusion models*) peuvent être utilisés pour générer de nouvelles images de radiographies thoraciques pour les classes sous-représentées. Contrairement aux GANs, qui peuvent parfois souffrir de problèmes de répétition d'images similaires et de manque de variété, les modèles de diffusion adoptent une méthode graduelle. Cette méthode consiste à modifier progressivement une image en ajoutant du bruit à chaque étape, ce qui permet d'ajouter des détails subtils tout en préservant les caractéristiques importantes de l'image. Cette approche contrôlée vise à produire des images de haute qualité en ajoutant progressivement du bruit et préservant les détails, contribuant ainsi à équilibrer de manière plus efficace les données d'entraînement, notamment pour les classes sous-représentées.

En outre, il est essentiel d'étudier la performance des modèles sur d'autres types de modalités d'imagerie telles que la tomodensitométrie (CT-scan). Cela pourrait fournir des outils diagnostiques supplémentaires et augmenter les chances de détection précoce des maladies pulmonaires. En effet, les CT-scans fournissent des images avec une résolution plus élevée et des détails plus fins des structures internes, permettant une meilleure visualisation des anomalies pulmonaires. L'utilisation des CT-scans en complément des radiographies thoraciques peut ainsi fournir une vue d'ensemble plus complète, contribuant à une gestion plus efficace des maladies pulmonaires et à des soins de santé plus précis.

De plus, l'incorporation des données cliniques en plus de l'analyse d'image présente un potentiel considérable pour renforcer l'efficacité des modèles, permettant ainsi d'offrir une vue d'ensemble plus complète de la santé du patient. Par exemple, les antécédents médicaux du patient (tels que les maladies antérieures), les facteurs de mode de vie (tels

que le tabagisme, la consommation d'alcool, le régime alimentaire) ainsi que les conditions préexistantes (comme le diabète, l'hypertension ou les maladies cardiovasculaires) peuvent avoir une influence significative sur la manifestation et la progression des maladies pulmonaires. En tenant compte de ces facteurs, les modèles pourraient offrir des évaluations plus précises et adaptées selon le profil du patient, améliorant ainsi la qualité des soins et la prise en charge des patients.

Par ailleurs, il est important d'explorer des techniques supplémentaires d'extraction de caractéristiques pour améliorer la capacité du modèle à identifier les artefacts dans les images radiographiques thoraciques et à optimiser ses performances pour la classification multi-label des maladies pulmonaires.

LISTE DE PUBLICATIONS

[1] Hage Chehade, A., Abdallah, N., Marion, JM., Oueidat, M. & Chauvet, P., Lung and colon cancer classification using medical imaging : a feature engineering approach. *Physical and Engineering Sciences in Medicine* 45, 729–746, 2022. DOI : 10.1007/s13246-022-01139-x

[2] Hage Chehade, A., Abdallah, N., Marion, J. M., Hatt, M., Oueidat, M. & Chauvet, P., A Systematic Review : Classification of Lung Diseases from Chest X-Ray Images Using Deep Learning Algorithms. *SN Computer Science*, 5(4), 405, 2024. DOI : 10.1007/s42979-024-02751-2

[3] Hage Chehade, A., Abdallah, N., Marion, JM., Chéhadé, K., Oueidat, M. & Chauvet, P., Improving Lung Disease Classification from Chest X-ray Images using an Efficient Clustering Approach. *ICBES 2024*. DOI : 10.11159/icbes24.142

[4] Hage Chehade, A., Abdallah, N., Marion, J. M., Oueidat, M. & Chauvet, P., CycleGAN-Based Approach for Enhanced Segmentation of Lung Disease Areas in Chest X-Ray Images. Article accepté par la conférence IEEE IPAS 2025.

[5] Hage Chehade, A., Abdallah, N., Marion, JM., Oueidat, M. & Chauvet, P., Evaluating the Impact of View Position in X-Ray Imaging for the Classification of Lung Diseases. Article soumis à *IEEE Transactions on Biomedical Engineering*.

[5] Hage Chehade, A., Abdallah, N., Marion, JM., Hatt, M., Oueidat, M. & Chauvet, P., Advancing Chest X-Ray Diagnostics : A Novel CycleGAN-Based Preprocessing Method for Enhanced Lung Disease Classification in ChestX-Ray14. Article soumis à *Computer Methods and Programs in Biomedicine*.

[6] Hage Chehade, A., Abdallah, N., Marion, JM., Hatt, M., Oueidat, M. & Chauvet, P., CycleGAN-Based Chest X-ray Image Segmentation Approach for Enhanced Multi-Label Chest Disease Classification. Article soumis à *Artificial Intelligence In Medicine*.

LISTE DES ABRÉVIATIONS

AI	Artificial Intelligence
AP	AnteroPosterior
AUC	Area Under the Curve
CAD	Computer-Aided Diagnosis
CBAM	Convolutional Block Attention Module
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Network
CT	Computed Tomography
CXR	Chest X-Ray
CycleGAN	Cycle-Consistent Generative Adversarial Network
DL	Deep Learning
FN	False Negative
FP	False Positive
GAN	Generative Adversarial Network
GLCM	Gray Level Co-occurrence Matrix
GLDM	Gray Level Dependence Matrix
GLRLM	Gray Level Run Length Matrix
GLSZM	Gray Level Size Zone Matrix
GNN	Graph Neural Network
HE	Histogram Equalization
HEF	High-Frequency Emphasis Filtering
HCA	Hierarchical Cluster Analysis
HOG	Histogram of Oriented Gradients
JSRT	Japanese Society of Radiological Technology
MC	Montgomery County X-ray
ML	Machine Learning
MLP	Multi Layers Perceptron
MRI	Magnetic Resonance Imaging

NLP	Natural Language Processing
NGTDM	Neighbouring Gray Tone Difference Matrix
PA	PosteroAnterior
RF	Random forest
ROI	Region Of Interest
SHAP	SHapley Additive exPlanations
TB	Tuberculosis
TN	True Negative
TP	True Positive
UM	Unsharp Masking
XGBoost	Extreme Gradient Boosting

BIBLIOGRAPHIE

- [1] Tao Han, Virginia Xavier Nunes, Luís Fabrício De Freitas Souza, Adriell Gomes Marques, Iágson Carlos Lima Silva, Marcos Aurélio Araujo Ferreira Junior, Jinghua Sun, and Pedro P Reboucas Filho. Internet of medical things—based on deep learning techniques for segmentation of lung and stroke regions in ct scans. *IEEE Access*, 8 :71117–71135, 2020.
- [2] Rogelio Ruzcko NMI Tobias, Luigi Carlo M De Jesus, Matt Ervin G Mital, Sandy C Lauguico, Marielet A Guillermo, Edwin Sybingco, Argel A Bandala, and Elmer P Dadios. Cnn-based deep learning model for chest x-ray health classification using tensorflow. In *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*, pages 1–6. IEEE, 2020.
- [3] Qingji Guan and Yaping Huang. Multi-label chest x-ray image classification via category-wise residual attention learning. *Pattern Recognition Letters*, 130 :259–266, 2020.
- [4] Tomás Franquet. Imaging of community-acquired pneumonia. *Journal of thoracic imaging*, 33(5) :282–294, 2018.
- [5] Kh Tohidul Islam, Sudanthi NR Wijewickrema, Aaron Collins, and Stephen J O’Leary. A deep transfer learning framework for pneumonia detection from chest x-ray images. In *VISIGRAPP (5 : VISAPP)*, pages 286–293, 2020.
- [6] Subrato Bharati, Prajoy Podder, and M Rubaiyat Hossain Mondal. Hybrid deep learning for detecting lung diseases from x-ray images. *Informatics in Medicine Unlocked*, 20 :100391, 2020.
- [7] Yufei Jin, Huijuan Lu, Wenjie Zhu, and Wanli Huo. Deep learning based classification of multi-label chest x-ray images via dual-weighted metric loss. *Computers in Biology and Medicine*, 157 :106683, 2023.
- [8] Yan-Wei Lee, Sheng-Kai Huang, and Ruey-Feng Chang. Chexgat : A disease correlation-aware network for thorax disease diagnosis from chest x-ray images. *Artificial Intelligence in Medicine*, 132 :102382, 2022.

- [9] FM Javed Mehedi Shamrat, Sami Azam, Asif Karim, Kawsar Ahmed, Francis M Bui, and Friso De Boer. High-precision multiclass classification of lung disease through customized mobilenetv2 from chest x-ray images. *Computers in Biology and Medicine*, 155 :106646, 2023.
- [10] Domantas Kuzinkovas and Sandhya Clement. The detection of covid-19 in chest x-rays using ensemble cnn techniques. *Information*, 14(7) :370, 2023.
- [11] Vinayakumar Ravi, Vasundhara Acharya, and Mamoun Alazab. A multichannel efficientnet deep learning-based stacking ensemble approach for lung disease detection using chest x-ray images. *Cluster Computing*, 26(2) :1181–1203, 2023.
- [12] Shuxuan Xie, Zengchen Yu, and Zhihan Lv. Multi-disease prediction based on deep learning : a survey. *Computer Modeling in Engineering & Sciences*, 128(2) :489–522, 2021.
- [13] Adem Tekerek and Ismael Abdullah Mohammed Al-Rawe. A novel approach for prediction of lung disease using chest x-ray images based on densenet and mobilenet. *Wireless Personal Communications*, pages 1–15, 2023.
- [14] Lakshmi Sudha Kondaka, M Thenmozhi, K Vijayakumar, and Rashi Kohli. An intensive healthcare monitoring paradigm by using iot based machine learning strategies. *Multimedia Tools and Applications*, 81(26) :36891–36905, 2022.
- [15] Doctissimo. Les poumons : définition, anatomie et maladies — doctissimo.fr. <https://www.doctissimo.fr/sante/atlas-corps-humain/question-corps-humain/role-poumons>. [Consulté le 7 mars 2024].
- [16] Futura. Définition | Poumon | Futura Santé — futura-sciences.com. <https://www.futura-sciences.com/sante/definitions/biologie-poumon-7024/>. [Consulté le 7 mars 2024].
- [17] <https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancer-du-poumon/Les-poumons>. [Consulté le 7 mars 2024].
- [18] *World Health Organization. Pneumonia in children - 11 November 2022. (Consulté le 7 mars 2024).*
- [19] Infection pulmonaire - Infection des poumons | Ramsay Santé — ramsaysante.fr. <https://www.ramsaysante.fr/vous-etes-patient-en-savoir-plus-sur-ma-pathologie/infection-pulmonaire>. [Consulté le 7 mars 2024].

- [20] Venceslau Hespanhol and Cristina Bárbara. Pneumonia mortality, comorbidities matter? *Pulmonology*, 26(3) :123–129, 2020.
- [21] Pneumonie - symptômes, causes, traitements et prévention - VIDAL — vidal.fr. <https://www.vidal.fr/maladies/voies-respiratoires/pneumonie.html>. [Consulté le 7 mars 2024].
- [22] Épanchement pleural - Épanchement pleural - Manuels MSD pour le grand public — msdmanuals.com. <https://www.msdmanuals.com/fr/accueil/troubles-pulmonaires-et-des-voies-a%C3%A9riennes/maladies-de-la-pl%C3%A8vre-et-du-m%C3%A9diastin/%C3%A9panchement-pleural>. [Consulté le 8 mars 2024].
- [23] Épanchement pleural — chuv.ch. <https://www.chuv.ch/fr/chuv-home/patients-et-familles/specialites-medicales/atlas-medical-thematique/thorax/epanchement-pleural#:~:text=Dans%20la%20cage%20thoracique%2C%20les,de%20liquide%20entre%20les%20membranes>. [Consulté le 8 mars 2024].
- [24] Pleural effusion – A collection of fluid between the lung and chest wall that may compress the ... | Cuerpo humano anatomia, Imagenes de medicos, Anatomia del cuerpo — pinterest.fr. <https://www.pinterest.fr/pin/320811173431372141/>. [Consulté le 8 mars 2024].
- [25] Epanchement pleural — chuv.ch. <https://www.chuv.ch/fr/chirurgie-thoracique/cht-home/patients-et-famille/affections-du-thorax/autres-affections/epanchement-pleural>. [Consulté le 8 mars 2024].
- [26] J Pastré, S Roussel, O Sanchez, et al. Pleural effusion : diagnosis and management. *La Revue de Medecine Interne*, 36(4) :248–255, 2014.
- [27] Berthold Jany and Tobias Welte. Pleural effusion in adults—etiology, diagnosis, and treatment. *Deutsches Ärzteblatt International*, 116(21) :377, 2019.
- [28] Maribel Botana Rial, Javier Pérez Pallarés, Enrique Cases Viedma, Francisco Julián López González, José Manuel Porcel, María Rodríguez, Beatriz Romero Romero, Luis Valdés Cuadrado, Victoria Villena Garrido, and Rosa Cordovilla Pérez. Diagnosis and treatment of pleural effusion. recommendations of the spanish society of pulmonology and thoracic surgery. update 2022. *Archivos de bronconeumologia*, 59(1) :27–35, 2023.

- [29] Frank Gaillard. Consolidation | Radiology Reference Article | Radiopaedia.org — radiopaedia.org. <https://radiopaedia.org/articles/consolidation?lang=us>. [Consulté le 8 mars 2024].
- [30] The Healthline Editorial Team. Lung Consolidation : Treatment, vs Pleural Effusion, and More — healthline.com. <https://www.healthline.com/health/lung-consolidation>. [Consulté le 8 mars 2024].
- [31] Radiographie thoracique — ottawaheart.ca. <https://www.ottawaheart.ca/fr/examen-intervention/radiographie-thoracique>. [Consulté le 9 mars 2024].
- [32] Chest X-ray Quality — radiologymasterclass.co.uk. https://www.radiologymasterclass.co.uk/tutorials/chest/chest_quality/chest_xray_quality_projection#top_2nd_img. [Consulté le 9 mars 2024].
- [33] <https://www.ameli.fr/maine-et-loire/assure/sante/examen/imagerie-medicale/deroulement-scanner#:~:text=Qu'est%2Dce%20qu',%C3%A9tudier%20en%20effectuant%20des%20rotations>. [Consulté le 9 mars 2024].
- [34] Strainchamps Evrard, Mouchel. imagerie 3D en tomodesitom&xE9;trie — utc.fr. https://www.utc.fr/master-qualite/public/publications/qualite_et_biomedical/UTC/master_mts/2004-2005/projets/reconstruction3d/trois.html. [Consulté le 9 mars 2024].
- [35] Mohammad Osama Hussein Yonso. Right lower lobe pneumonia | Radiology Case | Radiopaedia.org — radiopaedia.org. <https://radiopaedia.org/cases/right-lower-lobe-pneumonia-7>. [Consulté le 10 mars 2024].
- [36] Chest X-ray Abnormalities — radiologymasterclass.co.uk. https://www.radiologymasterclass.co.uk/tutorials/chest/chest_pathology/chest_pathology_page4. [Consulté le 10 mars 2024].
- [37] Chest X-ray Abnormalities — radiologymasterclass.co.uk. https://www.radiologymasterclass.co.uk/tutorials/chest/chest_pathology/chest_pathology_page3#top_1st_img. [Consulté le 10 mars 2024].
- [38] Reabal Najjar. Right lower lobe consolidation | Radiology Case | Radiopaedia.org — radiopaedia.org. <https://radiopaedia.org/cases/right-lower-lobe-consolidation-1>. [Consulté le 10 mars 2024].
- [39] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. Chexnet :

- Radiologist-level pneumonia detection on chest x-rays with deep learning. (2017) *arXiv preprint arXiv :1711.05225*, 2017.
- [40] Pranav Rajpurkar et al. Deep learning for chest radiograph diagnosis : A retrospective comparison of the cheXnext algorithm to practicing radiologists. *PLoS medicine*, vol. 15(11) :e1002686, 2018.
- [41] Aya Hage Chehade, Nassib Abdallah, Jean-Marie Marion, Mathieu Hatt, Mohamad Oueidat, and Pierre Chauvet. A systematic review : Classification of lung diseases from chest x-ray images using deep learning algorithms. *SN Computer Science*, 5 :405, 2024.
- [42] Stefan Jaeger, Sema Candemir, Sameer Antani, Yi-Xiáng J. Wáng, Pu-Xuan Lu, and George Thoma. Two public chest x-ray dataset for computer-aided screening of pulmonary diseases. *Quantitative imaging in medicine and surgery*, 4(6) :475–477, 2014.
- [43] Paul Mooney. Chest X-Ray Images (Pneumonia). [Online] Available : <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>, 2019.
- [44] Junji Shiraishi, Shigehiko Katsuragawa, Junpei Ikezoe, Tsuneo Matsumoto, Takeshi Kobayashi, Ken-ichi Komatsu, Mitate Matsui, Hiroshi Fujita, Yoshie Kodera, and Kunio Doi. Development of a digital image database for chest radiographs with and without a lung nodule : receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *American Journal of Roentgenology*, 174(1) :71–74, 2000.
- [45] Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silvana Ciurea-Illcus, Chris Chute, Henrik Marklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, et al. Chexpert : A large chest radiograph dataset with uncertainty labels and expert comparison. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 590–597, 2019.
- [46] Alistair E. W. Johnson, Tom J. Pollard, Seth J. Berkowitz, Nathaniel R. Greenbaum, Matthew P. Lungren, Chih ying Deng, Roger G. Mark, and Steven Horng. MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Scientific Data*, 6, 2019.
- [47] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8 : Hospital-scale chest x-ray database and bench-

- marks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pages 2097–2106, 2017.
- [48] Olaide Nathaniel Oyelade, Absalom El-Shamir Ezugwu, and Haruna Chiroma. Covframenet : An enhanced deep learning framework for covid-19 detection. *Ieee Access*, 9 :77905–77919, 2021.
- [49] Chunli Qin, Demin Yao, Yonghong Shi, and Zhijian Song. Computer-aided detection in chest radiography based on artificial intelligence : a survey. *Biomedical engineering online*, 17 :1–23, 2018.
- [50] Xiaojun Xu, Youren Wang, Guoshi Yang, and Yanli Hu. Image enhancement method based on fractional wavelet transform. In *2016 IEEE International Conference on Signal and Image Processing (ICSIP)*, pages 194–197. IEEE, 2016.
- [51] SK Savitha and NC Naveen. Algorithm for pre-processing chest-x-ray using multi-level enhancement operation. In *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 2182–2186. IEEE, 2016.
- [52] Sivaramakrishnan Rajaraman, Sema Candemir, Zhiyun Xue, Philip O Alderson, M Kohli, Joseph Abuya, George R Thoma, and S Antani. A novel stacked generalization of models for improved tb detection in chest radiographs. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 718–721. IEEE, 2018.
- [53] Ilena Gabriella, Stella A Kamarga, and Agung W Setiawan. Early detection of tuberculosis using chest x-ray (cxr) with computer-aided diagnosis. In *2018 2nd international conference on biomedical engineering (IBIOMED)*, pages 76–79. IEEE, 2018.
- [54] Mohd Hanafi Ahmad Hijazi, Stefanus Kieu Tao Hwa, Abdullah Bade, Razali Yaa-kob, and Mohammad Saffree Jeffree. Ensemble deep learning for tuberculosis detection using chest x-ray and canny edge detected images. *IAES International Journal of Artificial Intelligence*, vol. 8(4) :429–435, 2019.
- [55] Magudeeswaran Veluchamy and Bharath Subramani. Image contrast and color enhancement using adaptive gamma correction and histogram equalization. *Optik*, 183 :329–337, 2019.

-
- [56] Khairul Munadi, Kahlil Muchtar, Novi Maulina, and Biswajeet Pradhan. Image enhancement for tuberculosis detection using deep learning. *in IEEE Access*, vol. 8, :217897–217907, 2020.
- [57] Sajid Ali Khan, Shariq Hussain, and Shunkun Yang. Contrast enhancement of low-contrast medical images using modified contrast limited adaptive histogram equalization. *Journal of Medical Imaging and Health Informatics*, 10(8) :1795–1803, 2020.
- [58] Tawsifur Rahman, Amith Khandakar, Yazan Qiblawey, Anas Tahir, Serkan Kiranyaz, Saad Bin Abul Kashem, Mohammad Tariqul Islam, Somaya Al Maadeed, Susu M Zughair, Muhammad Salman Khan, et al. Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images. *Computers in biology and medicine*, 132 :104319, 2021.
- [59] Khalid El Asnaoui and Youness Chawki. Using x-ray images and deep learning for automated detection of coronavirus disease. *Journal of Biomolecular Structure and Dynamics*, 39(10) :3615–3626, 2021.
- [60] Ahmed Iqbal, Muhammad Usman, and Zohair Ahmed. An efficient deep learning-based framework for tuberculosis detection using chest x-ray images. *Tuberculosis*, 136 :102234, 2022.
- [61] Abeer A AbdElhamid, Eman AbdElhalim, Mohamed A Mohamed, and Fahmi Khalifa. Multi-classification of chest x-rays for covid-19 diagnosis using deep learning algorithms. *Applied Sciences*, 12(4) :2080, 2022.
- [62] Alaa S Al-Waisy, Shumoos Al-Fahdawi, Mazin Abed Mohammed, Karrar Hameed Abdulkareem, Salama A Mostafa, Mashaal S Maashi, Muhammad Arif, and Begonya Garcia-Zapirain. Covid-chexnet : hybrid deep learning framework for identifying covid-19 virus in chest x-rays images. *Soft computing*, 27(5) :2657–2672, 2023.
- [63] Yupu Song, Cailin Li, Shiyang Xiao, Han Xiao, and Baoyun Guo. Unsharp masking image enhancement the parallel algorithm based on cross-platform. *Scientific Reports*, 12(1) :20175, 2022.
- [64] Aya Hage Chehade, Nassib Abdallah, Jean-Marie Marion, Mohamad Oueidat, and Pierre Chauvet. Lung and colon cancer classification using medical imaging : A feature engineering approach. *Physical and Engineering Sciences in Medicine*, 45(3) :729–746, 2022.

- [65] Santanu Roy, Kanika Bhalla, and Rachit Patel. Mathematical analysis of histogram equalization techniques for medical image enhancement : a tutorial from the perspective of data loss. *Multimedia Tools and Applications*, 83(5) :14363–14392, 2024.
- [66] Shubhi Kansal, Shikha Purwar, and Rajiv Kumar Tripathi. Image contrast enhancement using unsharp masking and histogram equalization. *Multimedia Tools and Applications*, 77 :26919–26938, 2018.
- [67] Tawsifur Rahman, Amith Khandakar, Yazan Qiblawey, Anas Tahir, Serkan Kiranyaz, Saad Bin Abul Kashem, Mohammad Tariqul Islam, Somaya Al Maadeed, Susu M Zughaier, Muhammad Salman Khan, et al. Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images. *Computers in biology and medicine*, 132 :104319, 2021.
- [68] Mohamed Y Adam, Mozamel M Saeed, and A Ahmed Al Samani. Medical image enhancement application using histogram equalization in computational libraries. *International Journal of Computer Science and Telecommunications*, 6(1) :7–12, 2015.
- [69] Pankaj Kandhway, Ashish Kumar Bhandari, and Anurag Singh. A novel reformed histogram equalization based medical image contrast enhancement using krill herd optimization. *Biomedical Signal Processing and Control*, 56 :101677, 2020.
- [70] Anas M Tahir, Yazan Qiblawey, Amith Khandakar, Tawsifur Rahman, Uzair Khurshid, Farayi Musharavati, Mohammad Tariqul Islam, Serkan Kiranyaz, Somaya Al-Maadeed, and Muhammad EH Chowdhury. Deep learning for reliable classification of covid-19, mers, and sars from chest x-ray images. *Cognitive computation*, pages 1–21, 2022.
- [71] Aleka Melese Ayalew, Ayodeji Olalekan Salau, Yibeltal Tamyalew, Bekalu Tadele Abeje, and Nigus Woreta. X-ray image-based covid-19 detection using deep learning. *Multimedia Tools and Applications*, 82(28) :44507–44525, 2023.
- [72] Giulio Siracusano, Aurelio La Corte, Michele Gaeta, Giuseppe Cicero, Massimo Chiappini, and Giovanni Finocchio. Pipeline for advanced contrast enhancement (pace) of chest x-ray in evaluating covid-19 patients by combining bidimensional empirical mode decomposition and contrast limited adaptive histogram equalization (clahe). *Sustainability*, 12(20) :8573, 2020.

-
- [73] Anggi Muhammad Rifai, Suwanto Raharjo, Ema Utami, and Dhani Ariatmanto. Analysis for diagnosis of pneumonia symptoms using chest x-ray based on mobilenetv2 models with image enhancement using white balance and contrast limited adaptive histogram equalization (clahe). *Biomedical Signal Processing and Control*, 90 :105857, 2024.
- [74] Agus Maman Abadi, Dhoriva Urwatul Wutsqa, and Leonardus Ragil Pamungkas. Detection of lung cancer using radiograph images enhancement and radial basis function classifier. In *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 1–6. IEEE, 2017.
- [75] Vinayak Sharma, Sachin Kumar Gupta, Kaushal Kumar Shukla, et al. Deep learning models for tuberculosis detection and infected region visualization in chest x-ray images. *Intelligent Medicine*, 2023.
- [76] Sema Candemir, Stefan Jaeger, Wilson Lin, Zhiyun Xue, Sameer Antani, and George Thoma. Automatic heart localization and radiographic index computation in chest x-rays. In *Medical Imaging 2016 : Computer-Aided Diagnosis*, volume 9785, pages 302–309. SPIE, 2016.
- [77] Changmiao Wang, Ahmed Elazab, Jianhuang Wu, and Qingmao Hu. Lung nodule classification using deep feature fusion in chest radiography. *Computerized Medical Imaging and Graphics*, 57 :10–18, 2017.
- [78] Alexey A Novikov, Dimitrios Lenis, David Major, Jiří Hladvka, Maria Wimmer, and Katja Bühler. Fully convolutional architectures for multiclass segmentation in chest radiographs. *IEEE transactions on medical imaging*, 37(8) :1865–1876, 2018.
- [79] W Dai, N Dong, Z Wang, X Liang, H Zhang, and EP Xing. Structure correcting adversarial network for organ segmentation in chest x-rays. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support ; Springer : Cham, Switzerland*, pages 263–273, 2018.
- [80] Han Liu, Lei Wang, Yandong Nan, Faguang Jin, Qi Wang, and Jiantao Pu. Sdfn : Segmentation-based deep fusion network for thoracic disease classification in chest x-ray images. *Computerized Medical Imaging and Graphics*, 75 :66–73, 2019.
- [81] Sheng Chen, Yaqi Han, Jinqiu Lin, Xiangyu Zhao, and Ping Kong. Pulmonary nodule detection on chest radiographs using balanced convolutional neural network

- and classic candidate detection. *Artificial Intelligence in Medicine*, 107 :101881, 2020.
- [82] Bingzhi Chen, Zheng Zhang, Jianyong Lin, Yi Chen, and Guangming Lu. Two-stream collaborative network for multi-label chest x-ray image classification with lung segmentation. *Pattern Recognition Letters*, 135 :221–227, 2020.
- [83] Julio Mendoza and Helio Pedrini. Detection and classification of lung nodules in chest x-ray images using deep convolutional neural networks. *Computational Intelligence*, 36(2) :370–401, 2020.
- [84] Barath Narayanan Narayanan and Russell C Hardie. A computationally efficient unet architecture for lung segmentation in chest radiographs. In *2019 IEEE National Aerospace and Electronics Conference (NAECON)*, pages 279–284. IEEE, 2019.
- [85] Barath Narayanan Narayanan, Venkata Salini Priyamvada Davuluru, and Russell C Hardie. Two-stage deep learning architecture for pneumonia detection and its diagnosis in chest radiographs. In *Medical Imaging 2020 : Imaging Informatics for Healthcare, Research, and Applications*, volume 11318, pages 130–139. SPIE, 2020.
- [86] Tawsifur Rahman, Amith Khandakar, Muhammad Abdul Kadir, Khandaker Rejaul Islam, Khandakar F Islam, Rashid Mazhar, Tahir Hamid, Mohammad Tariqul Islam, Saad Kashem, Zaid Bin Mahbub, et al. Reliable tuberculosis detection using chest x-ray with deep learning, segmentation and visualization. *IEEE Access*, 8 :191586–191601, 2020.
- [87] Yujia Xu, Hak-Keung Lam, and Guangyu Jia. Manet : A two-stage deep learning method for classification of covid-19 from chest x-ray images. *Neurocomputing*, 443 :96–105, 2021.
- [88] Ahmed Iqbal, Muhammad Usman, and Zohair Ahmed. Tuberculosis chest x-ray detection using cnn-based hybrid segmentation and classification approach. *Biomedical Signal Processing and Control*, 84 :104667, 2023.
- [89] Feidao Cao and Huaici Zhao. Automatic lung segmentation algorithm on chest x-ray images based on fusion variational auto-encoder and three-terminal attention mechanism. *Symmetry*, 13(5) :814, 2021.
- [90] Agughasi Victor Ikechukwu and S Murali. Cx-net : an efficient ensemble semantic deep neural network for roi identification from chest-x-ray images for copd diagnosis. *Machine Learning : Science and Technology*, 4(2) :025021, 2023.

-
- [91] S Arvind, Jitendra V Tembhurne, Tausif Diwan, and Parul Sahare. Improved light weight deep cnn based u-net for the semantic segmentation of lungs from chest x-rays. *Results in Engineering*, 17 :100929, 2023.
- [92] Feng Li, Roger Engelmann, Lorenzo Pesce, Samuel G Armato, and Heber MacMahon. Improved detection of focal pneumonia by chest radiography with bone suppression imaging. *European radiology*, 22(12) :2729–2735, 2012.
- [93] Hieu Xuan Nguyen and Tin Thanh Dang. Ribs suppression in chest x-ray images by using ica method. In *5th International Conference on Biomedical Engineering in Vietnam*, pages 194–197. Springer, 2015.
- [94] Wei Yang, Yingyin Chen, Yunbi Liu, Liming Zhong, Genggeng Qin, Zhentai Lu, Qianjin Feng, and Wufan Chen. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. *Medical image analysis*, 35 :421–433, 2017.
- [95] Yu Gordienko, Peng Gang, Jiang Hui, Wei Zeng, Yu Kochura, Oleg Alienin, Oleksandr Rokovyi, and Sergii Stirenko. Deep learning with lung segmentation and bone shadow exclusion techniques for chest x-ray analysis of lung cancer. In *Advances in Computer Science for Engineering and Education, International conference on computer science, engineering and education applications (ICCSEEA)*, pages 638–647. Springer, 2018.
- [96] Sivaramakrishnan Rajaraman, Ghada Zamzmi, Les Folio, Philip Alderson, and Sameer Antani. Chest x-ray bone suppression for improving classification of tuberculosis-consistent findings. *Diagnostics*, 11(5) :840, 2021.
- [97] Luyi Han, Yuanyuan Lyu, Cheng Peng, and S Kevin Zhou. Gan-based disentanglement learning for chest x-ray rib suppression. *Medical Image Analysis*, 77 :102369, 2022.
- [98] Di Xu, Qifan Xu, Kevin Nhieu, Dan Ruan, and Ke Sheng. An efficient and robust method for chest x-ray rib suppression that improves pulmonary abnormality diagnosis. *Diagnostics*, 13(9) :1652, 2023.
- [99] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [100] Jürgen Schmidhuber. Deep learning in neural networks : An overview. *Neural networks*, 61 :85–117, 2015.

- [101] Jun Gao, Qian Jiang, Bo Zhou, and Daozheng Chen. Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis : An overview. *Mathematical Biosciences and Engineering*, 16(6) :6536–6561, 2019.
- [102] Shi Dong, Ping Wang, and Khushnood Abbas. A survey on deep learning and its applications. *Computer Science Review*, 40 :100379, 2021.
- [103] Ana Barragán-Montero, Umair Javaid, Gilmer Valdés, Dan Nguyen, Paul Desbordes, Benoit Macq, Siri Willems, Liesbeth Vandewinckele, Mats Holmström, Fredrik Löfman, et al. Artificial intelligence and machine learning for medical imaging : A technology review. *Physica Medica*, 83 :242–256, 2021.
- [104] Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do, and Kaori Togashi. Convolutional neural networks : an overview and application in radiology. *Insights into imaging*, 9 :611–629, 2018.
- [105] Lingga Aji Andika, Hasih Pratiwi, and Sri Sulistijowati Handajani. Convolutional neural network modeling for classification of pulmonary tuberculosis disease. In *Journal of Physics : Conference Series*, volume 1490, page 012020. IOP Publishing, 2020.
- [106] Rukshan Pramoditha. Convolutional Neural Network (CNN) Architecture Explained in Plain English Using Simple Diagrams — towardsdatascience.com. [Consulté le 20 mars 2024].
- [107] Rehenuma Lazin, Xinyi Shen, and Emmanouil Anagnostou. Estimation of flood-damaged cropland area using a convolutional neural network. *Environmental Research Letters*, 16(5) :054011, 2021.
- [108] Huo Yingge, Imran Ali, and Kang-Yoon Lee. Deep neural networks on chip-a survey. In *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pages 589–592. IEEE, 2020.
- [109] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net : Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015 : 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.
- [110] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

-
- [111] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [112] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words : Transformers for image recognition at scale. *arXiv preprint arXiv :2010.11929*, 2020.
- [113] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9650–9660, 2021.
- [114] Sukhendra Singh, Manoj Kumar, Abhay Kumar, Birendra Kumar Verma, Kumar Abhishek, and Shitharth Selvarajan. Efficient pneumonia detection using vision transformers on chest x-rays. *Scientific Reports*, 14(1) :2487, 2024.
- [115] Tete Xiao, Mannat Singh, Eric Mintun, Trevor Darrell, Piotr Dollár, and Ross Girshick. Early convolutions help transformers see better. *Advances in neural information processing systems*, 34 :30392–30400, 2021.
- [116] Haiping Wu, Bin Xiao, Noel Codella, Mengchen Liu, Xiyang Dai, Lu Yuan, and Lei Zhang. Cvt : Introducing convolutions to vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 22–31, 2021.
- [117] Samira Sajed, Amir Sanati, Jorge Esparteiro Garcia, Habib Rostami, Ahmad Keshavarz, and Andreia Teixeira. The effectiveness of deep learning vs. traditional methods for lung disease diagnosis using chest x-ray images : A systematic review. *Applied Soft Computing*, page 110817, 2023.
- [118] Sangjoon Park, Gwanghyun Kim, Yujin Oh, Joon Beom Seo, Sang Min Lee, Jin Hwan Kim, Sungjun Moon, Jae-Kwang Lim, Chang Min Park, and Jong Chul Ye. Self-evolving vision transformer for chest x-ray diagnosis through knowledge distillation. *Nature communications*, 13(1) :3848, 2022.
- [119] Michail Mamalakis, Andrew J Swift, Bart Vorselaars, Surajit Ray, Simonne Weeks, Weiping Ding, Richard H Clayton, Louise S Mackenzie, and Abhirup Banerjee. Denrescov-19 : A deep transfer learning network for robust automatic classification of covid-19, pneumonia, and tuberculosis from x-rays. *Computerized Medical Imaging and Graphics*, 94 :102008, 2021.

- [120] Rakesh Chandra Joshi, Saumya Yadav, Vinay Kumar Pathak, Hardeep Singh Malhotra, Harsh Vardhan Singh Khokhar, Anit Parihar, Neera Kohli, D Himanshu, Ravindra K Garg, Madan Lal Brahma Bhatt, et al. A deep learning-based covid-19 automatic diagnostic framework using chest x-ray images. *Biocybernetics and Biomedical Engineering*, 41(1) :239–254, 2021.
- [121] Mohammad Rahimzadeh and Abolfazl Attar. A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2. *Informatics in medicine unlocked*, 19 :100360, 2020.
- [122] Somnath Rakshit, Indrajit Saha, Michal Wlasnowolski, Ujjwal Maulik, and Dariusz Plewczynski. Deep learning for detection and localization of thoracic diseases using chest x-ray imagery. In *Artificial Intelligence and Soft Computing : 18th International Conference, ICAISC 2019, Zakopane, Poland, June 16–20, 2019, Proceedings, Part II 18*, pages 271–282. Springer, 2019.
- [123] Sebastian Guendel, Sasa Grbic, Bogdan Georgescu, Siqui Liu, Andreas Maier, and Dorin Comaniciu. Learning to recognize abnormalities in chest x-rays with location-aware dense networks. In *Iberoamerican Congress on Pattern Recognition*, pages 757–765. Springer, 2018.
- [124] Jing Xu, Hui Li, and Xiu Li. Ms-anet : deep learning for automated multi-label thoracic disease detection and classification. *PeerJ Computer Science*, 7 :e541, 2021.
- [125] Arjun Chaudhary, Abhishek Hazra, and Prakash Chaudhary. Diagnosis of chest diseases in x-ray images using deep convolutional neural network. In *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pages 1–6. IEEE, 2019.
- [126] Ivo M Baltruschat, Hannes Nickisch, Michael Grass, Tobias Knopp, and Axel Saalbach. Comparison of deep learning approaches for multi-label chest x-ray classification. *Scientific reports*, 9(1) :1–10, 2019.
- [127] Harsh Sharma, Jai Sethia Jain, Priti Bansal, and Sumit Gupta. Feature extraction and classification of chest x-ray images using cnn to detect pneumonia. In *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pages 227–231. IEEE, 2020.
- [128] Nada M Elshennawy and Dina M Ibrahim. Deep-pneumonia framework using deep learning models based on chest x-ray images. *Diagnostics*, 10(9) :649, 2020.

-
- [129] N Kumar, M Gupta, D Gupta, and S Tiwari. Novel deep transfer learning model for covid-19 patient detection using x-ray chest images. *Journal of ambient intelligence and humanized computing*, 14(1) :469–478, 2023.
- [130] Luca Brunese, Francesco Mercaldo, Alfonso Reginelli, and Antonella Santone. Explainable deep learning for pulmonary disease and coronavirus covid-19 detection from x-rays. *Computer Methods and Programs in Biomedicine*, 196 :105608, 2020.
- [131] Ilyas Sirazitdinov, Maksym Kholiavchenko, Tamerlan Mustafaev, Yuan Yixuan, Ramil Kuleev, and Bulat Ibragimov. Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. *Computers & electrical engineering*, 78 :388–399, 2019.
- [132] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.
- [133] Hojjat Salehinejad, Shahrokh Valaee, Tim Dowdell, Errol Colak, and Joseph Barfett. Generalization of deep neural networks for chest pathology classification in x-rays using generative adversarial networks. In *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 990–994. IEEE, 2018.
- [134] Ali Madani, Mehdi Moradi, Alexandros Karargyris, and Tanveer Syeda-Mahmood. Chest x-ray generation and data augmentation for cardiovascular abnormality classification. In *Medical imaging 2018 : Image processing*, volume 10574, pages 415–420. SPIE, 2018.
- [135] Sagar Kora Venu and Sridhar Ravula. Evaluation of deep convolutional generative adversarial networks for data augmentation of chest x-ray images. *Future Internet*, 13(1) :8, 2020.
- [136] Saleh Albahli and Ghulam Nabi Ahmad Hassan Yar. Ai-driven deep convolutional neural networks for chest x-ray pathology identification. *Journal of X-Ray Science and Technology*, 30(2) :365–376, 2022.
- [137] Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big data*, 3 :1–40, 2016.
- [138] Mohammad Tariqul Islam, Md Abdul Aowal, Ahmed Tahseen Minhaz, and Khalid Ashraf. Abnormality detection and localization in chest x-rays using deep convolutional neural networks. *arXiv preprint arXiv :1705.09850*, 2017.

- [139] Rahul Hooda, Ajay Mittal, and Sanjeev Sofat. Automated tb classification using ensemble of deep architectures. *Multimedia Tools and Applications*, 78(22) :31515–31532, 2019.
- [140] Thi Kieu Khanh Ho and Jeonghwan Gwak. Multiple feature integration for classification of thoracic disease in chest radiography. *Applied Sciences*, 9(19) :4130, 2019.
- [141] Asmaa Abbas, Mohammed M Abdelsamea, and Mohamed Medhat Gaber. Detrac : Transfer learning of class decomposed medical images in convolutional neural networks. *IEEE Access*, 8 :74901–74913, 2020.
- [142] Hieu H Pham, Tung T Le, Dat Q Tran, Dat T Ngo, and Ha Q Nguyen. Interpreting chest x-rays via cnns that exploit hierarchical disease dependencies and uncertainty labels. *Neurocomputing*, 437 :186–194, 2021.
- [143] Sinno Jialin Pan. Transfer learning. *Learning*, 21 :1–2, 2020.
- [144] Cheng Wang, Delei Chen, Lin Hao, Xuebo Liu, Yu Zeng, Jianwei Chen, and Guokai Zhang. Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access*, 7 :146533–146541, 2019.
- [145] Abdullah Irfan, Akash L Adivishnu, Antonio Sze-To, Taher Dehkharghanian, Shahrar Rahnamayan, and Hamid R Tizhoosh. Classifying pneumonia among chest x-rays using transfer learning. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 2186–2189. IEEE, 2020.
- [146] Rudrajit Choudhuri and Amit Paul. Multi class image classification for detection of diseases using chest x ray images. In *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 769–773. IEEE, 2021.
- [147] Anuj Rohilla, Rahul Hooda, and Ajay Mittal. Tb detection in chest radiograph using deep learning architecture. *ICETETSM-17*, pages 136–147, 2017.
- [148] Tawsifur Rahman, Muhammad EH Chowdhury, Amith Khandakar, Khandaker R Islam, Khandaker F Islam, Zaid B Mahbub, Muhammad A Kadir, and Saad Kashem. Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray. *Applied Sciences*, 10(9) :3233, 2020.
- [149] Vo Trong Quang Huy and Chih-Min Lin. An improved densenet deep neural network model for tuberculosis detection using chest x-ray images. *IEEE Access*, 2023.

- [150] Sangheum Hwang, Hyo-Eun Kim, Jihoon Jeong, and Hee-Jin Kim. A novel approach for tuberculosis screening based on deep convolutional neural networks. In *Medical imaging 2016 : computer-aided diagnosis*, volume 9785, pages 750–757. SPIE, 2016.
- [151] Paras Lakhani and Baskaran Sundaram. Deep learning at chest radiography : automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, 284(2) :574–582, 2017.
- [152] Rabia Emhamed Al Mamlook, Shengfeng Chen, and Hanin Fawzi Bzizi. Investigation of the performance of machine learning classifiers for pneumonia detection in chest x-ray images. In *2020 IEEE International Conference on Electro Information Technology (EIT)*, pages 098–104. IEEE, 2020.
- [153] Anuja Kumar Acharya and Rajalakshmi Satapathy. A deep learning based approach towards the automatic diagnosis of pneumonia from chest radio-graphs. *Biomedical and Pharmacology Journal*, 13(1) :449–455, 2020.
- [154] Yuting Yang, Gang Mei, and Francesco Piccialli. A deep learning approach considering image background for pneumonia identification using explainable ai (xai). *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2022.
- [155] Shagun Sharma and Kalpna Guleria. A deep learning based model for the detection of pneumonia from chest x-ray images using vgg-16 and neural networks. *Procedia Computer Science*, 218 :357–366, 2023.
- [156] Isabel Bush. Lung nodule detection and classification. *Report, Stanford Computer Science*, 2016.
- [157] Worawate Ausawalaithong, Arjaree Thirach, Sanparith Marukatat, and Theerawit Wilaiprasitporn. Automatic lung cancer prediction from chest x-ray images using the deep learning approach. In *2018 11th Biomedical Engineering International Conference (BMEiCON)*, pages 1–5. IEEE, 2018.
- [158] R Takemiya, S Kido, Y Hirano, and S Mabu. Detection of pulmonary nodules on chest x-ray images using r-cnn. In *International Forum on Medical Imaging in Asia 2019*, volume 11050, pages 147–152. SPIE, 2019.
- [159] Ioannis D Apostolopoulos and Tzani A Mpesiana. Covid-19 : automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and engineering sciences in medicine*, 43(2) :635–640, 2020.

- [160] Fatima M Salman, Samy S Abu-Naser, Eman Alajrami, Bassem S Abu-Nasser, and Belal AM Alashqar. Covid-19 detection using artificial intelligence. *International Journal of Academic Engineering Research (IJAER)*, 4 :18–25, 2020.
- [161] Kabid Hassan Shibly, Samrat Kumar Dey, Md Tahzib-Ul Islam, and Md Mahbubur Rahman. Covid faster r-cnn : A novel framework to diagnose novel coronavirus disease (covid-19) in x-ray images. *Informatics in Medicine Unlocked*, 20 :100405, 2020.
- [162] Rachna Sethi, Monica Mehrotra, and Dhaarna Sethi. Deep learning based diagnosis recommendation for covid-19 using chest x-rays images. In *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pages 1–4. IEEE, 2020.
- [163] Harsh Panwar, PK Gupta, Mohammad Khubeb Siddiqui, Ruben Morales-Menendez, and Vaishnavi Singh. Application of deep learning for fast detection of covid-19 in x-rays using ncovnet. *Chaos, Solitons & Fractals*, 138 :109944, 2020.
- [164] Dipayan Das, KC Santosh, and Umapada Pal. Truncated inception net : Covid-19 outbreak screening using chest x-rays. *Physical and engineering sciences in medicine*, 43(3) :915–925, 2020.
- [165] Parnian Afshar, Shahin Heidarian, Farnoosh Naderkhani, Anastasia Oikonomou, Konstantinos N Plataniotis, and Arash Mohammadi. Covid-caps : A capsule network-based framework for identification of covid-19 cases from x-ray images. *Pattern Recognition Letters*, 138 :638–643, 2020.
- [166] Emtiaz Hussain, Mahmudul Hasan, Md Anisur Rahman, Ickjai Lee, Tasmi Tamanna, and Mohammad Zavid Parvez. Corodet : A deep learning based classification for covid-19 detection using chest x-ray images. *Chaos, Solitons & Fractals*, 142 :110495, 2021.
- [167] M Emin Sahin. Deep learning-based approach for detecting covid-19 in chest x-rays. *Biomedical Signal Processing and Control*, 78 :103977, 2022.
- [168] Sammy V Militante, Nanette V Dionisio, and Brandon G Sibbaluca. Pneumonia and covid-19 detection using convolutional neural networks. In *2020 Third International Conference on Vocational Education and Electrical Engineering (ICVEE)*, pages 1–6. IEEE, 2020.
- [169] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U Rajendra Acharya. Automated detection of covid-19 cases using deep

- neural networks with x-ray images. *Computers in biology and medicine*, 121 :103792, 2020.
- [170] Dina M Ibrahim, Nada M Elshennawy, and Amany M Sarhan. Deep-chest : Multi-classification deep learning model for diagnosing covid-19, pneumonia, and lung cancer chest diseases. *Computers in biology and medicine*, 132 :104348, 2021.
- [171] Ejaz Khan, Muhammad Zia Ur Rehman, Fawad Ahmed, Faisal Abdulaziz Alfouzan, Nouf M Alzahrani, and Jawad Ahmad. Chest x-ray classification for the detection of covid-19 using deep learning techniques. *Sensors*, 22(3) :1211, 2022.
- [172] Goram Mufarah M Alshmrani, Qiang Ni, Richard Jiang, Haris Pervaiz, and Nada M Elshennawy. A deep learning architecture for multi-class lung diseases classification using chest x-ray (cxr) images. *Alexandria Engineering Journal*, 64 :923–935, 2023.
- [173] Qingji Guan and Yaping Huang. Multi-label chest x-ray image classification via category-wise residual attention learning. *Pattern Recognition Letters*, 130 :259–266, 2020.
- [174] Jonathan Rubin, Deepan Sanghavi, Claire Zhao, Kathy Lee, Ashequl Qadir, and Minnan Xu-Wilson. Large scale automated reading of frontal and lateral chest x-rays using dual convolutional neural networks. *arXiv preprint arXiv :1804.07839*, 2018.
- [175] Imane Allaouzi and Mohamed Ben Ahmed. A novel approach for multi-label chest x-ray classification of common thorax diseases. *IEEE Access*, vol. 7 :64279–64288, 2019.
- [176] Pulkit Kumar, Monika Grewal, and Muktabh Mayank Srivastava. Boosted cascaded convnets for multilabel classification of thoracic diseases in chest radiographs. In *International conference image analysis and recognition, ICIAR 2018, Lecture Notes in Computer Science*, pages 546–552. Springer, 2018.
- [177] Hongyu Wang and Yong Xia. Chestnet : A deep neural network for classification of thoracic diseases on chest radiography. (2018) *arXiv preprint arXiv :1807.03058*, 2018.
- [178] Yuxing Tang, Xiaosong Wang, Adam P Harrison, Le Lu, Jing Xiao, and Ronald M Summers. Attention-guided curriculum learning for weakly supervised classification and localization of thoracic diseases on chest radiographs. In *International Workshop on Machine Learning in Medical Imaging*, pages 249–258. Springer, 2018.

- [179] Zongyuan Ge, Dwarikanath Mahapatra, Suman Sedai, Rahil Garnavi, and Rajib Chakravorty. Chest x-rays classification : A multi-label and fine-grained problem. *arXiv preprint arXiv :1807.07247*, 2018.
- [180] Zhe Li, Chong Wang, Mei Han, Yuan Xue, Wei Wei, Li-Jia Li, and Li Fei-Fei. Thoracic disease identification and localization with limited supervision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8290–8299, 2018.
- [181] Shaocong Mo and Ming Cai. Deep learning based multi-label chest x-ray classification with entropy weighting loss. In *2019 12th International Symposium on Computational Intelligence and Design (ISCID)*, volume 2, pages 124–127. IEEE, 2019.
- [182] Bingzhi Chen, Jinxing Li, Guangming Lu, and David Zhang. Lesion location attention guided network for multi-label thoracic disease classification in chest x-rays. *IEEE journal of biomedical and health informatics*, 24(7) :2016–2027, 2019.
- [183] Yanbo Ma, Qiu hao Zhou, Xuesong Chen, Haihua Lu, and Yong Zhao. Multi-attention network for thoracic disease classification and localization. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1378–1382. IEEE, 2019.
- [184] Congbo Ma, Hu Wang, and Steven CH Hoi. Multi-label thoracic disease image classification with cross-attention networks. In *International conference on medical image computing and computer-assisted intervention*, pages 730–738. Springer, 2019.
- [185] Zongyuan Ge, Dwarikanath Mahapatra, Xiaojun Chang, Zetao Chen, Lianhua Chi, and Huimin Lu. Improving multi-label chest x-ray disease diagnosis by exploiting disease and health labels dependencies. *Multimedia Tools and Applications*, 79(21) :14889–14902, 2020.
- [186] Abdelbaki Souid, Nizar Sakli, and Hedi Sakli. Classification and predictions of lung diseases from chest x-rays using mobilenet v2. *Applied Sciences*, 11(6) :2751, 2021.
- [187] Jiashi Zhao, Mengmeng Li, Weili Shi, Yu Miao, Zhengang Jiang, and Bai Ji. A deep learning method for classification of chest x-ray images. In *Journal of Physics : Conference Series*, volume 1848, page 012030. IOP Publishing, 2021.
- [188] Lifang Chen, Tengfei Mao, and Qian Zhang. Identifying cardiomegaly in chest x-rays using dual attention network. *Applied Intelligence*, 52(10) :11058–11067, 2022.

-
- [189] Mukesh Mann, Rakesh P Badoni, Harsh Soni, Mohammed Al-Shehri, Aman Chandra Kaushik, and Dong-Qing Wei. Utilization of deep convolutional neural networks for accurate chest x-ray diagnosis and disease detection. *Interdisciplinary Sciences : Computational Life Sciences*, pages 1–19, 2023.
- [190] Bingzhi Chen, Jinxing Li, Xiaobao Guo, and Guangming Lu. Dualhexnet : dual asymmetric feature learning for thoracic disease classification in chest x-rays. *Biomedical Signal Processing and Control*, 53 :101554, 2019.
- [191] Vinicius Teixeira, Leodécio Braz, Helio Pedrini, and Zanoni Dias. Dualanet : dual lesion attention network for thoracic disease classification in chest x-rays. In *2020 international conference on systems, signals and image processing (IWSSIP)*, pages 69–74. IEEE, 2020.
- [192] Saleh Albahli, Hafiz Tayyab Rauf, Abdulelah Algosaihi, and Valentina Emilia Balas. Ai-driven deep cnn approach for multi-label pathology classification using chest x-rays. *PeerJ Computer Science*, 7 :e495, 2021.
- [193] Mingyan Yang, Hisashi Tanaka, and Takayuki Ishida. Performance improvement in multi-label thoracic abnormality classification of chest x-rays with noisy labels. *International Journal of Computer Assisted Radiology and Surgery*, pages 1–9, 2022.
- [194] Li Yao, Eric Poblentz, Dmitry Dagunts, Ben Covington, Devon Bernard, and Kevin Lyman. Learning to diagnose from scratch by exploiting dependencies among labels. *arXiv preprint arXiv :1710.10501*, 2017.
- [195] Chaochao Yan, Jiawen Yao, Ruoyu Li, Zheng Xu, and Junzhou Huang. Weakly supervised deep learning for thoracic disease classification and localization on chest x-rays. In *Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics*, pages 103–110, 2018.
- [196] Yan Shen and Mingchen Gao. Dynamic routing on deep neural network for thoracic disease classification and sensitive area localization. In *International Workshop on Machine Learning in Medical Imaging*, pages 389–397. Springer, 2018.
- [197] Xiaoxuan Liu, Livia Faes, Aditya U Kale, Siegfried K Wagner, Dun Jack Fu, Alice Bruynseels, Thushika Mahendiran, Gabriella Moraes, Mohith Shamdas, Christoph Kern, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging : a systematic review and meta-analysis. *The lancet digital health*, 1(6) :e271–e297, 2019.

- [198] Suhail Raoof, David Feigin, Arthur Sung, Sabiha Raoof, Lavanya Irugulpati, and Edward C Rosenow III. Interpretation of plain chest roentgenogram. *Chest*, 141(2) :545–558, 2012.
- [199] René Hosch, Lennard Kroll, Felix Nensa, and Sven Koitka. Differentiation between anteroposterior and posteroanterior chest x-ray view position with convolutional neural networks. In *RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, volume 193, pages 168–176. Georg Thieme Verlag KG, 2021.
- [200] Seear Al-Dabooni and Donald Wunsch. Model order reduction based on agglomerative hierarchical clustering. *IEEE transactions on neural networks and learning systems*, 30(6) :1881–1895, 2018.
- [201] Alexandros Margaritis, Hilde Soenen, Erik Fransen, Georgios Pipintakos, Geert Jacobs, Johan Blom, et al. Identification of ageing state clusters of reclaimed asphalt binders using principal component analysis (pca) and hierarchical cluster analysis (hca) based on chemo-rheological parameters. *Construction and Building Materials*, 244 :118276, 2020.
- [202] Fatih Demir, Abdulkadir Sengur, and Varun Bajaj. Convolutional neural networks based efficient approach for classification of lung diseases. *Health information science and systems*, 8(1) :4, 2020.
- [203] FM Javed Mehedi Shamrat, Sami Azam, Asif Karim, Rakibul Islam, Zarrin Tasnim, Pronab Ghosh, and Friso De Boer. Lungnet22 : a fine-tuned model for multiclass classification and prediction of lung disease using x-ray images. *Journal of Personalized Medicine*, 12(5) :680, 2022.
- [204] Diederik P Kingma and Jimmy Ba. Adam : A method for stochastic optimization. *arXiv preprint arXiv :1412.6980*, 2014.
- [205] Jing Li, Kan Jin, Dalin Zhou, Naoyuki Kubota, and Zhaojie Ju. Attention mechanism-based cnn for facial expression recognition. *Neurocomputing*, 411 :340–350, 2020.
- [206] Manpreet Kaur, Jasdeep Kaur, and Jappreet Kaur. Survey of contrast enhancement techniques based on histogram equalization. *International Journal of Advanced Computer Science and Applications*, 2(7), 2011.
- [207] Malathy Jawahar, J Prassanna, Vinayakumar Ravi, L Jani Anbarasi, S Graceline Jasmine, R Manikandan, Ramesh Sekaran, and Suthendran Kannan. Computer-

- aided diagnosis of covid-19 from chest x-ray images using histogram-oriented gradient features and random forest classifier. *Multimedia Tools and Applications*, 81(28) :40451–40468, 2022.
- [208] Bin Li, Kaili Cheng, and Zhezhou Yu. Histogram of oriented gradient based gist feature for building recognition. *Computational intelligence and neuroscience*, 2016, 2016.
- [209] Ayodeji Olalekan Salau and Shruti Jain. Feature extraction : a survey of the types, techniques, applications. In *2019 international conference on signal processing and communication (ICSC)*, pages 158–164. IEEE, 2019.
- [210] Aleka Melese Ayalew, Ayodeji Olalekan Salau, Bekalu Tadele Abeje, and Belay Enyew. Detection and classification of covid-19 disease from x-ray images using convolutional neural networks and histogram of oriented gradients. *Biomedical Signal Processing and Control*, 74 :103530, 2022.
- [211] Muhammad Darwis, Liyando Hermawan Hasibuan, Mochammad Firmansyah, Nur Ahady, and Rizka Tiaharyadini. Implementation of k-means clustering algorithm in mapping the groups of graduated or dropped-out students in the management department of the national university. *JISA (Jurnal Informatika dan Sains)*, 4(1) :1–9, 2021.
- [212] Fayza Nayla Riyana Putri, Nur Cahyo Hendro Wibowo, and Hery Mustofa. Clustering of tuberculosis and normal lungs based on image segmentation results of chan-vese and canny with k-means. *Indonesian Journal of Artificial Intelligence and Data Mining*, 6(1) :18–28, 2023.
- [213] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.
- [214] Daniel Iglesias Morís, Jose Joaquim de Moura Ramos, Jorge Novo Buján, and Marcos Ortega Hortas. Data augmentation approaches using cycle-consistent adversarial networks for improving covid-19 screening in portable chest x-ray images. *Expert systems with applications*, 185 :115681, 2021.
- [215] Asaf Bar-El, Dana Cohen, Noa Cahan, and Hayit Greenspan. Improved cyclegan with application to covid-19 classification. In *Medical Imaging 2021 : Image Processing*, volume 11596, pages 296–305. SPIE, 2021.

- [216] Sukhendra Singh, Manoj Kumar, Abhay Kumar, Birendra Kumar Verma, and S Shitharth. Pneumonia detection with qcsa network on chest x-ray. *Scientific Reports*, 13(1) :9025, 2023.
- [217] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam : Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19, 2018.
- [218] Yan Han, Chongyan Chen, Liyan Tang, Mingquan Lin, Ajay Jaiswal, Song Wang, Ahmed Tewfik, George Shih, Ying Ding, and Yifan Peng. Using radiomics as prior knowledge for thorax disease classification and localization in chest x-rays. In *AMIA Annual Symposium Proceedings*, volume 2021, page 546. American Medical Informatics Association, 2021.
- [219] Michele Avanzo, Joseph Stancanello, and Issam El Naqa. Beyond imaging : the promise of radiomics. *Physica Medica*, 38 :122–139, 2017.
- [220] Michele Avanzo, Lise Wei, Joseph Stancanello, Martin Vallieres, Arvind Rao, Olivier Morin, Sarah A Mattonen, and Issam El Naqa. Machine and deep learning methods for radiomics. *Medical physics*, 47(5) :e185–e202, 2020.
- [221] Nikhil Pandey. Chest Xray Masks and Labels — kaggle.com. <https://www.kaggle.com/datasets/nikhilpandey360/chest-xray-masks-and-labels>. [Consulté le 2 mai 2024].
- [222] Nobuyuki Ostu. A threshold selection method from gray-level histograms. *IEEE Trans SMC*, 9 :62, 1979.
- [223] Hsin-Jui Chen, Shanq-Jang Ruan, Sha-Wo Huang, and Yan-Tsung Peng. Lung x-ray segmentation using deep convolutional neural networks on contrast-enhanced binarized images. *Mathematics*, 8(4) :545, 2020.
- [224] Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud GPM Van Stiphout, Patrick Granton, Catharina ML Zegers, Robert Gillies, Ronald Boellard, André Dekker, et al. Radiomics : extracting more information from medical images using advanced feature analysis. *European journal of cancer*, 48(4) :441–446, 2012.
- [225] Tamarisk du Plessis, Gopika Ramkilawon, William Ian Duncombe Rae, Tanita Botha, Neil Alexander Martinson, Sarah Alice Parry Dixon, Andre Kyme, and Mike Michael Sathekge. Introducing a secondary segmentation to construct a radio-

- mics model for pulmonary tuberculosis cavities. *La radiologia medica*, 128(9) :1093–1102, 2023.
- [226] Joost JM Van Griethuysen, Andriy Fedorov, Chintan Parmar, Ahmed Hosny, Nicole Aucoin, Vivek Narayan, Regina GH Beets-Tan, Jean-Christophe Fillion-Robin, Steve Pieper, and Hugo JWL Aerts. Computational radiomics system to decode the radiographic phenotype. *Cancer research*, 77(21) :e104–e107, 2017.
- [227] Iqbal H Sarker. Machine learning : Algorithms, real-world applications and research directions. *SN computer science*, 2(3) :160, 2021.
- [228] Pushpalatha C Bhat, Harrison B Prosper, Sezen Sekmen, and Chip Stewart. Optimizing event selection with the random grid search. *Computer Physics Communications*, 228 :245–257, 2018.
- [229] Christoph Molnar. *Interpretable machine learning*. Lulu. com, 2020.
- [230] Sebastian Gündel, Arnaud AA Setio, Florin C Ghesu, Sasa Grbic, Bogdan Georgescu, Andreas Maier, and Dorin Comaniciu. Robust classification from noisy labels : Integrating additional knowledge for chest radiography abnormality assessment. *Medical Image Analysis*, 72 :102087, 2021.

TABLE DES FIGURES

1.1	Anatomie des poumons humains.	18
1.2	Illustration de la pneumonie.	19
1.3	Illustration de l'épanchement pleural.	21
1.4	Procédure d'acquisition des radiographies thoraciques PA et AP.	26
1.5	Procédure d'acquisition d'une image scanner.	27
1.6	Visualisation de la pneumonie sur la radiographie thoracique.	28
1.7	Visualisation de l'épanchement pleural sur la radiographie thoracique.	29
1.8	Visualisation de la consolidation sur la radiographie thoracique.	29
2.1	Organigramme de l'application de l'apprentissage profond pour la classification des pathologies pulmonaires à partir des images de radiographie thoracique.	32
2.2	CXR originale et son amélioration en utilisant la méthode Unsharp Masking.	37
2.3	CXR originale et son amélioration en utilisant la méthode Histogram Equalization.	38
2.4	CXR originale et son amélioration en utilisant la méthode Contrast Limited Adaptive Histogram Equalization.	39
2.5	Le processus de la méthode High-Frequency Emphasis Filtering.	40
2.6	CXR originale et son amélioration en utilisant la méthode High-Frequency Emphasis Filtering.	41
2.7	Résultats de la segmentation.	42
2.8	Résultats de la suppression des structures osseuses.	43
2.9	Architecture d'un réseau de neurones convolutifs.	45
2.10	Principe de l'opérateur de convolution.	47
2.11	L'impact du padding sur la dimension de la carte de caractéristiques.	48
2.12	L'effet du stride sur la convolution et la dimension de la carte de caractéristiques.	48
2.13	Types de pooling.	49
2.14	Fonctions d'activation.	50

2.15	Architecture U-Net.	52
2.16	Exemple d'architecture du Vision Transformer.	53
2.17	Architecture du réseau antagoniste génératif (GAN).	56
3.1	Visualisation du nombre d'images dans chaque classe de l'échantillon de la base de données.	74
3.2	Exemples d'images normales et de pneumonie acquises à partir de la base de données ChestX-ray14.	75
3.3	Aperçu de la méthodologie proposée.	76
3.4	Dendrogramme produit par la HCA et répartition des données cliniques dans chaque cluster.	78
3.5	Amélioration des images CXR en utilisant le Masque Flou avec les paramètres optimaux déterminés pour chaque cluster.	80
3.6	Architecture de notre modèle CNN-Attention.	82
3.7	Exemples de visualisation de la localisation des maladies pulmonaires.	87
4.1	Exemples d'images de radiographie thoracique présentant des artefacts et des défauts de qualité.	90
4.2	Aperçu de la méthodologie proposée.	92
4.3	Exemple d'images améliorées à l'aide de la technique d'égalisation d'histogramme.	94
4.4	Matrices de confusion de notre approche de clustering pour les classes pneumonie et no finding.	98
4.5	Le modèle CycleGAN et la perte de cohérence cyclique.	100
4.6	L'architecture CycleGAN adaptée dans notre étude pour le scénario des images avec artefacts contre les images sans artefacts.	102
4.7	Images avec et sans artefacts obtenues par le modèle CycleGAN.	103
4.8	Convolutional Block Attention Module (CBAM).	105
4.9	Exemples de visualisation de la localisation des maladies sur les images de test de la pneumonie et de la consolidation.	113
5.1	Aperçu de la répartition des images de la base de données ChestX-ray14.	117
5.2	Aperçu de la méthodologie proposée.	119
5.3	L'architecture du réseau de segmentation U-Net.	120
5.4	Images de radiographie thoracique et leurs masques obtenus par U-Net.	121

TABLE DES FIGURES

5.5	Notre approche proposée pour la génération de masques de pathologie dans les images de radiographie thoracique.	122
5.6	L'architecture CycleGAN adaptée dans notre étude pour le scénario des images avec pathologies contre les images sans pathologies.	123
5.7	Images avec et sans pathologies obtenues par le modèle CycleGAN.	124
5.8	Exemples de génération de masques utilisant notre approche proposée.	125
5.9	Résultats du modèle SHAP montrant les caractéristiques les plus importantes pour la classification multi-label des deux maladies : Effusion et Infiltration.	140
5.10	Courbes ROC et scores AUC de toutes les maladies sur l'ensemble de test de la base de données ChestX-ray14 utilisant le modèle XGBoost et notre approche de génération de masques et de sélection des caractéristiques.	142

LISTE DES TABLEAUX

2.1	Comparaison des bases de données.	33
2.2	Performance des méthodes de classification de la tuberculose utilisées dans la littérature.	61
2.3	Performance des méthodes de classification de la pneumonie utilisées dans la littérature.	62
2.4	Performance des méthodes de classification des nodules pulmonaires utilisées dans la littérature.	62
2.5	Performance des méthodes de classification du Covid-19 utilisées dans la littérature.	63
2.6	Performance des modèles de classification multi-classe des pathologies pulmonaires utilisés dans la littérature.	65
2.7	Performance des modèles de classification multi-label utilisés dans la littérature sur les bases de données MIMIC-CXR et CheXpert.	67
2.8	Performance des modèles de classification multi-label utilisés dans la littérature sur la base de données ChestX-ray14.	68
3.1	AUC de chaque cluster et l'AUC pondérée avec et sans UM pour la classification binaire de la pneumonie.	83
3.2	AUC de la classification binaire des trois maladies avec et sans l'étape de clustering.	84

3.3	Comparaison des résultats obtenus avec d'autres méthodes existantes dans la littérature pour la classification binaire des trois maladies pulmonaires de la base de données ChestX-ray14.	86
4.1	Nombre d'images des classes de pneumonie et de no finding dans chaque cluster obtenues à l'aide du clustering par K-means.	97
4.2	Métriques de performance de notre approche de clustering pour les classes pneumonie et no finding.	98
4.3	AUC de la classification binaire des images de pneumonie et no finding en utilisant les différentes approches.	109
4.4	Performances du modèle DenseNet-121 pour les deux classes, pneumonie et no finding, après l'application du clustering, CycleGAN, et l'intégration du CBAM et des variables cliniques.	110
4.5	AUC de la classification binaire des images de consolidation et no finding en utilisant les différentes approches.	111
4.6	Performances du modèle DenseNet-121 pour les deux classes, consolidation et no finding, après l'application du clustering, CycleGAN, et l'intégration du CBAM et des variables cliniques.	111
4.7	Comparaison des résultats obtenus avec d'autres méthodes existantes dans la littérature pour la classification des deux maladies pulmonaires de la base de données ChestX-ray14.	112
5.1	Nombre d'images catégorisées par type de pathologie dans la base de données ChestX-ray14	118
5.2	Caractéristiques extraites pour chaque groupe de caractéristiques radio-miques.	129
5.3	Caractéristiques sélectionnées pour les différentes conditions pathologiques.	132
5.4	Impact de la méthode de génération de masques sur les résultats de classification multi-label de l'effusion et de l'infiltration. Les valeurs importantes sont en gras.	136
5.5	Comparaison des résultats de la classification multi-label de l'effusion et de l'infiltration en utilisant notre approche de génération de masque avec et sans sélection de caractéristiques. Les valeurs importantes sont en gras.	137

5.6	Résultats détaillés du modèle XGBoost utilisant notre approche de génération de masques avec sélection de caractéristiques (scénario 2), pour la classification multi-label de l'effusion et de l'infiltration.	139
5.7	Comparaison des résultats AUC de notre méthode avec d'autres méthodes existantes dans la littérature pour la classification multi-label des maladies thoraciques de la base de données ChestX-ray14. Les valeurs importantes sont en gras.	142

Titre : Détection et classification multi-label de maladies pulmonaires par apprentissage automatique à partir d'images de radiographie thoracique

Mot clés : Radiographie thoracique, Maladies pulmonaires, Apprentissage automatique, Classification, Segmentation, CycleGAN.

Résumé : Les maladies pulmonaires représentent une cause majeure de décès dans le monde, et le diagnostic précoce est crucial pour améliorer les chances de rétablissement. Les technologies d'Intelligence Artificielle ont ouvert des voies prometteuses dans le domaine biomédical. Ainsi dans cette thèse, des modèles d'IA sont utilisés pour améliorer la performance de classification des maladies pulmonaires à partir des images de radiographie thoracique. De nouvelles approches de prétraitement basées sur CycleGAN sont développées pour réduire l'effet du bruit causé par les artefacts tels que des dispositifs médicaux dans les radiographies thoraciques, ainsi

que pour générer des masques incluant les zones pathologiques dans les régions d'intérêt. Ensuite, une nouvelle approche de sélection de caractéristiques est développée pour identifier a priori les caractéristiques statistiquement les plus significatives avant la classification. Au-delà de l'analyse des images, les données cliniques associées sont également examinées pour affiner le modèle de classification selon le profil du patient, ce qui améliore l'efficacité diagnostique. Les avancées proposées présentent des résultats prometteurs améliorant la performance de la classification binaire et multi-label des maladies pulmonaires.

Title: Detection and multi-label classification of lung diseases from chest x-ray images using machine learning

Keywords: Chest X-ray, Lung disease, Machine Learning, Classification, Segmentation, CycleGAN.

Abstract: Lung diseases are a major cause of death worldwide, and early diagnosis is crucial to improve the chance of recovery. Artificial Intelligence technologies have opened promising avenues in the biomedical field. Thus, in this thesis, AI models are used to improve the classification performance of lung diseases from chest X-ray images. New preprocessing approaches based on CycleGAN are developed to reduce the noise effect caused by artifacts such as medical devices in chest X-rays, as well as to generate masks that in-

clude pathological areas within the regions of interest. Additionally, a new feature selection approach is developed to identify the statistically most significant features a priori before classification. Beyond image analysis, the associated clinical data are also examined to refine the classification model according to the patient's profile, enhancing diagnostic effectiveness. The proposed advancements show promising results in improving the performance of both binary and multi-label classification of lung diseases.