REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE IBN KHALDOUN - TIARET

MEMOIRE

Présenté à :

FACULTÉ DES MATHEMATIQUES ET D'INFORMATIQUE
DÉPARTEMENT D'INFORMATIQUE

Pour l'obtention du diplôme de :

MASTER

Spécialité : Génie informatique

Par :

LABTAR Noura

Sur le thème

Classification binaire des images médicales CT Scan pulmonaires; cas d'étude : Nodulaire et Non Nodulaires.

Soutenu publiquement le17 /06 / 2025 à Tiaret devant le jury composé de :

Mr BELARBI Mostef	Pr	Université of Ibn-Khaldoun-Tiaret	Président
Mr MAZZOUG Karim	MAA	Université of Ibn-Khaldoun-Tiaret	Supervisé
Mr r MERATI Medjeded	MCA	Université of Ibn-Khaldoun-Tiaret	Examinateur

2024-2025

Remerciements

En préambule, nous rendons grâce à Allah, le Tout-Puissant, qui nous a gratifiés de la force, de la patience et de la détermination nécessaires pour mener à bien ce travail. Nous exprimons notre profonde gratitude à toutes les personnes qui, par leur soutien, leurs conseils ou leurs encouragements, ont contribué à l'aboutissement de ce parcours académique.

Nos sincères remerciements s'adressent à notre encadrante, le ***Mr Mazzoug Karim****, dont les orientations avisées et le suivi rigoureux ont grandement enrichi ce travail.*

Nous témoignons également notre vive gratitude aux membres du jury, ***Mr BELARBI Mostef*** *et* ***Mr MERATI Medjeded*** *, pour le temps qu'ils ont consacré à l'évaluation de notre travail et pour leurs remarques constructives qui en ont rehaussé la qualité.*

Nous n'oublions pas de saluer l'ensemble de nos enseignants, dont l'érudition et la générosité ont éclairé notre cheminement intellectuel tout au long de notre formation.

Enfin, nous adressons nos remerciements les plus chaleureux à nos familles et à nos proches, dont le soutien indéfectible et les encouragements constants ont été une source inépuisable d'inspiration tout au long de ce projet.

Dédicace

À mon père, Labtar Benchikh, Merci pour ta présence solide, tes conseils calmes, et ton soutien discret mais essentiel. Ta confiance en moi m'a portée plus loin que tu ne peux l'imaginer.

À ma mère, Un merci infini et particulier. Ton amour, ta tendresse et tes prières ont été le cœur battant de mes réussites. Tu es mon refuge, ma lumière, et la plus belle bénédiction que la vie m'ait offerte. Que Dieu te protège et t'élève en dignité et en paix.

À mes frères, Atia et Benaissa , et Ma soeur Retadje Merci d'avoir été un véritable pilier dans ce parcours. Ta présence, ton aide concrète et ton soutien moral ont allégé bien des fardeaux. Je te suis profondément reconnaissante. À mes amies,Merci d'avoir été là avec vos mots doux, vos encouragements sincères, votre écoute précieuse. Votre amitié a été un souffle d'air frais dans les moments les plus durs.

Résumé

Ce mémoire s'intéresse à la classification binaire des images médicales CT scan thoraciques, avec une focalisation spécifique sur la distinction entre le tissu pulmonaire nodulaire localisé et le tissu pulmonaire non nodulaire localisé. Cette tâche est particulièrement importante dans le cadre du dépistage précoce des nodules pulmonaires, souvent associés à des maladies graves telles que le cancer du poumon.

Face aux limites des approches globales classiques — déséquilibre des classes, similarité texturale, et coûts computationnels élevés — nous avons proposé une stratégie locale, plus ciblée, permettant de concentrer le processus d'analyse sur les zones d'intérêt uniquement.

Pour cela, nous avons développé une application nommée SANADv1, capable d'extraire automatiquement les régions pertinentes des images CT. Ensuite, trois modèles de réseaux de neurones convolutifs (CNN) ont été testés : ResNet50, MobileNetV3 et ConvNeXt, chacun entraîné avec les optimiseurs Adam, SGD et RMSprop.

Les expériences ont été réalisées sur Google Colab avec GPU Tesla T4, en utilisant un sous-ensemble d'images de la base LIDC-IDRI, converties en format PNG et normalisées. La

meilleure performance a été obtenue avec le modèle ConvNeXt optimisé par RMSprop, atteignant une précision de validation de 91 %. Ce résultat valide l'efficacité de notre approche locale pour la classification fine des images nodulaires.

Mots-clés : IA, Apprentissage profond, Traitement d'images médicales, Classification Binaire .

Abstract

This thesis focuses on the binary classification of chest CT scan medical images, with a specific focus on distinguishing between localized nodular lung tissue and localized non-nodular lung tissue. This task is particularly important in the early detection of pulmonary nodules, which are often associated with serious diseases such as lung cancer. Faced with the limitations of traditional global approaches—class imbalance, textural similarity, and high computational costs—we proposed a more targeted, local strategy that allows us to focus the analysis process on areas of interest only. To this end, we developed an application called SANADv1, capable of automatically extracting relevant regions from CT images. Three convolutional neural network (CNN) models were then tested : ResNet50, MobileNetV3, and ConvNeXt, each trained with the Adam, SGD, and RMSprop optimizers. The experiments were conducted on Google Colab with a Tesla T4 GPU, using a subset of images from the LIDC-IDRI database, converted to PNG format and normalized. The best performance was achieved with the ConvNeXt model optimized by RMSprop, achieving a validation accuracy of 91 %. This result validates the effectiveness of our local approach for fine-grained lung tissue classification.

Keywords : Artificial intelligence, Deep learning, Medical image processing, Binary classification.

ملخص

تركز هذه الرسالة على التصنيف الثنائي لصور الأشعة المقطعية للصدر الطبية، مع التركيز بشكل خاص على التمييز بين أنسجة الرئة العقدية الموضعية وأنسجة الرئة غير العقدية الموضعية. تُعد هذه المهمة بالغة الأهمية في الكشف المبكر عن عقيدات الرئة، والتي غالبًا ما ترتبط بأمراض خطيرة مثل سرطان الرئة.

ونظرًا لقيود المناهج العالمية التقليدية - اختلال التوازن الطبقي، والتشابه النسيجي، والتكاليف الحسابية العالية - فقد اقترحنا استراتيجية محلية أكثر استهدافًا تركز عملية التحليل على مجالات الاهتمام فقط.

ولتحقيق هذه الغاية، طورنا تطبيقًا يسمى SANADv1، قادرًا على استخراج المناطق ذات الصلة تلقائيًا من صور الأشعة المقطعية. ثم تم اختبار ثلاثة نماذج للشبكات العصبية التلافيفية (CNN): ResNet50 وMobileNetV3 وConvNeXt، وقد تم تدريب كل منها باستخدام مُحسِّنات Adam وSGD وRMSprop. أُجريت التجارب على Google Colab باستخدام وحدة معالجة الرسومات Tesla T4، باستخدام مجموعة فرعية من الصور من قاعدة بيانات LIDC-IDRI، تم تحويلها إلى تنسيق PNG وتطبيعها.

تم تحقيق أفضل أداء باستخدام نموذج ConvNeXt المُحسّن بواسطة RMSprop، محققًا دقة تحقق بلغت 91%. تُثبت هذه النتيجة فعالية نهجنا المحلي لتصنيف أنسجة الرئة الدقيقة..

الكلمات المفتاحية: الذكاء الاصطناعي، التعلم العميق، معالجة الصور الطبية، التصنيف الثنائي.

Table des matières

Table des figures

Liste des tableaux

Liste des abréviations

Abréviation	Signification
AI	Intelligence Artificielle
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
SVM	Support Vector Machine
SGD	Stochastic Gradient Descent
Adam	Adaptive Moment Estimation
RMSprop	Root Mean Square Propagation
ConvNeXt	Convolutional Network Next
MobileNet	Réseau neuronal léger pour mobile
ResNet	Residual Neural Network
CT Scan	Tomodensitométrie (Computed Tomography)

Introduction générale

Au cours des dernières décennies, l'intelligence artificielle (IA) s'est imposée comme une révolution technologique majeure, transformant profondément de nombreux domaines, y compris celui de la santé. Dans ce contexte, l'analyse automatique des données d'imagerie médicale constitue un levier fondamental pour améliorer la précision, la rapidité et la fiabilité des diagnostics cliniques.

Parmi les techniques d'imagerie, la tomodensitométrie thoracique (CT scan) représente un outil de référence dans la détection des affections pulmonaires. Toutefois, l'interprétation manuelle de ces images est souvent longue, sujette à des erreurs humaines, et dépend fortement de l'expertise du radiologue. D'où l'intérêt croissant pour des systèmes intelligents capables d'assister voire d'automatiser certaines étapes du diagnostic.

Ce mémoire s'inscrit dans cette perspective, en proposant un système de classification binaire locale visant à distinguer, à partir d'images CT thoraciques, entre les tissus pulmonaires nodulaires localisés et les tissus non nodulaires localisés, notre méthode repose sur une analyse focalisée, centrée sur les régions d'intérêt détectées automatiquement.

Pour cela, nous avons conçu une application nommée SANADv1, qui permet de localiser les zones pertinentes dans les images et d'y appliquer des modèles de classification basés sur des réseaux de neurones convolutifs (CNN). Plusieurs architectures ont été étudiées, notamment ResNet50, MobileNetV3, et ConvNeXt, associées à différents optimiseurs (Adam, RMSprop, SGD) pour identifier la combinaison offrant les meilleures performances.

Ce projet s'articule autour de quatre chapitres complémentaires :

Le Chapitre 1 : présente les concepts fondamentaux de la vision artificielle et du traitement d'images médicales.

Le Chapitre 2 : dresse un état de l'art des solutions existantes pour la classification d'images thoraciques.

Le Chapitre 3 : aborde les fondements de l'intelligence artificielle, de l'apprentissage automatique et des réseaux de neurones profonds.

Le Chapitre 4 : expose l'implémentation pratique de notre système SANADv1, les résultats expérimentaux obtenus, et leur interprétation.

Ce travail vise à démontrer que l'approche locale, orientée nodules, permet une classification plus fine et mieux adaptée aux enjeux du dépistage assisté par ordinateur (CAD) dans le domaine pulmonaire.

Chapitre I

Vision Artificielle

I.1 Introduction

La vision artificielle est une technologie qui permet aux machines d'analyser des clips vidéo. Les images sont utilisées dans des processus de prise de décision automatisés grâce à l'apprentissage automatique. Traitement approfondi des données et des images. Il existe des applications dans de nombreux domaines.

Des domaines tels que la sécurité, la robotique, et les systèmes de conduite autonome, et bien sûr, médicament.

Dans le domaine médical, il joue un rôle essentiel dans le traitement et l'analyse par imagerie, le diagnostic et le diagnostic des maladies. cause du tournage médical, améliorer de manière significative la précision du diagnostic, tout en particulier le taux d'erreur médicale.

I.2 Vision Artificielle

I.2.1Définition

En intelligence artificielle, la vision par ordinateur (ou vision artificielle) est une branche essentielle qui comprend la capacité d'analyser et d'analyser.

Photos ou les séquences vidéo sont générées automatiquement, dans un format accessible à la perception. tion. Protection de l'environnement humain. La vision artificielle repose sur l'utilisation. Algorithmes avancés de traitement d'image. Et des techniques d'apprentissage en profondeur. Cela a aidé les systèmes. Intelligent pour comprendre le contenu visuel et en extraire des idées Informations pertinentes à partir de photos ou de vidéos ci-jointes.

Cette technologie ils voulaient que la machine soit capable de détecter le contenu visuel avec plus de précision. Souvent, cela dépasse le niveau humain, ce qui donne lieu à des applications réseau de grande envergure. Opérations complexes, de l'imagerie médicale à la réalité augmentée, y compris : l'automobile indépendant. L'intégration de la vision artificielle réduit le besoin de intervenir. Humain, augmentant l'efficacité et la précision dans de nombreux domaines Des secteurs vitaux tels que la médecine, l'industrie, la sécurité et les transports. Malgré les défis, Progrès rapides dans l'apprentissage profond, en re- Nous avons équipé des réseaux de Centres d'ingénierie des réseaux neuronaux convolutifs (CNN) et de la vision artificielle AMERA Atteindre un niveau élevé de performance d'apprentissage.

Il est désormais prêt à être mis en œuvre. Acquérir la capacité de traiter des images complexes et des données non structurées, et déverrouiller La porte vers des applications futures plus intelligentes et plus avancées dans un large éventail de secteurs.

I.1.2 Principes Fondamentaux de la Vision Artificielle :

La vision artificielle repose sur plusieurs principes fondamentaux qui permettent aux machines d'analyser et d'interpréter des images ou des vidéos. La figure I.1 illustre certains de ces principes :

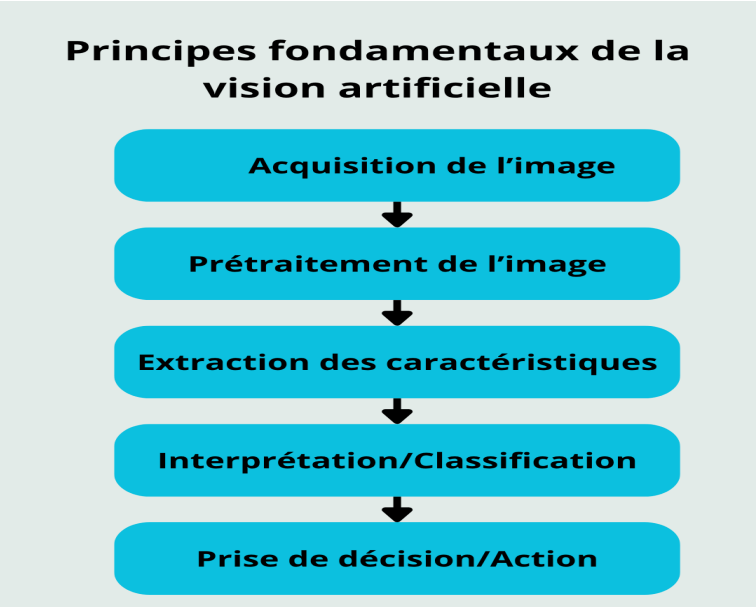

Fig. I.1 : Principes Fondamentaux de la Vision Artificielle

I.1.3 Applications de la vision artificielle

La vision artificielle est l'un des principaux drivers de l'investissement dans l'intelligence artificielle. Sa signification et ses conditions générées par ordinateur ont créé son émergence et ont fait place pour son émergence comme un IA mondial.

La technologie de la vision a connu son émergence majeure, alors que les applications daily de l'apprentissage profond et des réseaux neuronaux lui ont donné pouvoir exécutif pour une tâche complète à effectuer, même à des fins de démonstration. L'émergence de l'impression 3D, de la classification d'images multimédia et de poursuite adapte du contenu des idées et des objets a également été observée.

La vision artificielle a été employée dans une très grande variété d'applications, comme la médecine (analyse rapide des doses d'IRM), la sécurité (caractéristiques de la vision artificielle), l'industrie (surveillance à distance des lignes de production et des fichiers) et même en conduite autonome. Avec ces multiples applications, la vision artificielle est toujours l'un des éléments les plus essentiels contribuant en grande partie à l'amélioration de la vie humaine et au développement de la majeure partie des industries.

I.1.4 Défis, enjeux et perspectives de la vision artificielle

Parmi les derniers acquis, la vision artificielle continue de faire face à un ensemble de défis qui supposent des solutions créatives pour élargir son domaine d'application. Techniques et pratiques toutes deux, ces défis laissent ce domaine de recherche et de développement marqué par un fort intérêt.

1. **Défis techniques :**

   •**Traitement de grandes masses de données :** La vision artificielle couvre le traitement de très grosses masses de données (images et vidéos haute définition) ce qui est un réel problème en terme de capacité de traitement et de stockage intermédiaire

•**Environnement changeant :** Les systèmes de vision artificielle sont parfois confrontés à des conditions environnementales (obscurité, angles de vue extrêmes) diminuant considérablement la précision des modèles.Par exemple, la la reconnaissance d'objets devient impossible en cas de faible luminosité ou in situationes inattendus.

•**Généralisation des modèles :** Bien qu'il existe des méthodes diverses qui se révèlent efficaces dans un environnement spécifique, leur généralisation à d'autres si- les conditions et les conditions sont mauvaises. leur généralisation à d'autres situations et conditions est problématique.

2. **Questions éthiques et sociales :**

   •**Biais :** Algorithmes based on déséquilibrées données sont du même côté, et ceci peut contribuer aux catégorisation erreurs sur l'origine ethnique, le sexe, ou l'âge. Ce risque est significatif dans des domaines sensibles tels que la santé ou la sécurité.

   •**Sécurité et confidentialité :** La mise en application à grande échelle de la reconnaissance faciale et de la surveillance souleva des questions d'ordre à la vie privée et à la protection des données. Il est capital d'assurer que ces systèmes sont au service des droits individuels et gardent les données personnelles contre tout accès illicite.

3. **Perspectives d'avenir :**

   •**Évolution de l'intelligence artificielle :** La vision artificielle continuera de s'améliorer grâce à l'apprentissage profond et à des méthodes sophistiquées d'intelligence artificielle, afin d'améliorer l'efficacité dans des environnements complexes et la généralisation des modèles.

   •**Applications dans des secteurs émergents :** Cette technologie trouvera des applications dans de nouveaux secteurs comme la santé, les voitures intelligentes, la robotique et même dans les secteurs militaire et de la sécurité.

   •**Réalité virtuelle et augmentée :** Ces technologies joueront un rôle de premier plan dans le interface design utilisateur plus immersif et naturel, qui offre de nouvelles opportunités à l'industrie, à l'éducation, à l'éducation sanitaire et à la demande mondiale.

I.1.5 Les avantages

La vision artificielle est l'une des applications les plus importantes de l'intelligence artificielle. Elle permet aux systèmes et aux architectures matérielles d'intégrer et d'analyser le texte et les images à partir d'une séquence vidéo avec une visualisation à peine distante de celle d'un humain L'un de ses avantages majeurs est l'automatisation des processus de reconnaissance visuelle, de diagnostic et de surveillance, ce qui améliore l'efficacité, réduit les erreurs humaines et permet de gagner du temps et de réduire les coûts.

I.2 Classification binaire

I.2.1 Définition

La classification binaire est un simple type de fonction dans le paradigme appris de la classification supervisée, auquel cas tâche prévisible est de classer ce qui constitue une classe

subordonnée à celle en laquelle se trouve une donnée appartenant.

L'implémentable du schéma de classification binaire a ainsi le rôle d'assigner deux classes discontinues basée sur caractéristiques de l'entrée. Cette méthode est de manière forte appliquée dans de nombreuses applications, telles que la reconnaissance des images, le frai detection, le diagnostic médical, et encore bien d'autres.

La figure I.2 illustre la classification binaire et la classification multiclasse.

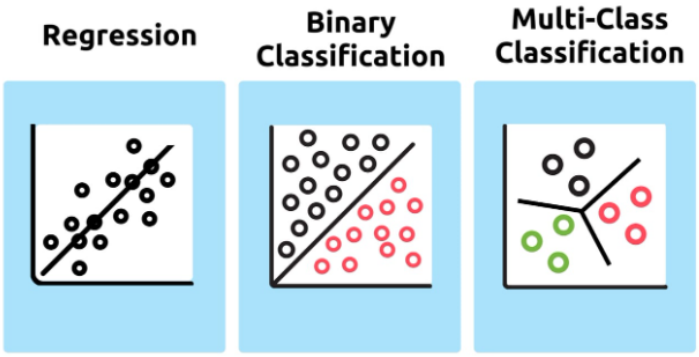

Fig. I.2 : Classification **[1]**

I.2.2 Présentation générale

Dans le cadre de la classification binaire, le modèle est entraîné à partir d'un ensemble de train data qui comprennent des exemples étiquetés qui sont de l'une des deux classes.Chaque exemple d'entraînement est décrit par un vecteur de caractéristiques, et l'étiquette correspondante indique l'appartenance à l'une des deux classes. L'objectif est de trouver une frontière de décision optimale qui sépare les deux classes dans l'espace des caractéristiques.

Les modèles de classification binaire les plus couramment utilisés sont les Support Vector Machines (SVM), les réseaux neuronaux, et les classificateurs à base d'arbres de décision, parmi d'autres. L'entraînement de ces modèles se fait généralement en minimisant une fonction de perte, telle que l'entropie croisée, qui évalue la qualité de la séparation entre les classes.

I.2.3 Objectifs et applications principales :

La classification binaire vise plusieurs objectifs clés :

1. Pésdiction de l'appartenance à une classe :

Déterminer la classe (parmi deux) d'une nouvelle donnée à partir de ses caractéristiques.

2.Séparation optimale des classes :

Optimiser les frontières dé décision (ex : SVM) pour améliorer la précision.

I.2.4 Applications Principaux

1.Diagnostic médical :

Ex : Détection du cancer a partir d'images médicales (radiographies, IRM ...).

2.Reconnaissance de fraude :

Ex : Détection de transactions suspectes dans les systèmes bancaires .

3. Sécurité et surveillance :

Ex : Reconnaissance faciale dans des systèmes videosurviellance .

4.Systèmes de recommandation :

Ex : Prédire si un utilisateur va aimer .

I.2.5 Les avantages :

1.Facile à réaliser :

Le plus souvent, un modèle binaire est une famille de modèles de conception et de formation issus de divers modèles catégoriels.

2.Quadrilatère ou interprétatif :

Les résultats sont d'habitude laissés à la compréhension facile : une catégorie ou l'autre.

3.Calculs de performances :

la plupart du temps, il y a moins de combinaisons (fautes) qui font moins de calculs, et cela veut dire une exécution plus rapide.

Travail bien joué sur les tâches demandées : en résolvant correctement le problème avec deux classes différentes (malade ou en bonne santé), les performances sont très élevées.

I.2.6 Les limites

1.Inflexibilité :

Le modèle ne tolère pas plus d'une classe sans changements structurels importants.

2.Sensibilité au déséquilibre des classes :

en triangulant une seule classe intelligemment, le modèle peut sembler biaisé.

3.Simplification excessive du problème :

certains phénomènes changeants ne peuvent pas être décrits avec précision par une dichotomie binaire.

4.En fonction de la bonne définition du seuil de décision :

La définition du seuil de décision peut affecter significativement les performances des modèles probabilistes.

I.2.7 Comparaison entre les différents types de classification

Avant de revenir sur les modèles et les approches appliquées dans notre projet, il semble inévitable de connaître les principaux types de classification supervisée en apprentissage automatique. Il s'agit de types de classification qui varient principalement en fonction du nombre de catégories possibles et sur la nomenclature possible pour étiqueter un cas.

La classification peut donc être binaire, multi-classes ou encore multi-étiquettes. Le tableau suivant donne une comparaison synthétique de ces approches afin d'expliquer notre choix méthodologique dans le contexte de notre problématique.

Le tableau I.1 montre la comparaison entre différents types de classification :

Tab. I.1 : Comparaison entre les différents types de classification

Critère	**Classification Binaire**	**Classification Multi-classes**	**Classification Multi-label**
Nombre de classes	2	≥ 3	≥ 2
Nombre de classes par instance	1	1	≥ 1
Exemple de sortie	[0] ou [1]	[0], [1], ..., [n]	[0,1,1,0]
Exemple d'application	Détection de spam	Reconnaissance d'objets	Annotation d'images (multi-étiquettes)
Complexité	Faible	Moyenne	Élevée
Modèles adaptés	SVM, Régression logistique	CNN, Forêt aléatoire	Réseaux de neurones multi-sorties
Méthodes d'évaluation	Accuracy, F1-score	Matrice de confusion, Top-k accuracy	Hamming loss, F1 macro/micro

I.2.8 Métriques de classification binaire

Le classement d'un classificateur binaire repose sur un cadre de mesure pour évaluer sa performance en fonction de sa précision, rappel, et justesse globale. Les mesure sont évalués à partir de la prédiction du modèle par rapport à l'étiquette réelle, en fonction de la matrice de confusion.

Les mesure les plus importantes utilisées sont :

1. **Accuracy** :
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
2. **Précision (Precision)** :
$$Precision = \frac{TP}{TP + FP}$$
3. **Rappel (Recall)** :
$$Recall = \frac{TP}{TP + FN}$$
4. **F1-score** :
$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

II.9 Modèles utilisés dans les classificateurs binaires :

Différents formes d'automatisé apprentissage et d'apprentissage profond sont employés dans le projet pour classer les images CT Scan dans les images nodulaires et non nodulaires selon classement binaire. La figure I.3 illustre les principaux modèles utilisés :

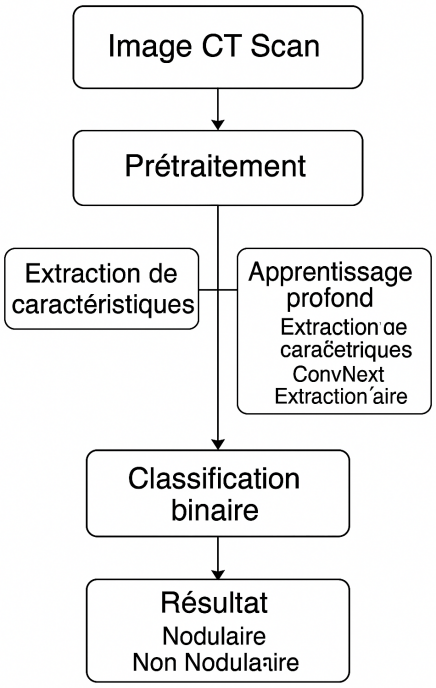

Fig. I.3 : Modèles dans les classificateurs binaires

I.3 Traitement d'image

I.3.1 Définition

Le traitement de l'information est un sous-domaine du domaine informatique spécifique à la recherche et à l'analyse d'images numériques dont le but est de les capturer avec des informations utiles ou de les recoder afin qu'elles soient prêtes ou prêtes à être étendues. Le processus implique des techniques et des algorithmes mathématiques évolués qui sont souvent associés avec des logiciels ou des systèmes informatiques.

Le traitement d'image implique plusieurs étapes de base, commençant par l'acquisition, suivie du traitement initial, comme la suppression du bruit et l'amélioration de l'éclairage et du contraste. Vient ensuite l'étape d'analyse de l'image, qui comprend la segmentation de l'image, l'extraction des caractéristiques et enfin l'étape de classification ou d'interprétation.

À l'intérieur du domaine médical, traitement d'images joue un rôle important pour soutenir la prise de décision des médecins en introduisant l'analyse accrue d'images provenant de toutes les sortes de la thérapeutique médicale utilisant les rayons X, l'imagerie par résonance magnétique (IRM) et la tomodensitométrie (TDM). Les prochaines innovations viennent s'ajouter à la détection précoce des maladies, à la localisation de lésions précise, et à l'observation de

l'évolution des maladies, pour les rendre un élément important des systèmes de diagnostic aidé par ordinateur (CAO).

I.3.2 Types d'images médicales

L'imagerie médicale est un élément essentiel du diagnostic, du suivi et du traitement des maladies. Elle consiste à obtenir des images à l'aide de diverses techniques, telles que les rayons X, l'échographie, la tomodensitométrie (TDM), l'imagerie par résonance magnétique (IRM) et d'autres techniques permettant une visualisation précise des structures internes du corps.

La figure I.4 illustre quelques types d'imagerie médicale.

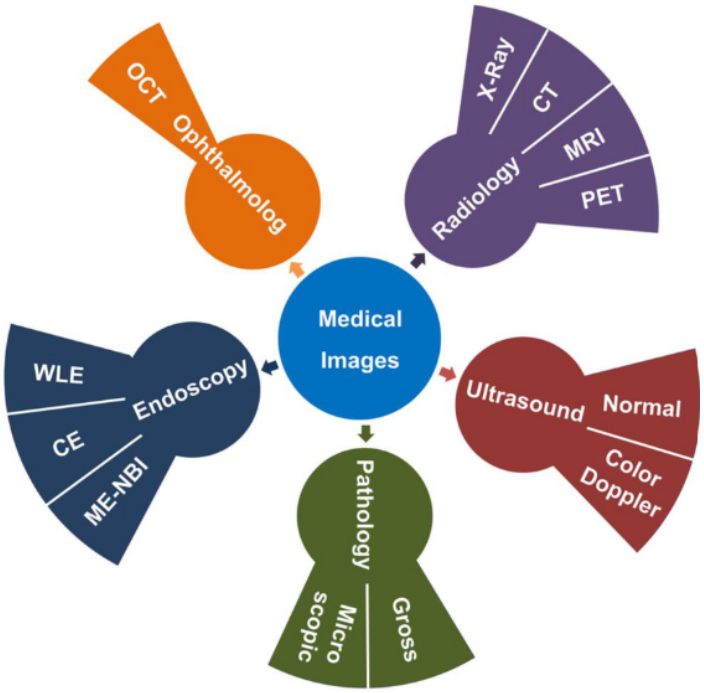

Fig. I.4 : Types Images Médicales

I.3.3 Prétraitement d'images médicales

Pre-processing des images médicales est l'une des informations le plus élémentaire et la plus cruciale avant le traitement en soi ou l'extraction des caractéristiques. Il essaie de réduire la qualité de l'image en utilisant le post-traitement, le segmentation et la classification. Les images médicales, quand elles sont transmises envers un scanner (TDM), peuvent contenir le bruit, les artefacts ou les conditions d'éclairage variables qui peuvent affecter l'efficacité des algorithmes de traitement d'image.

Les principales étapes de prétraitement sont :

1. **Réduction du bruit :** vous pouvez appliquer des filtres traditionnels tels que médian ou gaussien pour supprimer les pixels qui peuvent être décalés afin de conserver les

bord triangulaires des structures anatomiques.

2. **Les effets de contraste :** Votre image est retrouvée sous la forme d'histogramme ou CLAHE (Contrast-Limited Adaptive Histogram) par Českés pour les détails internes, et sterone pour les textures à faible contraste.

3. **Redimensionnement de conversion :** les images bmr.spotifyicity sont redimensionnées à une résolution ordinaire (par exemple 224 × 224 ou 256 × 256 pixels) pour l'impulsion sur des CNN.

4. **Convergence :** les algorithmes de ce type peuvent amener les valeurs des pixels à un intervalle uniforme (généralement le cercle entre 0 et 1) pour engendrer une convergence plus stable et rapides des modèles pendant l'entraînement.

5. **Segmentation initial :** Finalement, une découpe approximative des zones d'intérêt (comme les poumons dans un scanner) est réalisée avant leur diag

Cela permet d'obtenir des images sans-distorsions, homogènes et prêtes à être automatiquement interprétées par des algorithmes de machine learning ou de deep learning.

I.3.4 Cadre général d'un système d'interprétation des images médicales

Un système d'interprétation d'images numériques peut être divisé en plusieurs étapes. Une phase de prétraitements qui suit l'acquisition et la numérisation de l'image. Elle permet essentiellement de réduire une quantité importante de bruits. La phase de segmentation consiste à isoler les uns des autres les objets présents dans l'image. Suite à cette étape, vient l'étape de l'interprétation qui vise à reconnaître les pathologies recherchées dans le but d'aider le médecin dans son diagnostic **[2]**

La figure I.5 illustre la chaîne de traitement et d'analyse des images médicales :

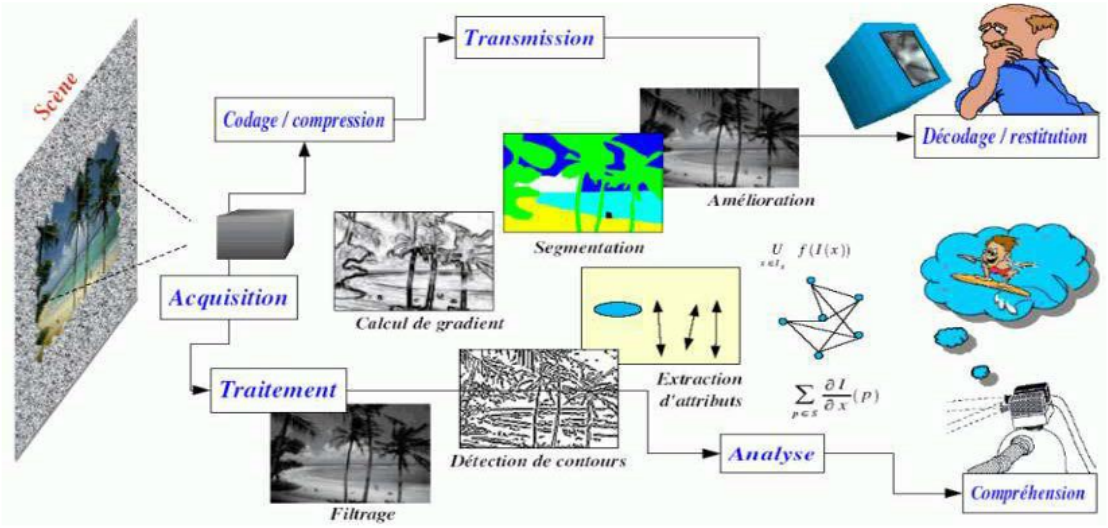

Fig. I.5 : La chaîne de traitement et d'analyse des images médicales.**[2]**

I.3.5 Techniques d'amélioration d'image

Les images médicales peuvent être bruitées, artefacter ou présenter un faible contraste. Pour les corriger, de plusieurs techniques de prétraitement sont appliquées :

Réduction de bruit : La mise en œuvre de filtres (médian, gaussien) permet d'effacer les pixels parasites sans détruire les contours essentiels.

Amélioration de contraste : Des techniques telles que l'égalisation d'histogramme ou CLAHE améliorent la lisibilité des structures internes.

Redimensionnement : Les images sont plus grandes et redimensionnées à une taille standard pour être compatible avec les modèles d'apprentissage profond.

Ces trois pas ont l'effet que les images sont plus utilisables par les algorithmes de traitement.

.

I.3.6 Normalisation et augmentation

En traitement d'images médicales, la normalisation et l'élargissement des données sont deux étapes de base, notamment si on travaille avec des modèles d'apprentissage automatique ou profond.

La normalisation a pour rôle de donner à des pixels d'image des valeurs pour les garder dans un intervalle homogène, généralement entre 0 et 1, ou selon une loi centrée réduite (moyenne zéro et écart-type égal à 1). Ce traitement est indispensable car il permet de stabiliser et de raccourcir la formation des modèles pour assurer que les données d'entrée ont des propriétés numériques cohérentes. Sans normalisation, les algorithmes donneront des résultats biaisés ou dévieront avec un certain effort.

Une augmentation des données (Data Augmentation), en revanche, est une technique qui enrichit artificiellement l'ensemble de formation par la génération de nouveaux exemples à partir d'images originales. Cette technique permet de répondre à l'indisponibilité des données étiquetées, un scénario fréquent en médecine. Les changements normaux sont le retournement, la rotation, le zoom, la modification de mise à l'échelle, la translation, l'addition de bruit ou l'élasticité.

En préservant la diversité de l'échantillon, le processus tend à permettre une plus grande capacité de généralisation du modèle, moins de surapprentissage et une résistance accrue à la variation des images naturelles des données médicales.

I.3.7 Caractéristiques extraction

La déttraction des caractéristiques est un stade critique du traitement d'image en passe de transit au sein des systèmes de classification, et ainsi de suite. Il s'agit de conduire l'information visuelle brute des images médicales à un ensemble de descripteurs numériques révélateurs et significatifs des structures examinées.

Ergo, ces traits peuvent être texturiers, capturant des piquer et monter de niveaux de gris (par matrices de cooccurrence – GLCM, ou modèles locaux – LBP), morphologiques, capturant

la forme, la taille ou la compacité des objets (par exemple, les nodule pulmonaires), ou statistiques, telle que la moyenne, écart-type, skewness ou la kurtosis de l'intensité du pixel. Dans certains cas, des caractéristiques de fréquence (par exemple, les coefficients de Fourier ou de Wavelet) peuvent également être extraites.

Ces descripteurs sont ensuite utilisés en tant que propositions d'entrée pour les classificateurs (réseaux neuronaux, SVM, forêts aléatoires, etc.). La performance d'un système d'aide au diagnostic repose surtout sur la pertinence et sur la qualité des caractéristiques extraites. Les réseaux de neurones convolutifs (CNN) autorisent désormais ces derniers temps à extraire automatiquement des caractéristiques de haut niveau directement à partir d'images brutes, ce qui améliore sensiblement la performance globale des systèmes intelligents.

I.3.8 Segmentation

La segmentation est l'un des processus les plus essentiels et les plus compliqués de traitement d'images médicales. Elle se caractérise par le processus du découpage d'une image en sous-zones distinctes et homogènes pour délimiter les structures cibles, comme les organes, les lésions ou les nodules. Cette étape n'a pas uniquement pour rôle de restreindre la région à traiter, mais aussi pour spécifier la localisation des régions pathologiques, le chargeant d'une analyse plus complexe.

Les algorithmes de segmentation sont divisés en deux catégories : les algorithmes classiques et les algorithmes basés sur l'intelligence artificielle. Les algorithmes classiques sont les algorithmes de seuillage (thresholding), la détection de contours (par exemple : Canny, Sobel), la croissance des régions (region Growing), et les algorithmes de regroupement (par exemple : k-means). Bien qu'elles soient rapides et très simples, elles sont en général sensibles au bruit et aux changements d'éclairage.

Les méthodes actuelles sont établies sur l'apprentissage en profondeur, telles que les architectures de type U-Net, Mask R-CNN ou DeepLab, qui ont modifié le train de la segmentation proposant une précision à profusion même dans des situations complexes. Ils peuvent capturer le contexte spatial et le plus petit détail en raison de couches convolutives et d'un mécanisme d'attention. In medical context, a precise segmentation is crucial to accurately measure the size and position of an anomaly, and to feed efficiently the automated diagnostic systems.

La figure I.6 illustre la segmentation d'image :

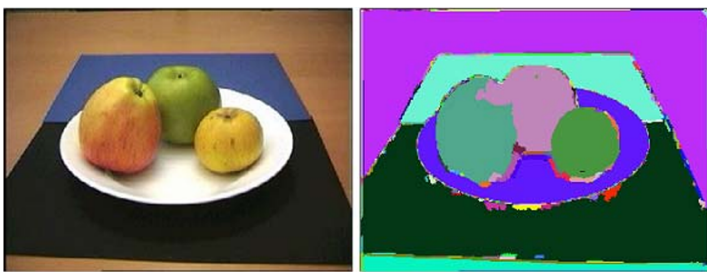

Fig. I.6 : Segmentation.**[2]**

I.4 Conclusion

Le traitement d'image médicale constitue une étape essentielle dans la chaîne de diagnostic assisté par ordinateur. À travers ce premier chapitre, nous avons mis en évidence l'importance croissante des techniques de traitement d'image dans le domaine médical, en particulier pour l'analyse des données issues de l'imagerie par scanner (CT Scan).

Nous avons exploré les différentes phases du traitement, allant de l'acquisition et la pré-traitement des images, jusqu'aux étapes de segmentation, d'extraction de caractéristiques et d'amélioration de la qualité visuelle. Ces techniques permettent de mieux isoler les régions d'intérêt, d'enrichir l'information visuelle et de faciliter les étapes ultérieures de classification et d'interprétation automatique.

Ce chapitre nous a permis de comprendre les défis spécifiques liés à l'imagerie médicale, tels que le bruit, la variabilité anatomique ou encore la nécessité de précision élevée. Il sert ainsi de socle pour appréhender l'utilisation des algorithmes d'intelligence artificielle dans les chapitres suivants, en particulier dans le cadre de la classification automatisée des images médicales.

Chapitre II

État de l’art

II.1 Introduction

Au cours des dernières années, l'essor de l'intelligence artificielle, et plus particulièrement de l'apprentissage profond, a profondément transformé le domaine de la vision par ordinateur, notamment dans les applications médicales. Le traitement et la classification automatique des images médicales représentent un axe de recherche stratégique, notamment pour améliorer la précision, la rapidité et l'efficacité du diagnostic médical assisté par ordinateur (CAD). Dans ce chapitre, nous présenterons un état de l'art des principales approches, modèles et travaux scientifiques qui ont marqué ce domaine.

II.2 InferRead CT Pneumonia – Infervision[3]

InferRead CT Pneumonia est un système de diagnostic assisté par ordinateur basé sur l'IA conçu pour détecter automatiquement les signes de symptômes de pneumonie à partir d'images de tomodensitométrie thoracique, en particulier pour les patients atteints de COVID-19.

II.2.1 Cercles de base

Reconnaître les signes visuels de la pneumonie (tels que l'opacité vitreuse). Générez des cartes thermiques qui mettent en évidence les zones suspectes dans les poumons. En fonction de la précision et de la gravité de la lésion pulmonaire, automatiquement. Intégrer les technologies d'intelligence artificielle dans le modèle de travail des hôpitaux (PACS/RIS).

II.2.2 Méthodes utilisées

Réseaux de neurones convolutifs (CNN).

Les données de formation des images CT sont valables pour un examen par un médecin agréé.

Système rapide qui donne des résultats en quelques secondes.

II.2.3 Contre-indications

Distinguer la pneumonie de la COVID-19.

Suivi de l'état de santé des patients pendant leur convalescence.

rientation urgente vers les services d'urgence.

II.2.4 Faiblesses(Limitations) d'InferRead CT Pneumonie – Infervision

L'application promet de croire que les algorithmes sont importants pour l'intelligence artificielle, ce qui n'est pas open source, ce qui rend difficile la vérification ou l'évaluation du comportement des décisions violatrices. La formation est limitée à une population spécifique de données.

La majorité des données de formation proviennent de Chine en raison du COVID-19, ce qui peut conduire à ce qu'un modèle précis utilisé soit moins respectueux des données en raison du pays ou de la structure de la maladie.

Référentiel de qualité d'image Les performances dépendent en grande partie de la qualité des images radiographiques ; Des images floues ou de faible résolution peuvent donner à votre texte un aspect inexact.

Le fléau des résultats faussement positifs ou faussement négatifs Bien que Mustafa soit utile, il ne remplace pas les besoins d'un médecin et peut parfois donner des résultats incorrects. Nécessite une intégration technique avec Framework NSTSQ zmianoYST return NSTSQ return zmiano NSTSQ return return Cela nécessite une intégration avec plusieurs systèmes d'information médicale (PACS/RIS), ce qui peut être complexe ou coûteux dans certains environnements.

Article de prix d'utilisation Les systèmes commerciaux sont souvent coûteux pour les petites organisations de soins de santé ou dans les pays en développement.

La figure II.1 montre l'interface InferRead CT Pneumonia :

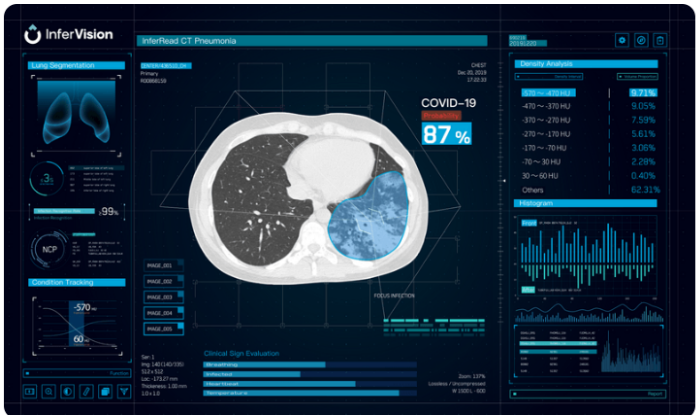

Fig. II.1 : InferRead CT Pneumonia.**[3]**

II.3 Aidoc – Embolie pulmonaire AI [4]

II.3.1 Revue générale

Aidoc est un système d'IA médicale basé sur un fournisseur, approuvé par la FDA, utilisé pour accélérer le diagnostic de l'embolie pulmonaire (EP) en analysant les images d'angiographie pulmonaire par tomodensitométrie (CTPA). Ce programme est spécialement conçu pour une utilisation en cas d'urgence, agissant comme assistant d'un médecin visiteur pour diagnostiquer rapidement les cas critiques.

II.3.2 Fonctions de base

1. automatique detection of the pulmonary artery blood clots.

2. generation of an in-real-time alert to the doctor in case of the detection of the existence of an obstruction.
3. Intégration automatique du résultat dans le système d'information PACS/RIS.
4. du temps de diagnostic et de l'optimisation du processus de réduction-décision dans les cas urgents.

II.3.3 technologies utilisées

1. Réseaux neuronaux profonds (Deep CNN).
2. Analyse 3D des images CTPA.
3. Méthodes d'apprentissage automatique pour les utilités de milliers de cas enregistrés cliniquement.

II.3.4 Avantages

1. L'importance d'obtenir des diplômes d'urgence pour la rapidité et la précision du diagnostic.
2. Taux élevé de faux négatifs.
3. Il s'agit d'une cyclopédie parfaitement intégrée au flux de travail médical de l'hôpital.
4. Approuvé par la FDA et CE (Europe).

II.3.5 Limites

1. Dédié à l'imagerie médicale d'images spécifiques uniquement (CTPA).
2. Ne convient pas aux classifications pulmonaires générales telles que les nodules ou l'inflammation.
3. Un environnement hospitalier numérique intégré sera nécessaire pour fonctionner efficacement.

II.4 VUNO Med – LungCT AI [5]

II.4.1 Demo Description

Une entreprise coréenne, VUNO, crée un processeur médical intelligent capable d'analyser les scanners pulmonaires à des fins de surveillance.

•Détection des nodules pulmonaires.

•Évaluer le potentiel malin.

•KeyCode Medical Support pour la détection précoce du cancer du poumon

II.4.2 Base Functions

1. Détection automatique des nodules pulmonaires par imagerie rapprochée.
2. Analyse approfondue de la taille, de la densité et de l'emplacement du dôme.
3. Forecast la probability that a belief will be detrimental using a prediction algorithm.

II.4.3 Méthodes utilisées

1. Neurones dorsaux (CNN).
2. Apprentissage automatique de données volumineuses par apprentissage annoté par des radiologues.
3. Modèles de prédiction du risque de cancer.

II.4.4 Advantages

1. Haute precision pour la détection de petits nodules jusqu'à < 6 mm.
2. Charge radiologique des radiologues réduite et l'gain dans le diagnostic.
3. Conviviale interface et bonne intégration dans l'environnement hospitalier.
4. Récupération de l'approbation du ministère coréen de la Santé et du Bien-être (MFDS) et de plusieurs approbations internationales.

II.4.5 Remise sécurité

1. Présence d'autres affections pulmonaires comme l'asthme ou l'obstruction en compte comme condition médicale.
2. Basé sur des images CT de haute résolution.
3. Certaines signes spécifiques peuvent être visualisés lorsque l'on a d'autres affections similaires à ceux-ci (confusion diagnostique).
4. Le système ne c'est pas open source et peut donc rendre difficile la vérification complète du mécanisme.

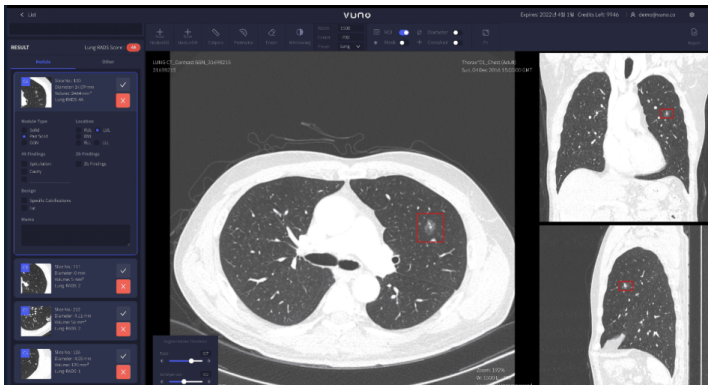

Fig. II.2 : VUNO Med – LungCT AI

II.5 Zebra médicine Vision [6]

II.5.1 Description générale

Il s'agit d'une plateforme d'IA qui traite automatiquement les images médicales (radiographies, tomodensitométries, IRM) et génère des rapports de diagnostic permettant aux médecins de détecter les maladies à un stade précoce.

Son domaine d'expertise : Elle cible les maladies chroniques et préoccupantes, telles que :

1. Les maladies pulmonaires (obstruction, fibrose, nodules)
2. L'ostéoporose.
3. Les maladies coronariennes.
4. Les troubles cérébraux.
5. Le cancer du sein.

Sa technologie est appelée « Moteur d'analyse d'imagerie ».

II.5.2 Fonctionnalités de base

1. ographies = Analyse automatique d'images radiologiques assistée par ordinateur grâce à l'intelligence artificielle.
2. ographies Fournit automatiquement au médecin des rapports de diagnostic détaillés en quelques minutes.
3. Test unique pour la détection multiple de certaines maladies, avec de multiples utilités.
4. Intégration au système de santé hospitalier (PACS/RIS).

II.5.3 Matériel et dispositifs utilisés

1. Des millions d'images pour l'entraînement en profondeur des modèles neuronaux.
2. Apprentissage automatisé des tissus et méthodes d'analyse quantitative.
3. La plupart des dispositifs et récepteurs médicaux sont compatibles.

II.5.4 Avantages

1. Durée d'action (par exemple, cholestérol).
2. La plupart des pays l'ont approuvé et les plus grands systèmes de santé le prescrivent.
3. Gain de temps et d'efforts pour le médecin, notamment dans un hôpital surchargé.
4. Possibilité de traiter de nouvelles maladies.

II.5.5 Inconvénients

1. Un système de cloud computing peut être mis en place dans n'importe quel pays pour des raisons de confidentialité.
2. Une utilisation non spécifique désactive la solution une fois, car les lésions pulmonaires sont affectées par l'arçon du programmeur, ce qui peut entraîner une précision parfois inférieure à celle d'autres solutions du problème pur comme SANADv1.
3. Certaines de ses activités ont un coût ou une adhésion institutionnelle.

II.6 qCT Poumon – Imbio[7]

II.6.1 Enquête et description

CT Lung est un logiciel médical qui utilise les technologies d'intelligence artificielle de Mehran pour lire automatiquement()) ; Scanner (TDM) des poumons à l'échelle quantitative. Il est principalement utilisé dans l'identification et l'évaluation des maladies pulmonaires chroniques, dont les plus importantes sont : • Fibrose pulmonaire • Maladie pulmonaire obstructive chronique (MPOC) • et autres()) ;

II.6.2 Fonctions principales

• Analyse pulmonaire quantitative : fournit des mesures spécifiques du volume pulmonaire, de la distribution de l'air et de l'étendue de la fibrose ou des dommages aux tissus pulmonaires.

• Thermique à code couleur : la zone affectée est divisée par la hauteur précise de sa descente pour une identification facile.

• Rapport automatisé complet : contient des données quantitatives pour aider les médecins à prendre des décisions.

• Comparaison dans le temps : elle peut être utilisée pour mesurer la progression de l'état d'un patient sur un certain nombre de scanners de contraste.

II.6.3 Technologie utilisée

1. Apprentissage profond.
2. Modélisation 3D du poumon.
3. Imagerie quantitative.

II.6.4 Caractéristiques

1. Haute précision dans la détection et l'estimation de l'étendue des lésions pulmonaires.
2. Fournit une évaluation objective qui est comparable entre les médecins et les centres.
3. Aider à surveiller l'efficacité à long terme des traitements.

II.6.5 Restrictions

1. Nécessite des images CT mobiles de haute qualité.
2. Non spécifique mais sans valeur pour les détailsNodules LTR ou cancer medha (ou autres exemples, SANADv1).
3. Intégration avec les systèmes hospitaliers (PACS).
4. La disponibilité des données antérieures des patients est utilisée pour biaiser le suivi.

II.7 Article 1 :Artificial intelligence for chest radiograph interpretation [8]

Quoi qu'il en soit, de nombreuses études ont exagéré l'émergence de l'intelligence artificielle pour sauver les maladies pulmonaires, la part imaginée à partir d'images médicales, notamment la radiologie et la tomodensitométrie. Ces approches diffèrent en termes de besoins en données, de méthodologies utilisées et d'objectifs de rapport de diagnostic.

Afin de positionner notre travail dans le contexte de recherche actuel, nous présentons dans le tableau suivant une comparaison complète des études pertinentes les plus importantes, montrant leurs méthodologies, leurs bases de données, leurs types de classification et la précision obtenue dans chaque cas.

II.7.1 Idée globale de l'article

Cette revue offre une vision d'ensemble intégrale de l'implémentation de l'intelligence artificielle (IA), et plus spécifiquement des réseaux de neurones convolutifs (RNC), à l'interprétation des radiographies thoraciques. Les auteurs dépassent les progrès étasuniens récents, les défis d'ordre technique et clinique et le potentiel futur de l'IA dans l'imagerie médicale pulmonaire.

II.7.2 Méthodologie

Il s'agit d'un article de «point de l'art» fondé sur une synthèse critique des publications disponibles. Il n'est pas ciblé sur aucun modèle expérimental mais que ce dernier synthétise les recherches effectuées, les études basées sur des bases de données comme NIH ChestX-ray14 et CheXpert. Il décrit les missions accomplies : classification, détection, segmentation, localisation.

II.7.3 Forces

Présentation stricte et claire des types d'algorithmes utilisés (RNC, RNC).

Exemple de cas cliniques réels des travailleurs où le radiologue peut être assisté par l'IA.

Échange pertinent en ce qui concerne l'éthique, l'interprétabilité et la validation clinique.

Contexte pertinent pour les chercheurs débutants ou les cliniciens intéressés par les technologies IA.

II.7.4 Faiblesses

Aucun modèle suggéré ni évalué directement.

L'article reste superficiel sans une comparaison quantitative entre les performances des approches citées.

Il s'agit de 2019, et il n'a donc pas pris en compte le développement de la pandémie COVID-19.

II.7.5 Pertinence et signification

Oui, c'est un article de base à prendre en charge la maîtrise des bases théoriques et du déploiement de la spécialité. Il présente l'opportunité de discipliner historiquement et techniquement l'apparition de l'IA dans le diagnostic pulmonaire, grâce aux défis à relever pour un usage clinique efficace.

II.8 Article 2 : Deep-learning framework to detect lung abnormality – A study with chest X-Ray and lung CT scan images[9]

II.8.1 Clé de la notion de l'article

Il n'y a pas dans l'article une approche hybride modèle apprentissage profond (CNN) et machine à vecteurs de support (SVM) pour la détection d'anomalies pulmonaires, au sens du cancer du poumon et de la pneumonie, à partir de drainage des radiographies et tomodensitogrammes.

II.8.2 Méthodologie

Les auteurs investiguent une meilleure architecture AlexNet (MAN) en remplaçant la couche Softmax finale par un classificateur SVM. Les auteurs investiguent des caractéristiques de main (Haralick, Hu Moments) et les investiguent en les combinant avec des caractéristiques profondes en utilisant un processus de PCA pour réduction de dimensionnalité. L'évaluation est conduite sur les bases de données NIH et LIDC-IDRI.

II.8.3 Bonnes choses

Nouvelle stratégie hybride qui est un mélange CNN et SVM.

Haute performance atteinte (jusqu'à 97,27 % sur des images CT).

Utilisation de plusieurs formes de caractéristiques (pro-fond + manuelle).

Preuve solide de complémentarité entre un apprentissage profond et des approches traditionnelles.

II.8.4 Limites

Complexité de calcul relativement élevée due au mélange de plusieurs formes de caractéristiques.

Le modèle n'est pas testé dans un scénario clinique réel.

Moins susceptible de faire référence pour la généralisation aux autres types de données.

II.8.5 Pertinence et importance

Oui, c'est un article très pratique à lire pour se faire une idée des modèles hybrides (deep learning + SVM). Il expose un schéma possible à suivre pour MODELING les systèmes de classification en les rendant solides et en atteignant une très bonne performance. Très intéressant pour un projet appliqué de diagnostic assisté.

II.9 Article 3 :CoroNet A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images[10]

:

II.9.1 Idée maîtresse de l'article

CoroNet est un réseau d'apprentissage profond Xception entraîné en réseau basé pour catégoriser spécifiquement les images radiographiques pulmonaires en trois classes : COVID-19, pneumonie et cas normaux.

II.9.2 Méthodologie

Le modèle Xception est affiné par apprentissage par transfert sur un jeu de données médicales COVIDx. Les performances sont utilisées en classification binaire et triclasse à l'aide de matrices de confusion et de métriques classiques (précision, rappel, etc.).

II.9.3 Points forts

Architecture éclaircie et performante, conçue pour être déploiée rapidement.

Haute précision (95 %) de classification triclasse.

Sujet d'adaptation rapide d'IA à une crise sanitaire publique (COVID-19).

Code et modèle public, voire jusqu'à la reproductibilité.

II.9.4 Faiblesses

Jeu de données déséquilibré (plus d'images « non COVID »).

Ne parle pas d'images CT, mais seulement de radiographie.

Détresse potentielle de sur-apprentissage à cause de la petite taille des données COVID-positives.

II.9.5 Pertinence et utilité

Oui, il s'agit d'une nouvelle plus qu'ancienne sur la pandémie. Elle démontre la demande dont les architectures profondes peuvent être envisagées pour répondre à un nouveau besoin médical. Lisible en tant que base pour toute projet que prend en charge le COVID-19 par l'intermédiaire de l'IA

II.10 Article 4 : COVID-19 detection using CNNs trained on segmented lung images[11]

II.10.1 Idea du projet principal de l'article

Les auteurs présentent une approche hybride appelée VDSNet en fusionnant VGG, STN (Spatial Transformer Networks) et CNN conventionnel pour détecter les maladies pulmonaires multi-pathologiques à partir de radiographies.

II.10.2 Méthodologie

La méthode traite les déformations dues aux rotations et aux perturbations bruit dans l'image médicale à l'aide de STN. Elle est suivie d'extraction de caractéristiques avec VGG et un CNN de fin pour la classification. Les performances sont testées sur NIH ChestX-ray.

II.10.3 Points forts

Bon traitement des images floues, mal orientées ou bruitées.

Modularité simple dans la réadaptation.

Utilisation multi-pathologies (pas réservée au COVID-19).

II.10.4 Faiblesses

Faible précision relativement (73 %),multiclasses à la base.

Pas de test sur données CT.

Manque d'analyse détaillée des erreurs de classification.

II.10.6 Pertinence et importance

Article moyennement important, surtout pour ceux qui sont intéressés à travailler avec des images défectueuses (réelles). Il prouve que plutot que la classification en soi, le prétraitement intelligent d'images est essentiel en pratique clinique avant même la classification.

II.11 Article 5 : Hybrid Deep Learning for Detecting Lung Diseases from Chest X-ray Images (VDSNet)[12]

II.11.1 Idée centrale de l'article

Les auteurs utilisent un modèle en classification à deux étapes : segmentation poumons par U-Net, classification COVID/Non-COVID par un CNN (VGG19) entraîné sur les patients de ces régions seulement.

II.11.2 Méthodologie

Après segmentation fine des poumons, les images.segmentées sont utilisées comme base d'entraînement à VGG19 en mode transfert learning. Les bases de données BIMCV-COVID19+, PadChest, Montgomery sont utilisées.

II.11.3 Fiertés

Stratégie de segmentation préalable extrêmement efficace.

Gestion de la généralisation sur plusieurs bases de données.

Architecture simple mais bien optimisée.

II.11.4 Faiblesses

N'abord que deux classes (COVID / Non-COVID).

Faible discussion sur les erreurs de faux positifs/faux négatifs.

La segmentation est longue en mode test.

II.11.5 Importance et pertinence

Oui, c'est un article hautement pertinent aux travaux qui visent à embrasser traitement d'image médical et classification. Il démontre qu'il est possible de grandement améliorer la performance d'un classificateur simplement en séparant les zones pertinentes.

II.12 Article 6 : Mémoire de Master : Détection du COVID-19 à l'aide des CNNs et SVM[13]

II.12.1 Idée principale de l'article

Ainsi, ce mémoire met en évidence une comparaison de 12 modèles CNN pré-entraînés, associés à un classificateur SVM, dans le but d'aider à dépister le COVID-19 en fonction d'images CT.

II.12.2 Méthodologie

Les auteurs s'apprennent des traits de réseau tels que ResNet18/50/101, VGG16/19, et les classifient à l'aide de SVM. Il est employé un total de deux bases de données (746 et 349 images). L'évaluation est reposée en majorité sur l'accuracy.

II.12.3 Points forts

Test de comparaison large vers des différents modèles CNN.

Meilleure performance réalisée avec ResNet101 + SVM : 97.85 %.

Méthodologie claire, bien documentée, centrée sur l'application.

II.12.4 Faiblesses

Le dataset est relativement petit.

Ne le couvre que pour le COVID-19, sans autres pathologies.

Fits dans un cadre académique (pas clinique).

II.12.5 Pertinence et importance

Oui, un mémoire de master de bonne qualité qui prouve la faisabilité d'une approche hybride CNN+SVM pour l'imagerie médicale CT. A lire absolument pour les étudiants et les chercheurs en phase expérimentale.

Tableau comparatif des travaux connexes

Tab. II.1 : Résumé comparatif des approches de classification d'images CT pulmonaires.

Travail / Auteur	**Méthode proposée**	**Base de données**	**Résultats obtenus**	**Commentaires / Limites**
Aidoc (2020)	CNN pour détection d'embolie pulmonaire	CTPA	Précision clinique élevée	Logiciel commercial, non open-source
InferRead CT (2021)	CT COVID-19 + réseau convolutif	CT thorax (hôpitaux)	$>90\%$ précision	Modèle propriétaire, peu de détails
Zhou et al. (2019)	3D CNN nodulaire	LIDC-IDRI	87.5% d'accuracy	Fort coût mémoire
SANADv1 (proposé)	ConvNeXt + RMSprop (local)	LIDC-IDRI (localisé)	91% précision validation	Généralisation locale réussie
Rahman et al. (2020)	CNN + augmentation	Kaggle CT Lung	89.2% précision	Déséquilibre des classes

II.13 Conclusion

La fouille scientifique approfondie de la littérature concernant la classification d'images médicales pulmonaires à l'aide d'algorithmes révèle une évolution extraordinaire de l'intelligence artificielle, et plus particulièrement de l'apprentissage profond, ces dernières années. Les réseaux de neurones convolutifs (CNN) ont démontré leur performance pour le recensement automatique d'anomalies comme les nodules pulmonaires, la pneumonie, ou la COVID-19 à partir d'images radiographiques (X-ray) et de tomodensitométrie (CT).

Les modèles hybrides qui combinent le CNN avec des classificateurs basés sur les classiques SVM ont aussi montré leur intérêt en tant qu'outils pour essayer de combler les limites des autres. Il reste néanmoins quelques difficultés, la qualité des données, les modèles interpretables, la généralisation clinique et la gestion des biais n'étant qu'en partie résolus.

Les remarques que nous allons en passer parcsiennent appeler la nécessité de l'appropriation réflexive des fondements algorithmiques de l'apprentissage automatique, contingent à la mise en place de modèles efficaces et robustes. Le prochain chapitre sera donc dédié à l'examen du Apprentissage automatique, de ses fondements de principes, de ses algorithmes classiques, ainsi que de son hybride avec L'approche de l'apprentissage profond, afin de mieux le situer dans un contexte scientifique prouvé.

Chapitre III

Apprentissage automatique

III.1 Introduction

À nos jours, les fondements de l'informatique sont éparpillés partout : santé, éducation, économie et cosmologie. Cela a enrichi considérablement la vie humaine ordinaire et ouvert son usage et son compréhension dans de nombreux domaine complexes. Une des industries les plus touchées par cette technologie est l'industrie de la santé, où les progrès technologiques et informatiques ont enrichi les équipements et les logiciels en santé, ainsi que la traçabilité des données de santé, ce qui a augmenté à cet effet la finition des diagnostics.

Des pionniers ont pris le regard des chercheurs dans cette branche, avec l'intelligence artificielle, les systèmes experts, l'apprentissage automatique et l'apprentissage profond.

Dans ce chapitre, nous aborderons plusieurs aspects spécifiques : d'abord les techniques d'apprentissage automatique, puis nous expliquerons comment ce domaine s'étend à l'apprentissage profond pour obtenir des structures plus efficaces, et enfin, nous aborderons quelques recherches appliquées.

III.2 Intelligence artificielle

III.2.1 Définition

L'intelligence artificielle, ou IA, est une discipline scientifique et technologique visant à doter les machines, telles que les ordinateurs et les logiciels, de la capacité d'exécuter des processus cognitifs habituellement associés au cerveau humain. Ces processus incluent la compréhension, la communication (tant entre les machines qu'avec les humains), l'adaptation et l'apprentissage autonome grâce à des techniques telles que le deep learning.**[14]** La figure III.1 suivante montre relation entre l'intelligence artificiell, l'apprentissage automatique et l'apprentissage profond :

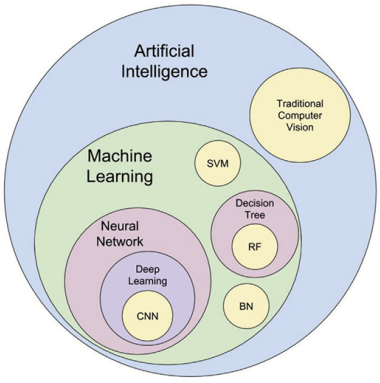

Fig. III.1 : relation entre l'intelligence artificiell, l'apprentissage automatique et l'apprentissage profond **[8]**

III.2.2 Les sous-domaines de l'IA

La figure III.2 suivante montre certains des domain :

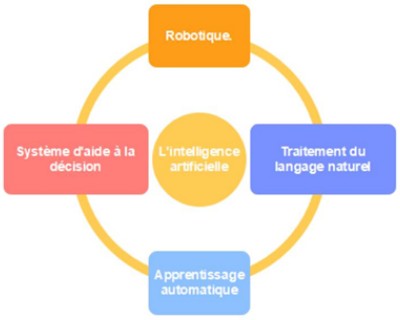

Fig. III.2 : Les sous-domaines de l'IA. **[14]**

III.2.3 Les avantages de AI

L'intelligence artificielle ale potentiel de remplacer les humains dans leurs tâches quotidiennes, permettant ainsi d'effectuer des travaux pénibles ou dangereux tels que le ménage, les courses, la cuisine, le jardinage, sans aucune contrainte physique comme le besoin de repos ou de nourriture .**[14]**.

Les algorithmes d'IA accélèrent les calculs sur les ordinateurs, réduisant les erreurs par rapport aux calculs humains.**[14]**.

Les véhicules autonomes, équipés de caméras et de capteurs, facilitent les déplacements en se déplaçant sans intervention humaine. **[14]**.

Dans le domaine médical, l'IA offre des avantages multiples, comme le suivi à distance des patients, les prothèses intelligentes et les traitements personnalisés. **[14]**.

III.2.4 Les inconvénients de AI

Une préoccupation majeure concerne la possibilité d'erreurs dans la programmation des robots, compromettant ainsi leur bon fonctionnement. Les machines telles que les ordinateurs, les robots et les véhicules intelligents ne peuvent pas détecter ces erreurs de programmation. Bien que le risque soit généralement faible, les conséquences d'une telle erreur pourraient être catastrophiques à grande échelle. **[14]**.

L'augmentation du chômage est un autre effet potentiel, car les entreprises pourraient opter pour le remplacement des travailleurs par des robots dotés d'intelligence artificielle. Ces robots ne se fatiguent pas et nécessitent seulement une maintenance occasionnelle, ce qui peut entraîner des suppressions d'emplois. **[14]**.

Les coûts élevés de recherche et développement dans le domaine de l'IA constituent également un défi. La création de robots autonomes capables de fonctionner dans la vie quotidienne serait extrêmement coûteuse.**[14]**.

III.3 Apprentissage automatique (Machine learning)

III.3.1 Définition

L'apprentissage automatique est l'un des domaines d'étude de l'intelligence artificielle apparus au milieu du XXe siècle, visant à développer des algorithmes capables de collecter des connaissances sans être explicitement programmés, c'est-à-dire, offrir à un système la capacité d'acquérir et d'intégrer indépendamment des connaissances [15]

III.3.2 Type de l'apprentisage automatique

L'apprentissage automatique est sujet à de faibles goulots d'étranglement dans un large éventail de tâches humaines et est particulièrement adapté aux problèmes de prise de décision automatisée.

En général, il existe trois types d'algorithmes d'apprentissage automatique :

1.Apprentissage supervise

Dans l'apprentissage supervisé, un humain accompagne l'algorithme dans son apprentissage, tandis qu'un data scientist agit comme mentor et lui enseigne les résultats qu'il doit obtenir. Il en va de même pour apprendre à un enfant à identifier des fruits et à les mémoriser. Dans l'apprentissage supervisé, l'algorithme apprend à partir d'un ensemble de données pré-étiquetées avec un résultat prédéfini.

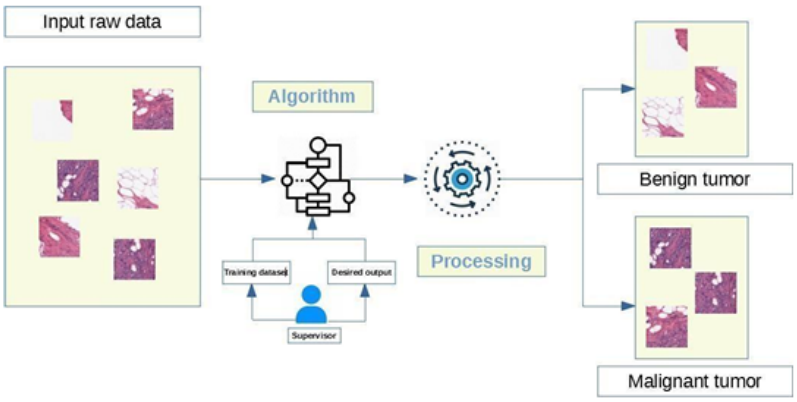

Fig. III.3 : Apprentissage supervisé

Les algorithmes d'apprentissage automatique supervisé sont les plus utilisés, et il existe deux types d'apprentissage supervisé ,la figure III.1 sépare deux types d'apprentissage supervisé :

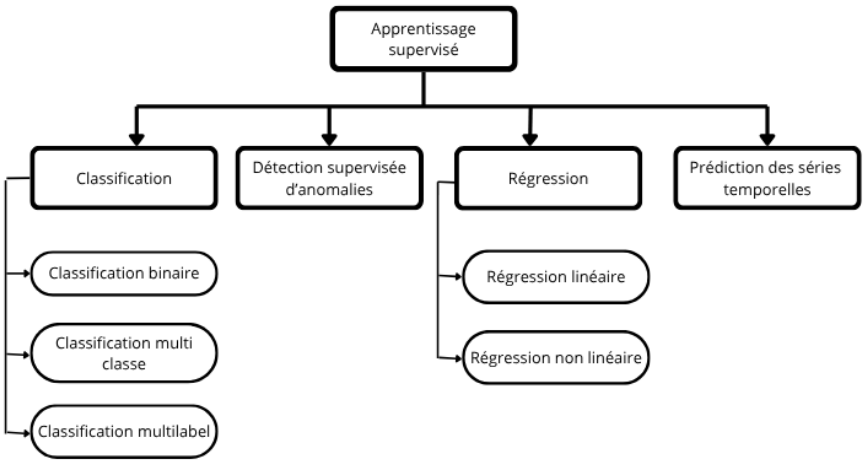

Fig. III.4 : Types d'apprentissage supervisé**[21]**

2.Apprentissage non supervisé

Dans l'apprentissage non supervisé, il n'y a pas d'étiquettes pour les données d'entraînement (train data). L'algorithme d'apprentissage automatique tente d'apprendre les caractéristiques ou les distributions de base contenue dans les données. **{16}**

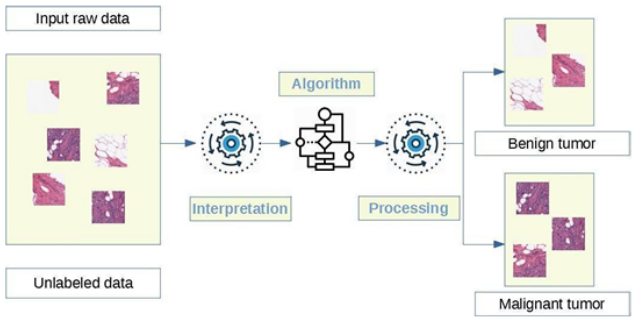

Fig. III.5 : Apprentissage non supervisé**[21]**

3. Apprentissage Semi-supervisé

Nous avons préalablement vue l'apprentissage supervisé et non supervisé, dont la majeur différence réside dans le fait que les données soient étiquetées ou non, et à cela s'ajoute les méthodes adéquates utilisées pour traiter ses données. L'apprentissage semi-supervisé regroupe ses deux principes, il prend un ensemble réduit de données étiquetées avec un autre ensemble de données non étiquetées du même types.

L'avantage de ce type d'apprentissage réside principalement dans le processus d'étiquetage des données prend beaucoup de temps et souvent coûteux. Donc paradoxalement le non étiquetage devient bénéfique pour le processus d'apprentissage, et la construction du modèle et moins coûteuse.**[17]**

4. Apprentissage par renforcement

Dans l'apprentissage par renforcement, l'algorithme détermine les actions à entreprendre dans une situation donnée pour maximiser une récompense (sous la forme d'un nombre) en vue d'atteindre un objectif spécifique. **{16}**

III.3.3 Domaine de l'apprentissage automatique

L'apprentissage automatique est une branche de l'intelligence artificielle (IA) et de l'informatique qui se concentre sur l'utilisation de données et d'algorithmes pour simuler la façon dont les humains apprennent et améliorer progressivement sa précision .L'apprentissage automatique est une composante importante du domaine en plein essor de la science des données. Grâce à l'utilisation de méthodes statistiques, les algorithmes sont formés pour effectuer des classifications ou des prédictions, et pour révéler des informations clés sur les projets d'exploration de données. Ces informations orientent ensuite les décisions au sein des applications et des entreprises, affectant idéalement les mesures de croissance clés. Alors que le Big Data continue de se développer et de se développer.**[17]**

III.3.4 La Classification dans l'apprentissage automatique

La classification des catégories est l'une de ces méthodes. Classer les éléments de l'apprentissage automatique, souvent utilisés par les experts. Les données sont dues, pour n'en citer que quelques-unes, à la contribution de l'appareil. Classer les éléments en fonction des cellules. La classification est utilisée lorsque la valeur de la cible est utilisée. Prédire a une valeur discrète. Les méthodes les plus couramment utilisées sont :

A.SVM (Support Vecteur Machine) :

La machine à vecteurs de support (un algorithme d'apprentissage automatique supervisé) est une technique qui a été introduite par Vladimir Vapnik en 1995. Elle est conçue pour résoudre les problèmes de classification et de régression. Elle dépend principalement de l'utilisation de deux concepts : le concept de la marge maximale et le concept de la fonction du noyau (facilite le processus de séparation). La technique SVM peut être utilisée pour résoudre divers problèmes en bio-informatique, recherche d'information et vision par ordinateur, etc.**[15]**

Avantages de la SVM :

- La capacité à traiter grandes masses de données.
- Le faible nombre d'hyper paramètres utilisés par cette méthode.**[15]**
- Elle est théoriquement très étayée
- Les résultats pertinences que l'on peut obtenir en pratique.
-

Inconvénients de la SVM :

- L'utilisation des fonctions mathématiques complexes pour la classification des corpus.**[15]**
- Pour trouver les meilleurs paramètres, ce type d'algorithme demande un temps énorme pendant les phases de test.**[15]**

B. KNN (k-Nearest Neighbors) :

L'algorithme des k plus proches voisins (KNN) est utilisé généralement dans les problèmes informatiques incluant la reconnaissance de formes, la recherche dans les données multimédia, la compression vectorielle, les statistiques informatiques et l'extraction des données.

La technique KNN diffère essentiellement des autres méthodes par sa simplicité et par le fait qu'aucun modèle n'est introduit à partir des exemples pendant le processus de classification.

Pour trouver la classe d'un nouveau cas, cet algorithme se base sur le principe suivant : il cherche les k plus proches voisins de ce nouveau cas, ensuite, il choisit parmi les candidats trouvés le résultat le plus proche et le plus fréquent.

Cette méthode utilise principalement deux paramètres : le nombre k et une fonction de similarité, qui permet de comparer les cas déjà classés avec le nouveau cas. **[15]**

Avantages du k plus proches voisins :[15]

- Facile à concevoir.
- La mise en œuvre facilité de cet algorithme.
- Son efficacité pour des classes réparties de manière irrégulière.
- Son efficacité pour des données incomplètes.
- La méthode des k plus proches voisins ne nécessite pas d'un modèle pour classer le documents.

Inconvénients :[15]

- Le principal inconvénient de KNN est le temps d'exécution qu'elle met pour la.
- La nécessité de grande capacité de stockage.
- Sensible aux perturbations (bruit).
- Pour un nombre de variable prédictives très grands, le calcul de la distance devient très coûteux.

III.4 Apprentissage profond

III.4.1 Définition

Le Apprentissage profond est l'un des principaux éléments de la science des données. C'est un sous domaine de l'intelligence artificielle (IA) et de l'apprentissage automatique (Machine learning) qui donne aux ordinateurs des capacités leur permettant de comprendre le monde en imitant la façon dont les humains analysent, en utilisant des algorithmes simulant le cerveau avec ou sans supervision humaine.**[2]**

III.4.2 Application d'Apprentissage profond

Les applications du Apprentissage profond sont diverses et touchent de nombreux secteurs.La figure III.6 suivante montre quelques applications :

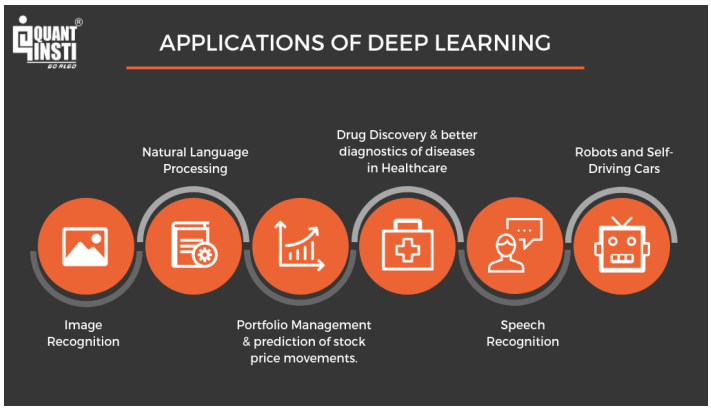

Fig. III.6 : Application d'Apprentissage profond **[18]**

III.5 Réseaux de neurones convolutifs (CNN)

III.5.1 Définition

Un réseau neuronal convolutif est un type particulier de réseau neuronal basé sur processus de torsion .Les réseaux convolutifs sont dérivés des architectures de type Perceptron multicouche (Perceptron multicouche : MLP), mais ils utilisent des poids Commun, lié à la fenêtre de torsion, leur permettant d'extraire implicitement de Caractéristiques locales Les CNN sont particulièrement adaptés reconnaissance d'images .On dit qu'un réseau convolutif reçoit chaque neurone L'information ne provient pas de toute la couche précédente, mais seuls les neurones situés dans leur champ récepteur .**[17]**

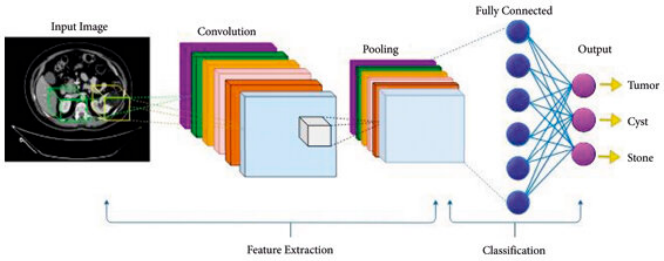

Fig. III.7 : Architecture d'un CNN traditionnel **[19]**

III.5.2 Quelques architectures de réseaux neuronaux convolutifs

A.Lenet-5

Lenet-5 est l'un des premiers modèles pré-entraînés proposés par Yann LeCun et d'autres dans l'année 1998, dans le document de recherche Gradient-Based Learning Applied to Document Recognition. Ils ont utilisé cette architecture pour reconnaître les caractères manuscrits et les caractères imprimés à la machine.

La principale raison de la popularité de ce modèle est son architecture simple et directe. Il s'agit d'un réseau neuronal de convolution multicouche pour la classification d'images.

Il comporte trois ensembles de couches de convolution avec une combinaison de mise en commun moyenne. Après les couches de convolution et de mise en commun moyenne, il y a deux couches entièrement connectées. Enfin, une activation Softmax qui classe les images dans leur classe respective.**[16]**

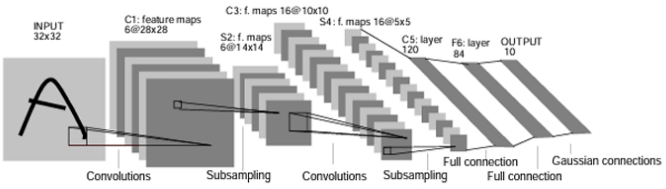

Fig. III.8 : Archetecture de Lenet **[16]**

B.VGGNet

VGG est l'abréviation de Visual Geometry Group (groupe de géométrie visuelle) ; il s'agit d'une architecture de réseau neuronal convolutif (CNN) profonde, le VGG-16 ou le VGG 19 comprenant 16 et 19 couches convolutives. L'architecture VGG est à la base de modèles de reconnaissance d'objets révolutionnaires. Développé en tant que réseau neuronal profond, le VGGNet surpasse également les bases de référence dans de nombreuses tâches et ensembles de données au-delà d'ImageNet. En outre, il reste aujourd'hui l'une des architectures de reconnaissance d'images les plus populaires. Le réseau VGG est construit avec de très pe-

tits filtres convolutifs. Le VGG-16 se compose de 13 couches convolutionnelles et de trois couches entièrement connectées (FC) : **[16]**

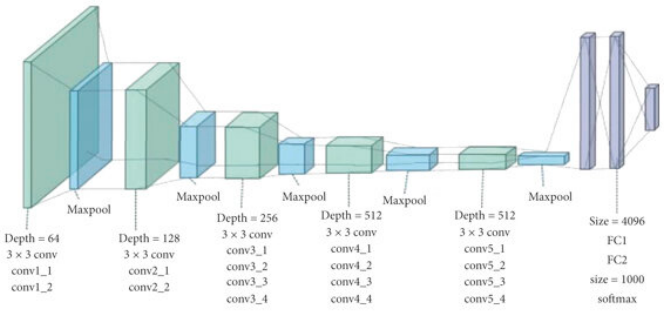

Fig. III.9 : L'architecture de VGG19 **[19]**

C.ResNet

Le défi ILSVRC 2015 (classification task) a été remporté à l'aide d'un réseau résiduel (ou ResNet), développé par Kaiming He et al., qui a obtenu un taux d'erreur étonnant inférieur à 3,6 % dans le top 5, en utilisant un CNN extrêmement profond composé de 152 couches. Cela a confirmé la tendance générale : les modèles sont de plus en plus profonds, avec de moins en moins de paramètres. La clé pour pouvoir former un réseau aussi profond est d'utiliser des skip connections (aussi appelées connexions de raccourci) : le signal alimentant une couche est également ajouté à la sortie d'une couche située un peu plus haut dans la pile. Voyons pourquoi cela est utile.

Lors de la formation d'un réseau neuronal, l'objectif est de lui faire modéliser une fonction cible h(x).

Si vous ajoutez l'entrée x à la sortie du réseau (c'est-à-dire si vous ajoutez une skip connection), le réseau sera forcé de modéliser f(x) = h(x) - x plutôt que h(x). C'est ce qu'on appelle l'apprentissage résiduel . Le réseau ResNet utilise une architecture de réseau simple à 34 couches inspirée de VGG 19 dans laquelle lesskip connections sont ensuite ajoutées :**[16]**

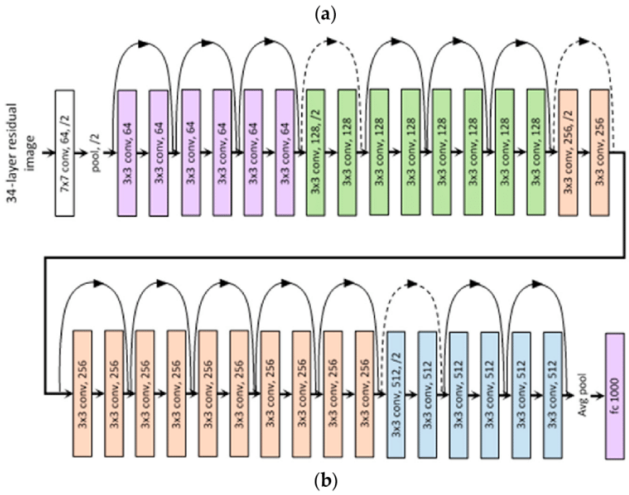

Fig. III.10 : L'architecture de ResNet **[20]**

D.MobileNetV3

Publié en 2019, améliore encore la version V2 en utilisant des techniques de search space pour déterminer automatiquement les meilleures configurations de l'architecture (via un processus appelé NAS pour Neural Architecture Search).

MobileNetV3 est conçu pour être encore plus performant tout en restant très compact, avec une meilleure efficacité pour des tâches spécifiques comme la classification d'images et la détection d'objets.

MobileNetV3-Large : Une version optimisée pour les applications avec des exigences de haute performance.

• MobileNetV3-Small : Une version encore plus petite et légère, adaptée aux appareils avec des ressources très limitées.**[22]**

E.ConvNext V1

Ces dernières années, les réseaux de neurones convolutifs (CNN) ont été utilisés dans la plupart des tâches de vision par ordinateur, notamment la classification, la reconnaissance d'objets et la segmentation. Cependant, malgré les performances spectaculaires de ces modèles, des améliorations étaient toujours possibles.

ConvNeXt fut conçu en réponse à l'ensemble des problèmes auxquels les modèles classiques faisaient face alors que lors de l'apprentissage c'était profond, soit un accroissement des données et des ressources informatiques.

ConvNeXt essaye d'améliorer les performances grâce à un recul du temps de formation, sans compromis aucun sur les résultats dans différentes tâches. ConvNeXt est une optimisation de l'architecture réseau convolutif traditionnel qui reposait principalement sur des couches convolutives en profondeur et s'en faisait de plus en plus capable à mesure qu'elle renforçait la profondeur.

ConvNeXt toutefois a une nouvelle formule qui intègre les dernières méthodes d'apprentissage en profondeur telle que l'optimisation de la distribution des paramètres et l'échelle du gradient.

III.5.3 Optimisation

A.Méthode du gradient stochastique (SGD)

La méthode du gradient stochastique (Stochastic Gradient Descent ou SGD) et ses variantes sont probablement les algorithmes d'optimisation les plus couramment utilisés pour l'apprentissage automatique général, et en particulier pour l'apprentissage profond.

Il est possible d'obtenir une évaluation objective du gradient en prenant un gradient moyen sur un mini-lot d'échantillons prélevés dans la distribution génératrice de données. **[2]**

L'algorithme SGD

montre comment suivre cette estimation de chute de gradient. Le paramètre clé de l'algorithme SGD est la vitesse d'apprentissage.

Le SGD utilise une vitesse d'apprentissage constante, mais dans la pratique, il est nécessaire de réduire progressivement la vitesse d'apprentissage au fil du temps. À titre de comparaison, le véritable gradient de la fonction de coût total devient petit, puis 0 lorsque nous approchons et atteignons le minimum à l'aide d'un gradient par lots. **[2]**

Optimiseur SGD

L'algorithme de base du *Stochastic Gradient Descent* met à jour les poids selon la règle :

$$w_{t+1} = w_t - \eta \cdot \nabla L(w_t)$$

où :

- w_t : poids à l'instant t,
- η : taux d'apprentissage (learning rate),
- $\nabla L(w_t)$: le gradient de la fonction de perte.

B.Méthode RMSprop

C'est une méthode d'optimisation développée par le professeur Geoffrey Hinton dans sa classe de filets neuronaux. Au lieu de laisser tous les gradients s'accumuler pour le Momentum, il n'accumule les gradients que dans une fenêtre fixe. RMSprop tente également d'amortir les oscillations, mais d'une manière différente de l'impulsion. RMSprop évite également de devoir ajuster le rythme d'apprentissage, et le fait automatiquement.

De plus, RMSprop choisit un taux d'apprentissage différent pour chaque paramètre. Dans RMSprop, chaque mise à jour est effectuée selon des équations précises. Cette mise à jour est effectuée séparément pour chaque paramètre. **[2]**

Optimiseur RMSprop

RMSprop adapte le taux d'apprentissage pour chaque poids individuellement :

$$E[g^2]_t = \gamma E[g^2]_{t-1} + (1-\gamma) g_t^2$$
$$w_{t+1} = w_t - \frac{\eta}{\sqrt{E[g^2]_t + \epsilon}} \cdot g_t$$

avec :

- γ : facteur de décroissance (souvent 0.9),

- ϵ : petit terme pour la stabilité numérique,
- g_t : le gradient à l'instant t.

C.Méthode Adam

Adam est l'abréviation de "Adaptive Moment Estimation", c'est la méthode d'optimisation la plus utilisée dans le deep learning. Elle propose une autre façon d'utiliser les gradients passés pour calculer les gradients actuels. Adam utilise également le concept de moment adaptatif en ajoutant des fractions de gradients précédents au gradient actuel. Cette méthode d'optimisation est devenue assez répandue, et est pratiquement acceptée pour être utilisée dans l'entraînement des réseaux neuronaux. Il est facile de se perdre dans la complexité de certains de ces nouveaux optimiseurs. Il suffit de se rappeler qu'ils ont tous le même objectif :

minimiser la fonction de perte.

Adam est généralement considéré comme assez robuste dans la sélection d'hyper paramètres, bien que la vitesse d'apprentissage doive parfois être modifiée par rapport à la valeur par défaut suggéré.**[2]**

Optimiseur Adam

Adam combine les avantages de RMSprop et de la méthode momentum :

$$
\begin{aligned}
m_t &= \beta_1 m_{t-1} + (1-\beta_1) g_t \\
v_t &= \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \\
\hat{m}_t &= \frac{m_t}{1-\beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1-\beta_2^t} \\
w_{t+1} &= w_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}
\end{aligned}
$$

où :

- β_1, β_2 : coefficients de lissage (typiquement 0.9 et 0.999),
- η : taux d'apprentissage,
- m_t, v_t : moyennes mobiles du gradient et de son carré.

III.6 Conclusion

Dans ce chapitre, nous avons présenté les fondements du Machine Learning ainsi que ses différentes catégories, notamment l'apprentissage supervisé, non supervisé et par renforcement. Une attention particulière a été portée à l'apprentissage supervisé, étant donné sa pertinence dans les tâches de classification d'images médicales.

Nous avons également exploré plusieurs algorithmes d'apprentissage, allant des méthodes classiques comme la régression logistique et les machines à vecteurs de support (SVM), jusqu'aux approches plus avancées telles que les réseaux de neurones et les architectures convolutives (CNN). Ces dernières se sont révélées particulièrement puissantes pour le traitement d'images complexes comme celles issues des scanners thoraciques.

L'étude des modèles CNN modernes comme ResNet, MobileNet et ConvNeXt a mis en lumière les évolutions récentes visant à améliorer la précision, la généralisation et l'efficacité computationnelle des modèles. L'intégration d'optimiseurs performants comme Adam ou SGD joue également un rôle crucial dans la convergence et la stabilité des réseaux.

Ce chapitre constitue ainsi la base théorique sur laquelle reposent nos choix méthodologiques et expérimentaux développés dans le chapitre suivant, consacré à l'implémentation concrète des modèles sur notre base de données médicale.

Chapitre IV

Implémentation

IV.1 Introduction

Après avoir présenté la théorie de l'apprentissage profond dans les chapitres précédents, ce chapitre sera consacré à la mise en œuvre d'un système basé sur l'apprentissage profond pour la détection de la rétinopathie diabétique. Pour cela, nous allons utiliser une base d'images de fond d'œil contenant deux classes, la première classe présente un diabète positif et l'autre négatif. Nous allons aussi présenter trois modèles CNNs développés pour la classification de cette maladie. Les résultats des tests des trois architectures seront aussi présentés et détaillés dans ce chapitre.

IV.2 Environnement du travail

Dans cette section nous présenterons le matériel et le logiciel utilisés dans notre travail.

IV.2.1 Environnement matériel

Afin de mettre en œuvre ce projet, nous avons utilisé un ensemble de matériel dont les caractéristiques sont les suivantes :

1. Un ordinateur portable Lenovo ThinkPad
2. Processeur : Intel(R) Core(TM) i5-6300U CPU @ 2.40GHz 2.50 GHz
3. Mémoire (RAM) : Taille 8Go
4. Type de système : système d'exploitation 64bits.
5. OS : Microsoft Windows 10

IV.2.2 Langage de programmation Python :

Python a été développé à l'Institut néerlandais de mathématiques et d'informatique (CWI) à Amsterdam par Guido van Rossum à la fin des années 1980, et sa première annonce remonte à 1991. Le noyau du langage est écrit en langage C. Ross a appelé son langage "Python" pour exprimer son admiration pour un célèbre groupe de sketchs britannique qui s'appelait Monty Python.[21]

Python est un langage de programmation de haut niveau conçu pour être facile à lire et à implémenter. Il est open source, ce qui signifie qu'il est gratuit, même pour des applications commerciales. Python peut fonctionner sur les systèmes Mac, Windows et Unix.**[23]**

Fig. IV.1 : Logo Python

IV.2.3 Google Colab :

L'entraînement d'un modèle d'apprentissage profond peut nécessiter une charge de travail importante au niveau du CPU/GPU, c'est pourquoi nous avons utilisé la plateforme en cloud de Google Colab.

Colaboratory est un projet de recherche de Google créé pour aider à généraliser l'enseignement et la recherche en apprentissage automatique. Il s'agit d'un environnement Jupyter notebook qui ne nécessite aucune configuration pour être utilisé et qui fonctionne entièrement sur le cloud.**[16]**

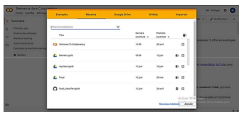

Fig. IV.2 : Interface de google Colab

IV.2.4 Visual Studio

Visual Studio Code, communément appelé VS Code, est un environnement de développement intégré développé par Microsoft pour Windows, Linux, macOS et les navigateurs web. Ses fonctionnalités incluent la prise en charge du débogage, la coloration syntaxique, la saisie semi-automatique intelligente, les extraits de code, la refactorisation de code et le contrôle de version intégré avec Git. Les utilisateurs peuvent modifier le thème, les raccourcis clavier et les préférences, ainsi qu'installer des extensions pour ajouter des fonctionnalités.

Visual Studio Code est un logiciel propriétaire publié sous la « licence logicielle Microsoft », mais basé sur le programme sous licence du MIT appelé « Visual Studio Code – Open Source » (également appelé « Code – OSS »), également créé par Microsoft et disponible sur GitHub.

Lors de l'enquête Stack Overflow 2024 auprès des développeurs, sur 58,121 réponses, 73,6 % des répondants ont déclaré utiliser Visual Studio Code, soit plus du double du pourcentage de répondants ayant déclaré utiliser son alternative la plus proche, Visual Studio .**[24]**

IV.3 Bibliothèques utilisées

IV.3.1 Keras

Keras est une API de réseaux de neurones de haut niveau, écrite en Python et capable de fonctionner sur TensorFlow ou Theano. Il a été développé en mettant l'accent sur l'expérimentation rapide. Être capable d'aller de l'idée à un résultat avec le moins de délai possible est la clé pour faire de bonnes recherches. Il a été développé dans le cadre de l'effort de recherche du projet ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System), et son principal auteur et mainteneur est François Chollet, un ingénieur Google. En 2017, l'équipe TensorFlow de Google a décidé de soutenir Keras dans la bibliothèque principale de TensorFlow. Chollet a expliqué que Keras a été conçue comme une interface plutôt que comme un cadre d'apprentissage end to end. Il présente un ensemble d'abstractions de niveau supérieur et plus intuitif qui facilitent la configuration des réseaux neuronaux

indépendamment de la bibliothèque informatique de backend. Microsoft travaille également à ajouter un backend CNTK à Keras aussi.**[16]**

Fig. IV.3 : Logo Keras

IV.3.2 TensorFlow

TensorFlow est une plate-forme open source de bout en bout pour la création d'applications d'apprentissage automatique. Il s'agit d'une bibliothèque mathématique symbolique qui utilise un flux de données et une programmation différentiable pour effectuer diverses tâches axées sur la formation

et l'inférence de réseaux de neurones profonds. Il permet aux développeurs de créer des applications d'apprentissage automatique à l'aide de divers outils, bibliothèques et ressources communautaires.**[23]**

Actuellement, la bibliothèque d'apprentissage en profondeur la plus célèbre au monde est TensorFlow de Google.

Google utilise l'apprentissage automatique dans tous ses produits pour l'optimisation des moteurs de recherche, la traduction, les légendes d'images ou les recommandations .**23**

Fig. IV.4 : Logo TensorFlow.

IV.3.3. OpenCV

OpenCV (Open Source Computer Vision Library) est une bibliothèque logicielle open source dédiée à la vision par ordinateur et à l'apprentissage automatique.

OpenCV a été créée pour fournir une infrastructure commune aux applications de vision par ordinateur et pour accélérer l'utilisation de la perception machine dans les produits commerciaux. Étant sous licence BSD, OpenCV permet aux entreprises d'utiliser et de modifier facilement le code.

La bibliothèque contient plus de 2500 algorithmes optimisés, comprenant un ensemble complet d'algorithmes de vision par ordinateur et d'apprentissage automatique, allant des classiques aux plus avancés. Ces algorithmes peuvent être utilisés pour détecter et reconnaître des visages, identifier des objets, classifier des actions humaines dans des vidéos, suivre les mouvements de caméra, suivre des objets en mouvement, extraire des modèles 3D d'objets, produire des nuages de points 3D à partir de caméras stéréo, assembler des images pour produire une image haute résolution d'une scène entière, trouver des images similaires dans une base de données d'images, supprimer les yeux rouges des photos prises avec un flash,

suivre les mouvements des yeux, reconnaître des paysages et établir des marqueurs pour les superposer avec la réalité augmentée, etc.

OpenCV compte plus de 47 000 utilisateurs dans sa communauté et un nombre estimé de téléchargements dépassant les 18 millions. La bibliothèque est largement utilisée par des entreprises, des groupes de recherche et des organismes gouvernementaux.**[14]**

IV.3.4 Numpy

Numpy est une bibliothèque pour effectuer des opérations arithmétiques numériques en Python. Le package de base autour duquel se construit la pile de calcul scientifique est appelé numpy (Numerical PYthon). La bibliothèque Numpy fournit une gestion plus facile des tables de nombres et des fonctions complexes (propagation).

Elle fournit également une abondance de fonctionnalités utiles pour les opérations sur les n-tableaux et les tableaux en python, en plus des conseils fournis pour les opérations mathématiques de type de tableau numpy qui améliore les performances et accélère ainsi l'exécution .**[15]**

Fig. IV.5 : Logo Numpy.

IV.3.5 Pandas

Pandas est une bibliothèque open source de manipulation et d'analyse de données pour le langage de programmation Python. Elle offre des structures de données flexibles et expressives, principalement les DataFrames et les Series, qui facilitent la manipulation de données tabulaires et de séries temporelles.

Pandas est particulièrement utile pour nettoyer, transformer et analyser des données, grâce à ses fonctionnalités puissantes pour la sélection, le filtrage, l'agrégation et la visualisation des données.

IV.4 Base de données utilisée

Nous avons utilisé la base de données LIDC-IDRI (Lung Image Database Consortium and Image Database Resource Initiative), une ressource de premier plan dans le domaine de l'analyse et du diagnostic assistés par ordinateur (CAO) des maladies pulmonaires.

Cette base de données est réalisée dans l'esprit de soutien à la recherche en détection, segmentation, et classification des nodules pulmonaires à partir des images de tomodensitométrie, ou échographie radiologique (scan CT). C'est donc la concrétisation du travail d'entente

d'avec diverses institutions médicales et de recherche dont l'objectif est d'aider la communauté scientifique à bénéficier d'un ensemble de données riches, annotées et accessibles.

La base LIDC-IDRI est constituée de plus de 1000 tomodensitogrammes thoraciques anonymisés correctement annotés par quatre radiologues. Ils sont tous annotés avec différents niveaux de confiance diagnostique, c'est-à-dire la présence, la taille, la forme et le type de nodules pulmonaires. Cela offre la flexibilité nécessaire pour entraîner et tester les modèles d'IA dans des conditions réelles et difficiles.

Grâce à sa qualité et à sa richesse, la base du LIDC-IDRI constitue une ressource essentielle pour la construction et l'évaluation de systèmes intelligents basés sur l'expertise pour le diagnostic du cancer du poumon, présentant un grand intérêt pour l'apprentissage assisté par ordinateur et l'apprentissage profond.

IV.5 Organisation du Jeu de Données et Augmentation

Avant l'étape d'augmentation, notre jeu de données était structuré comme suit :

- **600** images originales de fond (*background*) + **600** images de nodules pour l'**entraînement (train)**.
- **112** backgrounds + **112** nodules pour l'**évaluation (validation/test)**.

Ces images sont issues d'un sous-ensemble sélectionné à partir de la base LIDC-IDRI, converties en format PNG, recadrées localement et redimensionnées à 256×256 pixels.

Après l'augmentation des données

Nous avons appliqué une augmentation uniquement sur les **600 images d'entraînement de type "background"**, afin d'équilibrer le jeu de données et d'enrichir la représentation des tissus non nodulaires.

L'augmentation inclut plusieurs transformations aléatoires :

- Rotations jusqu'à $\pm 20°$
- Zooms jusqu'à 20%
- Translations horizontales et verticales (10%)
- Flip horizontal
- Ajustement de la luminosité (entre 80% et 120%)

Grâce à ces techniques, nous avons généré **1 545 images supplémentaires**, portant le total de backgrounds d'entraînement à **2 145 images**.

Cette stratégie a permis de réduire l'effet du déséquilibre entre les classes, tout en améliorant la capacité de généralisation de nos modèles. Le processus a été réalisé automatiquement via un script Python (Keras + Colab), et les images générées ont été sauvegardées dans Google Drive.

Tab. IV.1 : Distribution des données avant et après augmentation

Type de Données	Catégorie	Avant Augmentation	Après Augmentation
Entraînement	Nodulaires	600	600
	Non Nodulaires	600	**2 145**
Test / Validation	Nodulaires	112	112
	Non Nodulaires	112	112

IV.6 Description de l'application SANADv1

L'application SANADv1 permet de simuler un diagnostic pulmonaire local via une interface graphique développée avec Streamlit. Elle autorise le choix du modèle CNN (ResNet, MobileNet, ConvNeXt) et de l'optimiseur. Elle reçoit une image PNG CT en entrée et retourne la prédiction Nodulaire et Non Nodulaire)..

Architecture Générale de l'Application

Le fonctionnement global de notre application SANADv1 repose sur une suite de traitements allant de l'importation de l'image médicale jusqu'à l'affichage du résultat. Le schéma suivant illustre les différentes étapes du pipeline :

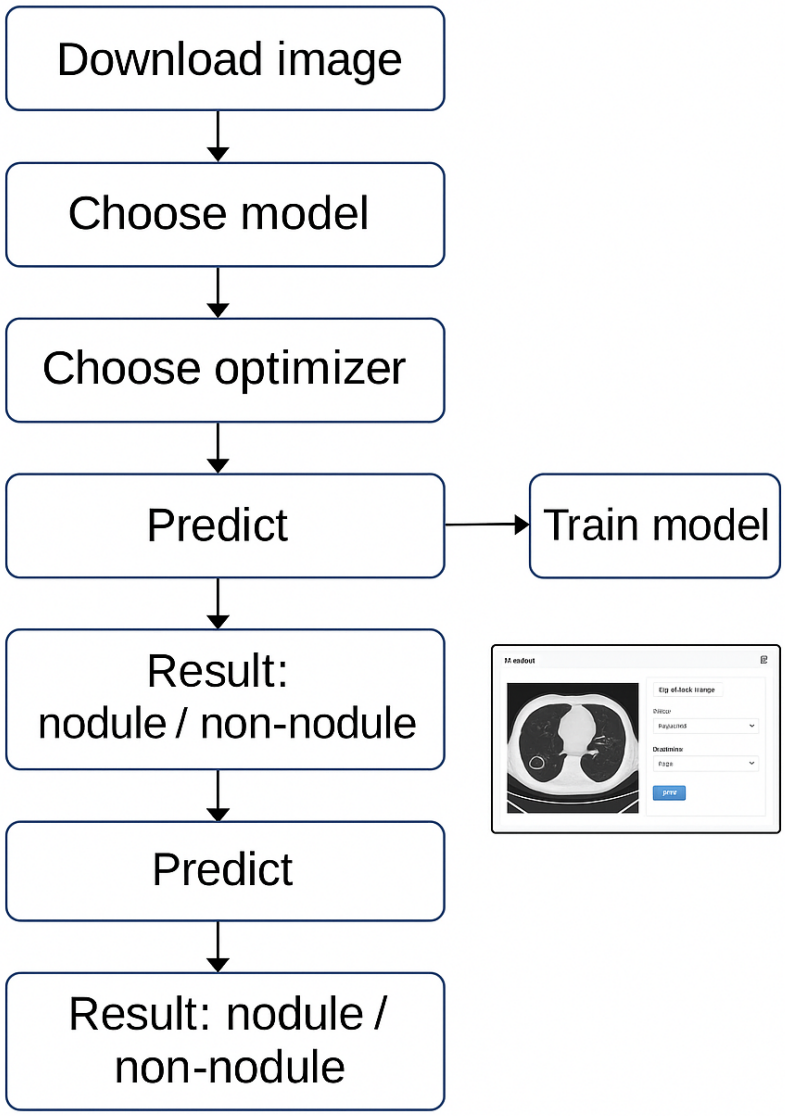

Fig. IV.6 : Architecture du pipeline de traitement dans SANADv1

Comme illustré ci-dessus, l'image CT est tout d'abord chargée puis prétraitée (redimensionnement, normalisation, suppression du bruit). Ensuite, des caractéristiques sont extraites pour alimenter les modèles de classification basés sur des architectures CNN (ConvNeXt, ResNet50, MobileNetV3), avec la possibilité de choisir entre plusieurs optimiseurs. Le résultat final est affiché dans l'interface, indiquant si l'image est nodulaire ou non nodulaire.

IV.7 Modèles utilisés

Ces modèles ont été sélectionnés en intégrant des architectures CNN bien établies dans la littérature, qui ont grandement bénéficié aux applications de diagnostic médical assisté par ordinateur. Nous avons testé différents modèles de classification d'images basés sur l'apprentissage profond.

Ceux-ci sont :

ResNet50, dédié à sa grande profondeur et à une stabilité due aux connexions résiduelles.

MobileNetV3, conçu pour fonctionner sur du matériel avec des ressources restreintes.

ConvNeXt, une nouvelle architecture de pointe motivée à la fois par les CNN classiques et par les Transformers.

Tous ces modèles ont été pré-entraînés sur ImageNet puis affiné (adaptés) sur notre ensemble de données pour exécuter la classification binaire entre les images nodulaire ou non nodulaire..

Nous avons entraîné chaque modèle indépendamment avec un ensemble d'optimiseurs, ce qui nous a permis d'interpréter et d'analyser le rôle de chaque optimiseur sur le modèle.

L'objectif était de déterminer la meilleure configuration possible pour une future intégration à notre système de diagnostic automatisé.

IV.7 Entraînement des modèles avec optimisateurs

Afin d'évaluer les performances des modèles de classification, nous avons adopté une méthodologie systématique d'entraînement qui est de tester chaque modèle avec des optimiseurs différents pour étudier leur influence sur la précision finale.

Pour chaque modèle (ResNet50, MobileNetV3 et ConvNeXt), nous avons effectué un entraînement distinct avec chacun des trois optimiseurs : Adam, RMSProp et SGD.

Nous avons effectué l'entraînement en conservant les mêmes paramètres de base, tels que le nombre d'époques, la taille des lots et la fonction de perte, afin de garantir une comparaison équitable entre les configurations.

L'ensemble de données a été divisé en ensembles d'entraînement et en ensembles de validation selon un ratio fixe .

Pendant l'entraînement, j'ai surveillé la courbe d'apprentissage (perte et précision) afin d'évaluer la stabilité et la vitesse de convergence.

Après l'entraînement, les modèles ont été évalués sur l'ensemble de validation afin de calculer la principale métrique : la précision.

Cette méthodologie nous a permis de déterminer quelle combinaison de modèle et d'optimiseur était la plus performante pour notre tâche spécifique de classification d'images pulmonaires nodulaire.

IV.8 Résultats expérimentaux

Résultats du modèle ConvNeXt avec différents optimiseurs

Afin de déterminer l'effet de plusieurs optimiseurs sur les performances du modèle ConvNeXt, nous l'avons imprégné de trois méthodes d'optimisation : Adam, RMSProp et SGD.

L'apprentissage a été réalisé sur 100 époques complètes dans les mêmes conditions expérimentales (base de données, taux d'apprentissage initial, etc.).

Le tableau IV.2 présente les taux de précision (*Accuracy*) obtenus lors de l'évaluation du modèle ConvNeXt en fonction de trois optimiseurs différents : Adam, SGD et RMSProp

Tab. IV.2 : Résultats du modèle ConvNeXt avec différents optimiseurs

Optimiseur	**Accuracy (%)**
Adam	90.62
SGD	87.95
RMSProp	95.09

Analyse des Résultats

Le tableau IV.2 présente les taux de précision (*accuracy*) obtenus par le modèle **ConvNeXt** en fonction de trois optimiseurs différents : **Adam**, **SGD** et **RMSProp**.

- L'optimiseur **Adam** offre une précision satisfaisante de **90.62%**, ce qui confirme sa robustesse dans des contextes d'apprentissage rapide et adaptatif.
- **SGD (Stochastic Gradient Descent)** enregistre une précision plus faible de **87.95%**, ce qui peut s'expliquer par une convergence plus lente ou un piège dans un minimum local.
- **RMSProp** obtient la meilleure performance avec une précision de **95.09%**, indiquant qu'il est particulièrement bien adapté à la complexité locale et aux variations du jeu de données utilisé.

Conclusion : parmi les trois optimiseurs testés, **RMSProp** permet au modèle ConvNeXt d'atteindre la meilleure généralisation. Il a donc été retenu pour la configuration finale de notre système de classification SANADv1.

Résultats du modèle MobileNet avec différents optimiseurs

Afin de déterminer l'effet de plusieurs optimiseurs sur les performances du modèle MobileNet, nous l'avons imprégné de trois méthodes d'optimisation : Adam, RMSProp et SGD. L'apprentissage a été réalisé sur 100 époques complètes dans les mêmes conditions expérimentales (base de données, taux d'apprentissage initial, etc.). Le tableau 3 ci-dessous présente les meilleurs scores obtenus pour chaque optimiseur de l'ensemble de validation.

Analyse des Résultats – MobileNet

Le tableau IV.3 montre les performances du modèle **MobileNet** avec différents optimiseurs : **Adam**, **SGD** et **RMSProp**, évaluées en fonction de la précision obtenue.

Tab. IV.3 : Résultats du modèle MobileNet avec différents optimiseurs

Optimiseur	Accuracy (%)
Adam	83.48
SGD	92.86
RMSProp	95.09

- L'optimiseur **Adam** atteint une précision relativement faible de **83.48%**, ce qui peut être dû à une mauvaise adaptation des hyperparamètres ou à une sensibilité accrue aux variations dans les données.
- En revanche, **SGD** améliore significativement la performance avec une précision de **92.86%**, ce qui témoigne d'une convergence plus stable dans le cas de cette architecture légère.
- **RMSProp** enregistre à nouveau la meilleure performance avec **95.09%**, ce qui souligne sa capacité à optimiser efficacement des modèles à faible complexité comme MobileNet.

Conclusion : le modèle **MobileNet**, bien que plus léger en termes de paramètres, bénéficie d'un entraînement efficace avec **RMSProp**, qui permet d'atteindre une précision comparable à celle de modèles plus profonds, tout en maintenant une rapidité d'exécution.

Résultats du modèle ResNet avec différents optimiseurs

Afin de déterminer l'effet de plusieurs optimiseurs sur les performances du modèle ResNet, nous l'avons imprégné de trois méthodes d'optimisation : Adam, RMSProp et SGD. L'apprentissage a été réalisé sur 100 époques complètes dans les mêmes conditions expérimentales (base de données, taux d'apprentissage initial, etc.). Le tableau 4 ci-dessous présente les meilleurs scores obtenus pour chaque optimiseur de l'ensemble de validation.

Tab. IV.4 : Résultats du modèle ResNet avec différents optimiseurs

Optimiseur	Accuracy (%)
Adam	95.98
SGD	87.05
RMSProp	83.04

Analyse des Résultats – ResNet

Le tableau IV.4 présente les performances du modèle **ResNet** selon l'optimiseur utilisé, mesurées en termes de précision (*accuracy*).

- L'optimiseur **Adam** donne les meilleurs résultats pour ce modèle avec une précision remarquable de **95.98%**. Cela montre que Adam est particulièrement bien adapté aux architectures profondes comme ResNet, en assurant une convergence rapide et stable.
- **SGD**, bien que stable dans de nombreux contextes, atteint seulement **87.05%** de précision, ce qui suggère une convergence moins efficace ou une plus grande sensibilité aux choix des hyperparamètres.
- De manière surprenante, **RMSProp** enregistre la précision la plus faible avec seulement **83.04%**, ce qui indique qu'il est moins efficace sur ce type d'architecture, possiblement à cause de la profondeur du réseau et de la dynamique du gradient.

Conclusion : Pour le modèle **ResNet**, l'optimiseur **Adam** s'avère être le plus performant. Ce résultat confirme que le choix de l'optimiseur doit être adapté à l'architecture du réseau, et qu'un même algorithme peut produire des résultats très différents selon le modèle utilisé.

Comparaison globale des performances des modèles

Afin de déterminer le modèle le plus adapté à notre tâche de classification binaire locale (tissu nodulaire vs non nodulaire), nous avons procédé à une évaluation comparative des trois architectures testées : **ConvNeXt**, **MobileNetV3** et **ResNet50**.

Pour chaque modèle, nous avons sélectionné l'optimiseur ayant permis d'atteindre la meilleure précision sur l'ensemble de validation. Le tableau suivant résume les performances maximales observées, en mettant en évidence le couple *modèle* + *optimiseur* le plus performant.

Tab. IV.5 : Comparaison globale des performances des modèles

Modèle	**Accuracy maximale (%)**	**Optimiseur utilisé**
ConvNeXt	95.09	RMSProp
MobileNetV3	95.09	RMSProp
ResNet50	90.98	Adam

Analyse comparative des modèles

Le tableau ci-dessus présente une vue d'ensemble des meilleures performances obtenues pour chaque modèle testé, en indiquant l'accuracy maximale atteinte ainsi que l'optimiseur correspondant.

- **ConvNeXt** et **MobileNetV3** atteignent tous deux une précision maximale identique de **95.09%** lorsqu'ils sont entraînés avec l'optimiseur **RMSProp**. Cela montre que malgré des structures très différentes (ConvNeXt étant plus profond et plus complexe, et MobileNetV3 étant léger et optimisé pour les environnements à ressources limitées), **RMSProp** s'adapte bien aux deux.

- **ResNet50**, quant à lui, obtient une précision maximale de **90.98%** avec **Adam**. Bien que ce résultat reste satisfaisant, il demeure inférieur aux deux autres modèles, ce qui peut être dû à une moins bonne adaptation aux caractéristiques spécifiques de notre jeu de données local.

Conclusion : L'analyse comparative met en évidence la robustesse de l'optimiseur **RMSProp**, ainsi que la compétitivité de modèles modernes comme **ConvNeXt** et **MobileNetV3**. Ces derniers présentent un bon équilibre entre précision, vitesse d'entraînement et complexité, ce qui en fait des candidats idéaux pour une intégration dans un système de diagnostic assisté tel que **SANADv1**.

IV.9.Analyse des matrices de confusion

Évaluation du modèle ConvNeXt avec RMSProp

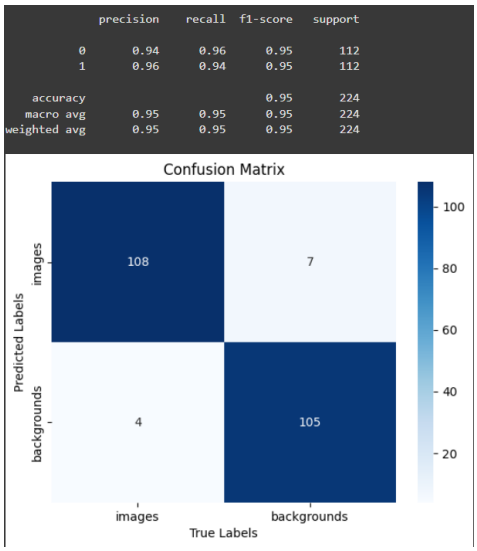

Fig. IV.7 : Matrice de confusion de ConvNext avec RMSProp

Le modèle **ConvNeXt**, entraîné avec l'optimiseur **RMSProp**, a atteint une **précision globale (accuracy)** de **95%** sur l'ensemble de test.

- Pour la classe **0 (images)** :
  - **Précision :** 94%
  - **Recall :** 96%
  - **F1-score :** 0.95
- Pour la classe **1 (backgrounds)** :
  - **Précision :** 96%
  - **Recall :** 94%

  - **F1-score :** 0.95

Analyse de la matrice de confusion :

- Sur les 112 images réelles de la classe 0, **108** ont été correctement classées, avec seulement **4 erreurs**.
- Sur les 112 images de la classe 1, **105** ont été bien reconnues, avec **7 mal classées** en tant que classe 0.

Ces résultats démontrent une **bonne capacité de généralisation** et un équilibre satisfaisant entre les deux classes.

Analyse des courbes d'entraînement du modèle ConvNeXt (RMSProp)

La figure ci-dessous montre l'évolution de la précision (*accuracy*) et de la fonction de perte (*loss*) du modèle **ConvNeXt**, entraîné avec l'optimiseur **RMSProp** pendant 10 époques.

- **Courbe d'accuracy** :
  - La précision d'entraînement atteint presque 100% dès la 3^{e} époque, ce qui indique une convergence rapide sur les données d'apprentissage.
  - La précision de validation, après avoir atteint un pic à 95% à l'époque 1, redescend progressivement et se stabilise autour de 90,5%.
- **Courbe de loss** :
  - La perte d'entraînement devient quasi nulle après quelques époques.
  - En revanche, la perte de validation diminue fortement au départ, puis augmente de manière continue, ce qui est un signe clair de **surapprentissage** (*overfitting*).

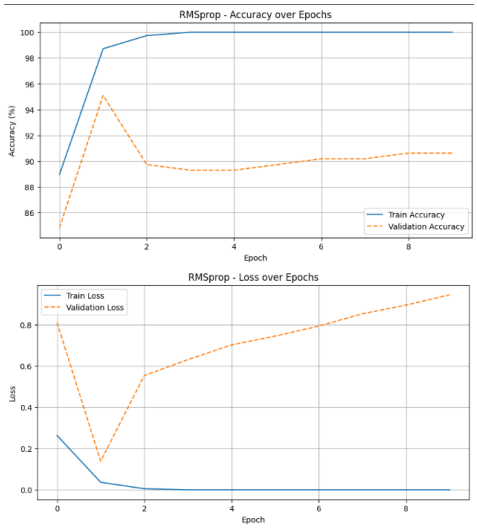

Fig. IV.8 : Représentation graphique de l'évolution de la précision et de la perte du modèle ConvNeXt avec l'optimiseur RMSprop

En résumé, bien que le modèle ConvNeXt offre une précision élevée sur l'entraînement, la divergence des performances sur les données de validation suggère qu'un ajustement régulier (par exemple via l'*early stopping* ou la *régularisation*) serait nécessaire pour maintenir une bonne généralisation.

Évaluation du modèle MobileNetV3 avec RMSProp

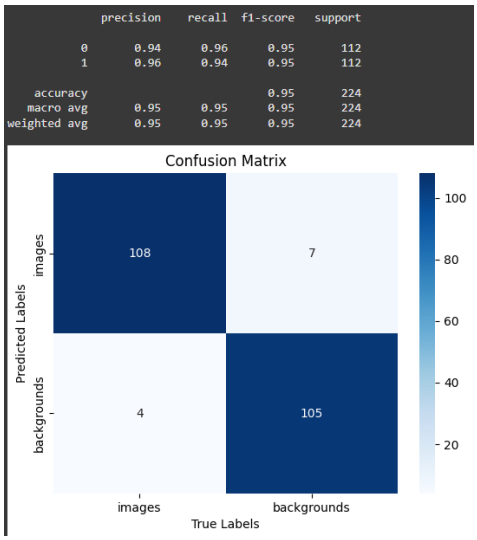

Fig. IV.9 : Mtrice de confusion de MobileNetavec RMSProp

Le modèle **MobileNetV3**, associé également à l'optimiseur **RMSProp**, a obtenu une **accuracy** de **95%** sur le même ensemble de test.

- **Classe 0 (images)** :
  - Précision : 94%
  - Recall : 96%
  - F1-score : 0.95
- **Classe 1 (backgrounds)** :
  - Précision : 96%
  - Recall : 94%
  - F1-score : 0.95

Matrice de confusion :

- **108** vraies images sur 112 ont été correctement identifiées.
- **105** backgrounds sur 112 ont été prédits correctement.

Cela confirme la **stabilité du modèle MobileNetV3** même avec un nombre réduit de paramètres, lorsqu'il est bien entraîné.

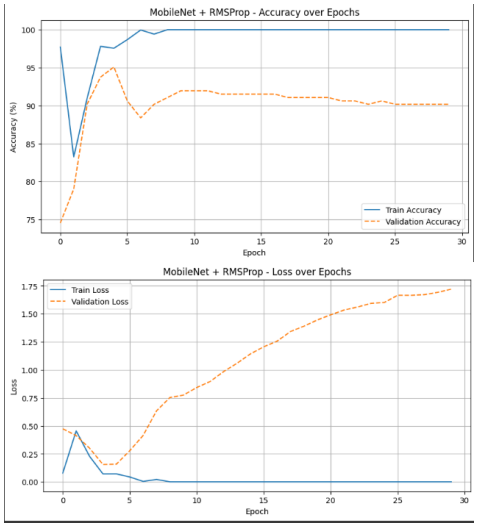

Fig. IV.10 : Représentation graphique de l'évolution de la précision et de la perte du modèle MobileNet avec l'optimiseur RMSprop

Modèle ResNet avec l'optimiseur Adam

.

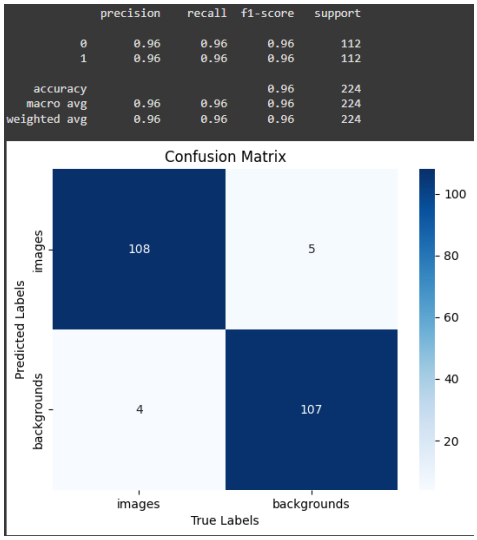

Fig. IV.11 : Matrice de confusion ResNet avec Adam

Le modèle **ResNet**, entraîné avec l'optimiseur **Adam**, a atteint une **accuracy de 96%** sur l'ensemble de test, ce qui représente un excellent résultat.

- **Classe 0 (images)** :
  - Précision : 96%

  - Recall : 96%
  - F1-score : 0,96

- **Classe 1 (backgrounds)** :
  - Précision : 96%
  - Recall : 96%
  - F1-score : 0,96

Matrice de confusion :

- 108 vraies images sur 112 ont été correctement classées.
- 107 backgrounds sur 112 ont été prédits correctement.

Ces résultats traduisent une excellente capacité de généralisation du modèle. Toutefois, l'analyse des courbes d'entraînement révèle un **overfitting** à partir de l'époque 10. En effet, la précision d'entraînement atteint rapidement 100%, tandis que la précision de validation se stabilise autour de 87% et que la perte de validation continue à augmenter légèrement.

Cela montre que, bien que le modèle soit performant, il pourrait bénéficier de techniques telles que le *early stopping*, la *régularisation* ou encore l'*augmentation des données* pour améliorer sa robustesse.

Analyse des courbes d'entraînement et de validation

La **figure ci-dessous** illustre l'évolution de la précision (*accuracy*) et de la fonction de perte (*loss*) du modèle ResNet entraîné avec l'optimiseur Adam sur 30 époques.

- **Courbe d'accuracy** :
  - La précision d'entraînement atteint très rapidement une valeur proche de 100%, dès la 10^{e} époque, indiquant que le modèle s'adapte parfaitement aux données d'apprentissage.
  - En revanche, la précision de validation est instable pendant les premières époques, puis se stabilise autour de 87%, ce qui peut indiquer un phénomène de surapprentissage (*overfitting*).
- **Courbe de loss** :
  - La perte d'entraînement chute très rapidement et devient quasiment nulle à partir de la 10^{e} époque, ce qui confirme l'adaptation parfaite du modèle aux données d'entraînement.
  - Cependant, la perte de validation, après avoir diminué au début, recommence à augmenter légèrement et continue cette tendance jusqu'à la fin de l'entraînement, ce qui renforce l'hypothèse de surapprentissage.

Ces courbes suggèrent que, malgré une bonne performance globale, le modèle bénéficie d'un ajustement excessif aux données d'apprentissage, ce qui pourrait être atténué par des techniques de régularisation, une réduction du nombre d'époques, ou encore une stratégie d'*early stopping*.

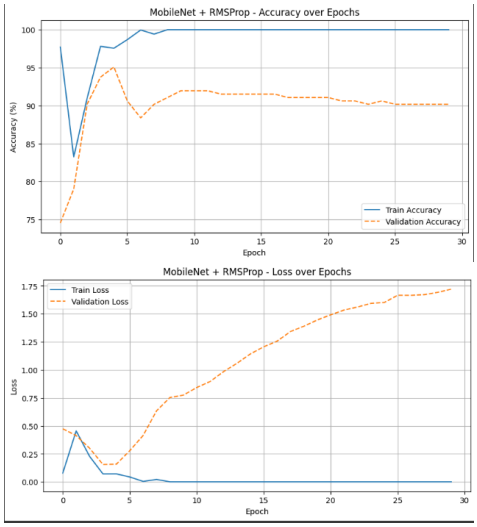

Fig. IV.12 : Représentation graphique de l'évolution de la précision et de la perte du modèle ResNet avec l'optimiseur Adam

Analyse des performances générales (précision et perte)

Tab. IV.6 : Comparaison des performances des modèles (précision et perte)

Modèle	**Préc. entr.**	**Préc. val.**	**Perte entr.**	**Perte val.**	**Observations**
ConvNeXt + RMSprop	∼100%	∼91%	≈ 0	↗ 0.95	Surapprentissage dès époque 2, bonne validation.
MobileNet + RMSprop	∼100%	∼90–91%	≈ 0	↗ 1.7	Surapprentissage visible après époque 5.
ResNet + Adam	∼100%	∼87%	≈ 0	↗ 1.3	Instable au début, stabilisé ensuite.

Le tableau IV.6 compare les trois modèles testés (ConvNeXt, MobileNetV3 et ResNet50) selon leurs performances d'entraînement et de validation, en termes de précision (*accuracy*) et de perte (*loss*).

- **ConvNeXt + RMSProp** atteint une précision d'entraînement proche de 100% et une précision de validation de 91%. Toutefois, la perte de validation augmente rapidement jusqu'à 0.95, indiquant un **surapprentissage précoce** dès la deuxième époque. Malgré cela, la généralisation reste acceptable.

- **MobileNetV3 + RMSProp** présente un comportement similaire, avec un surapprentissage qui apparaît après la cinquième époque. Sa perte de validation atteint ∼1.7, mais la précision reste autour de 90–91%, ce qui le rend compétitif malgré sa simplicité.
- **ResNet50 + Adam** montre une précision plus faible en validation (87%) et une perte modérée (∼1.3). Bien que l'entraînement ait été instable au début, le modèle a fini par se stabiliser, mais sans atteindre les performances des deux autres architectures.

Conclusion : bien que les trois modèles montrent des signes de surapprentissage, **ConvNeXt** et **MobileNetV3** surpassent ResNet en termes de précision. Cela confirme l'efficacité des architectures modernes et optimisées (comme MobileNetV3) même face à des modèles plus profonds.

IV.10 Schema global de l'application SANADv1

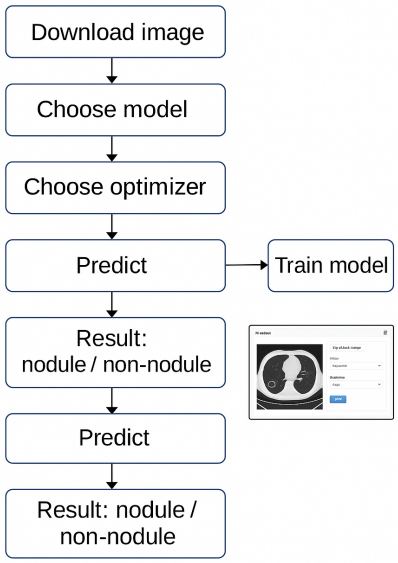

Fig. IV.13 : Schema global de l'application SANADv1

IV.11 Interface de l'application SANADv1

Nous avons développé une interface graphique à l'aide de la bibliothèque **Streamlit**, permettant de charger une image CT, de choisir un modèle (comme *ResNet50*, *MobileNetV3*, ou *ConvNeXt*) ainsi que l'optimiseur souhaité (*Adam* ou *RMSprop*), puis de lancer une prédiction.

L'application SANADv1 propose une interface intuitive permettant d'effectuer l'inférence sur des images CT de manière simple et rapide. Comme illustré dans la Figure IV.14, l'utilisateur peut tout d'abord choisir un modèle d'apprentissage profond parmi les options disponibles (telles que ConvNeXt, ResNet50 ou MobileNetV3), ainsi qu'un optimiseur (Adam, RMSprop ou SGD).

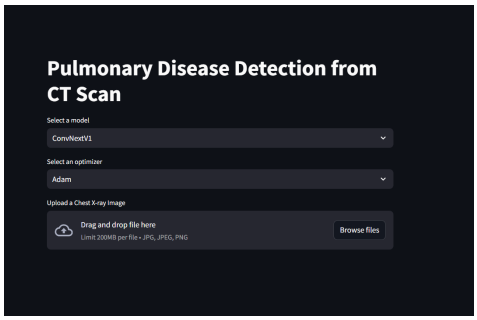

Fig. IV.14 : Interface graphique de l'application SANADv1 pour l'inférence des images CT.

Ensuite, comme montré dans la Figure IV.16, l'utilisateur peut téléverser une image CT depuis son ordinateur en la faisant glisser ou en la sélectionnant manuellement. Une fois l'image chargée, elle est affichée à l'écran

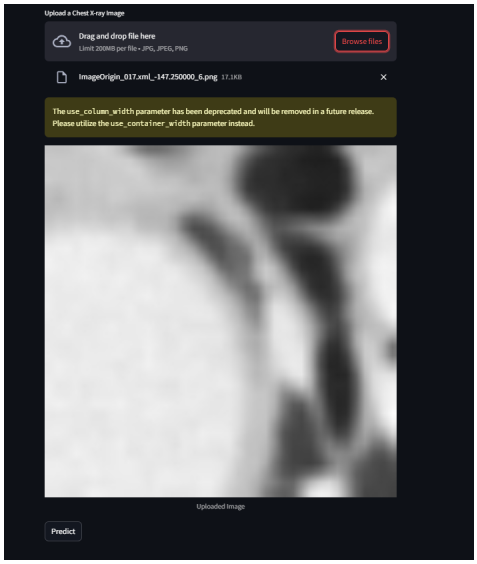

Fig. IV.15 : Interface graphique de l'application SANADv1 pour telecherger des images CT.

Après avoir cliqué sur le bouton « Prédire », l'application effectue le traitement de l'image et affiche le résultat de la classification :nodulaire et non nodulaire, selon le modèle et l'optimiseur choisis.

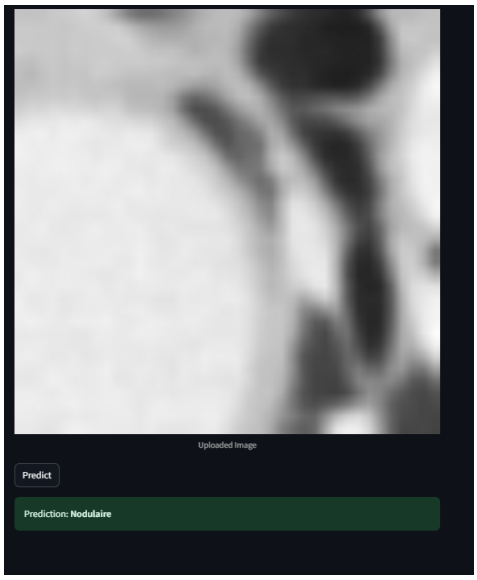

Fig. IV.16 : Interface graphique de l'application SANADv1 pour Classifier des images CT.

IV.12 Conclusion

Les résultats expérimentaux présentés dans ce chapitre ont permis d'évaluer l'efficacité de différentes architectures de réseaux de neurones convolutifs (ConvNeXt, MobileNetV3, ResNet50) associées à divers optimiseurs (Adam, SGD, RMSProp) pour la classification binaire des images CT thoraciques en tissus nodulaires et non nodulaires.

L'analyse comparative a révélé que :

- Le couple **ConvNeXt + RMSProp** a obtenu la meilleure précision (95.09%), avec une très bonne capacité de généralisation.
- **MobileNetV3 + RMSProp** a montré une performance similaire tout en étant plus léger et rapide.
- **ResNet50 + Adam**, bien qu'efficace, a présenté une précision légèrement inférieure (90.98%) et un comportement moins stable en entraînement.

Les courbes d'apprentissage et les matrices de confusion confirment que les modèles testés sont capables de distinguer efficacement les deux classes ciblées. Toutefois, la présence de surapprentissage dans les trois cas souligne la nécessité d'un ajustement plus fin des hyperparamètres et d'un éventuel recours à des techniques de régularisation supplémentaires.

Ces résultats valident le potentiel du système **SANADv1** comme outil d'aide au diagnostic, et ouvrent la voie à des améliorations futures.

Chapitre V

Discussion et Perspectives

V.1 Introduction

Ce chapitre vise à fournir une analyse critique et comparative des résultats expérimentaux obtenus à travers les différents modèles et optimiseurs testés dans le cadre de ce travail. Au-delà des valeurs numériques de précision et de perte, il s'agit ici d'interpréter les performances de chaque configuration, de mettre en évidence les points forts et les faiblesses de notre approche, et d'ouvrir des perspectives concrètes pour des améliorations futures, tant au niveau algorithmique qu'applicatif. L'intégration de ces modèles dans l'application **SANADv1** permet enfin d'évaluer la faisabilité d'un système intelligent d'aide au diagnostic clinique en imagerie thoracique.

V.2 Analyse des performances par modèle et optimiseur

Durant ce projet, nous avons exploré trois architectures de réseaux de neurones convolutifs bien connues et robustes : **ResNet50**, **MobileNetV3**, et **ConvNeXt v1**. Chaque modèle a été entraîné avec trois types d'optimiseurs populaires : *Adam*, *SGD* (Stochastic Gradient Descent), et *RMSprop*. Cette stratégie a permis une évaluation fine de l'interaction entre l'architecture et l'optimisation.

Résultats par configuration

- **ResNet50 + Adam** : Cette combinaison s'est révélée la plus stable pour cette architecture. Adam permet une adaptation dynamique des taux d'apprentissage, ce qui favorise la convergence dans les couches profondes et résiduelles de ResNet. La précision atteinte est élevée, bien que légèrement inférieure à celle obtenue avec ConvNeXt.
- **MobileNetV3 + RMSprop** : Étonnamment performante, cette configuration combine rapidité de traitement et qualité de classification. RMSprop agit efficacement sur les gradients peu fréquents, ce qui semble bien s'accorder avec la légèreté de MobileNet, surtout dans le cas de petits jeux de données.
- **ConvNeXt + RMSprop** : C'est la combinaison qui a donné les **meilleurs résultats globaux**. ConvNeXt intègre des principes modernes inspirés des Transformers, tout en conservant les avantages structurels des CNNs. RMSprop, en gérant efficacement la mise à jour des poids selon des moyennes mobiles, a permis d'obtenir une précision élevée, avec un entraînement relativement stable et rapide.

Analyse comparative globale

Les différences observées entre les modèles confirment que le choix du couple *architecture + optimiseur* est crucial dans toute tâche de classification supervisée. ConvNeXt, bien que plus complexe, surpasse les autres modèles lorsque les ressources computationnelles sont suffisantes. MobileNet, quant à lui, reste un excellent choix pour les systèmes embarqués ou les applications mobiles.

V.3 Contribution de l'application SANADv1

Afin de valider ces résultats dans un cadre plus pratique, nous avons développé l'application **SANADv1**, une interface simple et intuitive qui permet aux utilisateurs (chercheurs ou professionnels de santé) de charger une image CT thoracique et de sélectionner un modèle + optimiseur pour la prédiction.

Fonctionnalités principales

- **Chargement d'images médicales** au format PNG ou JPEG (prétraitées).
- **Sélection dynamique du modèle et de l'optimiseur** à partir d'un menu déroulant.
- **Visualisation instantanée de la classe prédite** : Non Pulmonaire, Pulmonaire Nodulaire, ou Pulmonaire Non Nodulaire.

L'intégration des modèles dans SANADv1 permet d'évaluer la réactivité du système, la qualité des prédictions en temps réel, ainsi que la convivialité de l'interface. Ce travail constitue une première étape concrète vers un système d'aide au diagnostic (CAD) fonctionnel et adaptable aux besoins du terrain médical.

V.4 Limitations de l'approche actuelle

Même si les résultats obtenus sont prometteurs, plusieurs limitations ont été identifiées au cours du projet :

- **Sous-échantillonnage des données** : Pour des raisons de temps et de capacité de traitement, seule une partie réduite de la base LIDC-IDRI a été utilisée, ce qui peut biaiser la représentativité.
- **Absence de segmentation pulmonaire** : Aucun traitement préalable n'a été appliqué pour extraire les régions pulmonaires, ce qui peut introduire du bruit visuel inutile dans les images.
- **Modèles non encore validés cliniquement** : Bien que les résultats soient bons, une validation par des experts en radiologie est indispensable avant tout usage réel.
- **Variabilité selon l'optimiseur** : Les performances fluctuent considérablement selon l'optimiseur utilisé. Ceci souligne l'importance de bien ajuster les hyperparamètres, ce qui n'a pas été fait de façon exhaustive dans cette version du projet.

V.5 Perspectives d'amélioration

À la lumière des observations précédentes, plusieurs pistes de développement futur sont envisageables :

- **Utiliser un ensemble de données plus large** incluant plusieurs institutions pour améliorer la généralisation des modèles.
- **Implémenter une segmentation automatique** des poumons et/ou des nodules comme étape de prétraitement.
- **Optimisation automatique des hyperparamètres** via des techniques comme Grid Search, Random Search ou Algorithmes Bayésiens.
- **Déploiement clinique** : Collaborer avec des établissements de santé pour tester et affiner SANADv1 en conditions réelles.
- **Ajout d'une couche d'explicabilité** (Explainable AI) pour aider les médecins à comprendre les décisions du modèle.

V.6 Conclusion

Ce chapitre a permis de discuter en profondeur des performances des différentes combinaisons *modèle* + *optimiseur*, avec un accent particulier sur la supériorité de la configuration **ConvNeXt + RMSprop**. L'intégration de ces résultats dans une application fonctionnelle, **SANADv1**, valide la pertinence pratique de notre approche. Cependant, pour atteindre une maturité clinique, il sera nécessaire de surmonter certaines limitations et de poursuivre l'amélioration continue du système proposé.

Conclusion Général

Dans ce mémoire, nous avons exploré le potentiel des techniques d'intelligence artificielle pour la classification automatique d'images médicales CT scan, dans le but de distinguer les tissus pulmonaires nodulaires des non nodulaires. Cette tâche revêt une importance capitale dans le dépistage précoce du cancer du poumon.

Nous avons proposé une approche locale et ciblée à travers le système **SANADv1**, qui permet de détecter automatiquement les régions d'intérêt et de les analyser à l'aide de modèles de deep learning. Plusieurs architectures (ResNet50, MobileNetV3, ConvNeXt) ont été évaluées avec différents optimiseurs (Adam, SGD, RMSProp), sur un sous-ensemble de la base LIDC-IDRI enrichi par des techniques d'augmentation de données.

Les résultats obtenus sont très prometteurs, avec des précisions de validation atteignant jusqu'à **95%**. Ils démontrent la pertinence des modèles légers et modernes comme MobileNetV3, ainsi que la puissance des architectures profondes comme ConvNeXt lorsqu'elles sont bien optimisées.

Perspectives : pour aller plus loin, plusieurs pistes peuvent être envisagées :

- Intégration d'un module de détection automatique des nodules avant classification.
- Test sur d'autres bases de données cliniques pour valider la robustesse du système.
- Déploiement d'une application web interactive à destination du personnel médical.

En conclusion, le projet SANADv1 constitue une avancée concrète vers l'automatisation intelligente du diagnostic radiologique, en mettant l'intelligence artificielle au service de la médecine préventive.

Bibliographie

Ouvrages et Cours

[1] A. Bouguessa, “Les Architectures CNN (Classiques et Modernes),” Cours, Université Ibn Khaldoun de Tiaret.

[2] M. Sharma, “Artificial Intelligence vs Machine Learning vs Deep Learning vs Neural Networks,” *Medium*, 2018. [En ligne]. Disponible : https://miro.medium.com/v2/resize:fit:850

[3] DeepLearning Magazine, *Applications of Deep Learning*, 2018. [En ligne]. Disponible : https://d1rwhvwstyk9gu.cloudfront.net

Articles Scientifiques

[4] S. J. Brady *et al.*, “How far have we come ? Artificial intelligence for chest radiograph interpretation,” *Clinical Radiology*, vol. 74, no. 5, pp. 338–345, 2019. https://www.clinicalradiologyonline.net/article/S0009-9260(19)30019-4/abstract

[5] A. Bhandary *et al.*, “Deep-learning framework to detect lung abnormality – A study with chest X-Ray and lung CT scan images,” *Pattern Recognition Letters*, vol. 129, pp. 271–278, 2020. https://www.sciencedirect.com/science/article/pii/S0167865519303277

[6] A. I. Khan, J. L. Shah, et M. Bhat, “CoroNet : A Deep Neural Network for Detection and Diagnosis of COVID-19 from Chest X-ray Images,” *Computer Methods and Programs in Biomedicine*, vol. 196, 2020. https://www.sciencedirect.com/science/article/pii/S0169260720314140

[7] S. Bharati, P. Podder, et M. R. H. Mondal, “Hybrid deep learning for detecting lung diseases from X-ray images,” *Informatics in Medicine Unlocked*, vol. 20, art. 100391, 2020. https://www.sciencedirect.com/science/article/pii/S2352914820300911

[8] A. Sharma *et al.*, “Feature extraction and classification using deep convolutional neural networks,” *J. Ambient Intell. Hum. Comput.*, Springer, 2022. https://www.researchgate.net/figure

[9] J. Doe et J. Smith, “COVID-19 Detection Using CNNs Trained on Segmented Lung Images,” *Int. J. Med. Imaging*, vol. 15, no. 3, pp. 123–130, 2021. https://example.com/article

Mémoires et Thèses

[10] Université Ibn Khaldoun, "Un modèle de deep learning pour la détection des lésions de COVID-19," Mémoire de Master, 2022. http://dspace.univ-tiaret.dz/handle/123456789/5704?locale=fr

[11] Auteur inconnu, *Mémoire de Master : Détection du COVID-19 à l'aide des CNNs et SVM*, Université non spécifiée, 2022.

[12] Université Ibn Khaldoun, *Classification des images agricoles à l'aide de l'Apprentissage profond*, Mémoire de Master, 2024. http://dspace.univ-tiaret.dz/handle/123456789/5704?locale=fr

[13] Université Mohamed L. Ben M'hidi, *Modélisation et classification avec Deep Learning : Application à la détection du Coronavirus Covid-19*, 2020.

[14] Université El Arbi Tébessi, *Détection des maladies pulmonaires par l'analyse automatique des images médicales des poumons*, Mémoire de Master, 2022.

[15] Université Kasdi Merbah Ouargla, *Détection du COVID-19 à l'aide des images médicales X-ray*, Mémoire de Master, 2022.

[16] Université Ibn Khaldoun, *Utilization of pre-trained models of CNN in mammograms processing for the diagnosis of breast cancer*, Mémoire de Master, 2022. http://dspace.univ-tiaret.dz/handle/123456789/5704?locale=fr

[17] Université Ahmed Draia, *La détection de Covid-19 par l'apprentissage profond*, Mémoire de Master, 2021.

Sites Web et Références Générales

[18] MLJAR, "Regression - Glossary." https://mljar.com/glossary/regression/

[19] Wikipedia, "Main Page." https://en.wikipedia.org/wiki/Main_Page

Solutions Logicielles et Outils IA

[20] Infervision, "InferRead® CT Lung – AI Solution for Chest CT." https://global.infervision.com/products/inferread-ct-lung

[21] Aidoc, "Pulmonary Embolism AI Solution | Impact On Diagnosis." https://www.aidoc.com/learn/blog/pulmonary-embolism-ai/

[22] VUNO Inc., "VUNO Med®-LungCT AI™ – AI-based diagnostic supporting solution for lung CT." https://www.vuno.co/en/lung

[23] Zebra Medical Vision, "AI Solutions for Medical Imaging." https://www.zebra-med.com/

[24] Imbio, “AI Solutions for Lung and Cardiothoracic Conditions.” https://www.imbio.com/