

Faculté des Sciences Exactes et de l'Informatique
Département de Mathématiques et d'Informatique
Filière : Informatique

MEMOIRE DE FIN D'ETUDES
Pour l'Obtention du Diplôme de Master en Informatique
Option : **Ingénierie des Systèmes d'Information**

THEME :
**Prédiction automatique du degré de sévérité
de tuberculose pulmonaire**

Etudiant(e) : « **ZAHAF Kheira** »

« **LATROCH Malika** »

Encadrant(e) : « Dr. HAMADI Abdelkader »

Dédicaces

Je dédie ce modeste travail aux personnes les plus chers dans ma vie : A l'homme de ma vie , mon exemple éternel , mon soutien moral et source de joie et de bonheur , celui qui sait toujours sacrifier pour me voir réussir , que Dieu soit te guide, à toi mon père .

A ma très chère et douce mère , qui ma comblé de son soutien et m'a voué un amour inconditionnel. Tu es pour moi un exemple de courage et de sacrifices continues. Que cet humble travail ,témoigne mon affection , mon éternel attachement et qu'il appelle sur moi ta continuelle bénédiction. A mes très chers frères et sœurs .

A toute la famille « LATROCH » et la famille «ZAHAF» mes aimables amies .

** LATROCH MALIKA, ZAHAF KHEIRA**

Remerciements

Nous remercions Allah le tout puissant , qui nous a donné la force et la patience et le courage pour l'accomplissement de ce travail et qui a récompensé nos prières.

Nous exprimons nos profondes gratitude et respectueuse reconnaissance à Notre encadreur Mr. « HAMADI Abdelkader » Pour sa bonne volonté d'accepter de nous encadrer, pour tout le temps qu'il Nous a octroyé et pour tous les conseils qu'il nous a prodigué. Pour sa directive précieuse, et pour la qualité de son suivi durant toute La période de notre projet.

Nous tenons à remercier les honorables membres du jury Qui ont bien voulu nous faire l'honneur d'évaluer ce manuscrit.

Nos vifs remerciements s'adressent également aux personnes qui nous ont toujours accompagné et aidé durant toute la vie.

Table des figures

1.1	La progression de la tuberculose pulmonaire	5
1.2	Les types de tuberculose pulmonaire	6
2.1	Schéma d'un neurone	14
2.2	L'architecture VGGNet	17
2.3	Le module Inception	17
2.4	L'architecture GoogleNet	18
2.5	l'architecture ResNet	18
2.6	Matrice de confusion	19
4.1	Système développé	31

Liste des tableaux

3.1	Résultats de la prédiction du degré de sévérité de tuberculose pulmonaire dans la tâche Tuberculosis de ImageCLEF 2018 . . .	29
4.1	Composition des collections d'apprentissage et validation de corpus	32
4.2	Matrice de confusion obtenue par le modèle <i>InceptionV3</i> sur la collection de validation.	33
4.3	Résultats en termes de précision moyenne(Accuracy) et de l'aire sous la courbe ROC (AUC) obtenus sur la collection de validation.	33

Résumé

La tuberculose (TBC) est une maladie infectieuse qui demeure une cause majeure de décès dans le monde. D'autre part, il y a eu un développement considérable ces dernières années, dans le domaine de l'apprentissage en profondeur, qui permet la classification d'images extrêmement hétérogènes. De plus, il a été vérifié que les réseaux de neurones convolutionnels profonds (CNN) sont des algorithmes très prometteurs pour diverses tâches visuelles. Nous nous intéressons dans ce travail à la prédiction automatique du degré de sévérité de la tuberculose pulmonaire via une analyse d'images tomodensitométriques. Cela permettra d'accélérer le diagnostic de la maladie à partir d'une image de radiologie sans avoir recours à des analyses médicales coûteuses. Vu la qualité des images de radiologie, cette étude s'avère beaucoup plus difficile qu'une analyse d'images ordinaires. Dans ce rapport nous nous intéressons à la prédiction automatique du degré de sévérité de tuberculose pulmonaire.

Dans ce rapport, nous nous intéressons à l'utilisation d'un réseau de neurones profond pour la prédiction automatique du degré de sévérité des cas tuberculeux. Nous proposons d'utiliser l'architecture Inception. Nos différentes expérimentations sont réalisées dans le contexte de la tâche Tuberculosis de ImageCLEF2019 [20].

Table des matières

Dédicaces	i
Remerciements	ii
Résumé	v
Table des figures	vi
Table des matières	1
Introduction générale	2
1 Tuberculose pulmonaire	4
1.1 Introduction	4
1.2 Tuberculose pulmonaire	4
1.3 Types de tuberculose pulmonaire	6
1.3.1 Tuberculose infiltrante	6
1.3.2 Tuberculose Focale	7
1.3.3 Granulome tuberculeux	7
1.3.4 Tuberculose miliaire	8
1.3.5 Tuberculose Fibro-cavernous	8
1.4 Conclusion	9
2 Apprentissage machine et imagerie médicale	10
2.1 Introduction	10
2.2 Imagerie médicale	10
2.2.1 Formats de fichiers d'images médicales	11
2.3 Apprentissage automatique	13
2.3.1 Apprentissage supervisé	13

2.3.2	Apprentissage par ensemble	15
2.3.3	Apprentissage profond	16
2.3.4	Apprentissage par transfert	18
2.3.5	Mesures d'évaluation	19
2.4	Utilisation de l'apprentissage machine pour l'analyse d'images	22
2.4.1	Apprentissage profond pour l'analyse d'images médicales	22
2.5	Conclusion	23
3	Prédiction automatique du degré de sévérité de tuberculose pulmonaire	24
3.1	Introduction	24
3.2	Difficulté et défis de l'analyse des images médicales	24
3.3	Compagne d'évaluation ImageCLEF	25
3.3.1	ImageCLEF 2017	26
3.3.2	ImageCLEF 2018	26
3.3.3	ImageCLEF 2019 [20]	26
3.4	Tâche Tuberculosis	26
3.4.1	Organisation de la tâche	26
3.4.2	Données	27
3.4.3	Mesures d'évaluation	27
3.5	Quelques travaux de prédiction du degré de sévérité de la tuberculose pulmonaire	28
3.5.1	Travail de UIIP_BioMed [25]	28
3.5.2	Travail de MEDGIFT [14]	28
3.5.3	Travail de MostaganemFSEI [18]	28
3.6	Résultats de prédiction du degré de sévérité de la tuberculose pulmonaire	29
4	Contribution et résultats	30
4.1	Introduction	30
4.2	Contribution	30
4.2.1	Extractions des slices	30
4.2.2	Filtrage	31
4.2.3	Apprentissage	32
4.3	Expérimentation et résultats	32
4.4	Conclusion	33
	Conclusion générale	34

Introduction générale

La tuberculose est une maladie infectieuse causée par une bactérie appelée *Bacillus mycobacterium-tuberculosis*. Avec une estimation de 10,4 millions de nouveaux cas tuberculeux et un taux de mortalité de 1,8 million dans le monde, cette maladie est restée l'une des dix premières causes de mortalité dans le monde en 2015. La tuberculose attaque les poumons, mais peut également affecter d'autres parties du corps. Un diagnostic précis et rapide est la clé du contrôle de cette maladie, mais les tests traditionnels de TBC produisent des résultats inexacts ou trop longs pour être définitifs.

L'apprentissage machine (automatique) est devenu l'un des principaux outils pour l'analyse d'images médicales. Les techniques d'apprentissage automatique constituent des solutions pour développer des outils permettant d'aider les médecins à diagnostiquer, à prédire le risque d'atteinte de maladies. Dans ce même contexte, des algorithmes impliquant l'apprentissage profond ont été testés pour diagnostiquer la présence ou l'absence de tuberculose. Les résultats obtenus ont été intéressants. En effet, les algorithmes ont atteint un taux de précision impressionnant de 96% [23] un résultat qui s'avère meilleur que celui de l'intervention de nombreux radiologues. Plusieurs chercheurs se sont intéressés à cette maladie, notamment dans le cadre du défi international ImageCLEF dans ses éditions de 2017 [12], 2018 [21] et 2019 [3] où une tâche spécifique nommée "Tuberculosis" a été défiée à la maladie de tuberculose.

Or, appliquer les techniques d'apprentissage automatique en imagerie médicale en général et spécialement pour le cas des images de radiologie 3D, s'avère très difficile à cause de plusieurs facteurs tels que : la taille des données, la difficulté d'extraire des descripteurs efficaces de ces données et rajoutons à cela l'ensemble des défis relatifs au domaine de l'apprentissage automatique en général.

Nous nous intéressons dans ce travail à la prédiction automatique du degré de sévérité de tuberculose pulmonaire via une analyse automatique des images tomodensitométriques de patients tuberculeux. Ce travail peut donner plusieurs avantages à la société, surtout dans le secteur de santé. En effet, résoudre cette problématique pourrait : accélérer le diagnostic de la maladie à partir d'une image de radiologique sans avoir recours à des analyses médicales coûteuses, assister ou couvrir l'absence de spécialistes et de radiologues.

Le reste du rapport est structuré dans trois chapitres. Le premier chapitre présente la maladie de tuberculose, ses types et symptômes. Le deuxième chapitre survole le domaine de l'imagerie médicale et présente quelques bases de l'apprentissage machine ainsi que son exploitation en imagerie médicale. Le dernier chapitre évoque quelques travaux abordant la problématique de prédiction du degré de sévérité de la tuberculose pulmonaire qui ont été réalisés dans le cadre de la campagne d'évaluation internationale ImageCLEF. Le chapitre 4 décrit l'évaluation de notre contribution en présentant les expérimentations effectuées dans le contexte du défi international ImageCLEF 2019 [20] en discutant les résultats obtenus. Le rapport se termine par une conclusion générale.

Chapitre 1

Tuberculose pulmonaire

1.1 Introduction

La tuberculose (TBC) est une maladie infectieuse transmissible et non immunisante, avec des signes cliniques variables. Elle est provoquée par une mycobactérie correspondant à différents germes et principalement Bacille de Koch (BK) [7]. Cette maladie affecte habituellement les poumons. C'est pourquoi nous choisissons d'axer notre étude sur le tuberculose pulmonaire.

Nous présentons dans ce qui suit de ce chapitre la maladie de tuberculose, ses types et ses symptômes.

1.2 Tuberculose pulmonaire

Tuberculose pulmonaire ou phthisie est une maladie infectieuse contagieuse due à une bactérie. La bactérie responsable de la tuberculose pulmonaire est transmise par voie aérienne, c'est la forme la plus courante de tuberculose (environ 70% des cas). Les bactéries détruisent les tissus pulmonaires, créant ainsi des cavités. La maladie reste localisée dans les poumons [10].

La tuberculose pulmonaire provient de la réactivation du foyer de primo-infection pulmonaire ; elle se traduit par une altération de l'état général (fièvre à prédominance vespérale, fatigue, amaigrissement), des sueurs nocturnes, une toux plus ou moins grasse, des crachats parfois sanglants (hémoptysies). La radiographie thoracique met en évidence des opacités (nodules) et des clartés (cavernes) dans les parties supérieure et postérieure du poumon.

La dissémination sanguine du bacille est plus rare et est à l'origine de la tuberculose miliaire. Des formes viscérales telles que la pleurésie ou la péricardite tuberculeuse sont possibles. Outre une fièvre et un amaigrissement, symptômes communs à toutes les atteints tuberculeuses, ces formes viscérales présentent chacune des caractéristiques particulières [6]. La figure 1.1 montre les étapes de l'évolution de la tuberculose dans les poumons.

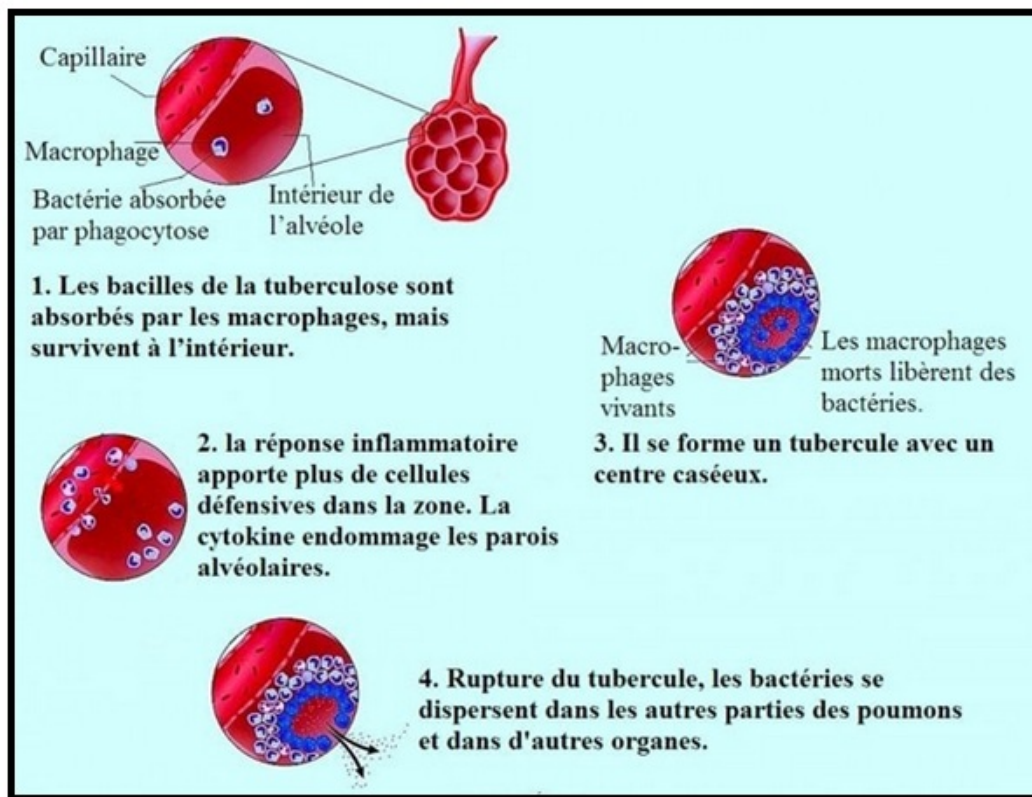


FIGURE 1.1 – La progression de la tuberculose pulmonaire

1.3 Types de tuberculose pulmonaire

On distingue cinq types de tuberculose pulmonaire : Infiltrante, Focale, Tuberculome, Miliaire et Fibro-Caverneuse.

La figure 1.2 montre la distribution de la maladie dans les poumons et la localisation de chaque type de tuberculose pulmonaire. La première rangée, montre les types de tuberculose pulmonaire Infiltrante, Focale, et Tuberculome (de gauche à droite). La deuxième rangée contient les types Milliaire et Fibro-caverneux.

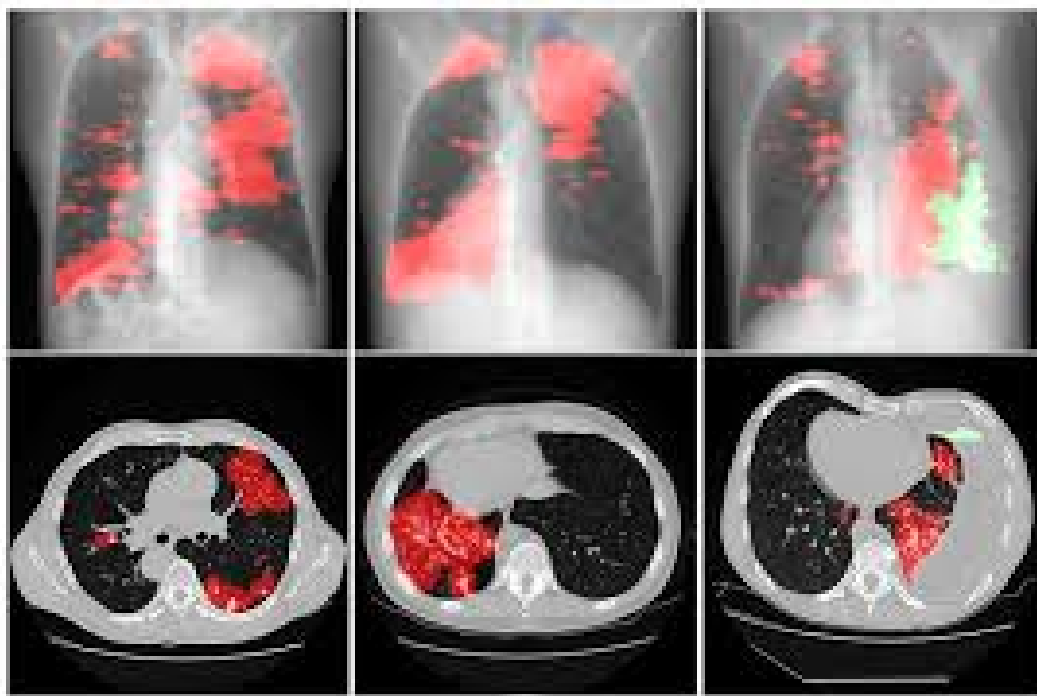


FIGURE 1.2 – Les types de tuberculose pulmonaire

1.3.1 Tuberculose infiltrante

La tuberculose pulmonaire infiltrante est une forme clinique de la tuberculose qui survient dans le contexte d'une hypersensibilité spécifique du tissu

pulmonaire et d'une augmentation significative de la réponse tissulaire exsudative dans la zone inflammatoire.

Les caractéristiques cliniques et morphologiques de la tuberculose infiltrante sont considérées comme des atteintes pulmonaires généralisées avec une tendance à la progression rapide du processus tuberculeux.

Le développement de la tuberculose infiltrante est associé à la progression de la tuberculose focale, à l'apparition et à l'expansion rapide de la zone d'infiltration autour des foyers de tuberculose récents ou anciens. La propagation de l'inflammation périfocale entraîne une augmentation significative du volume des lésions du tissu pulmonaire.

Un infiltrat tuberculeux est un complexe de foyers frais ou anciens avec une vaste zone d'inflammation perifocale. Les infiltrations sont plus souvent localisées dans les segments de 1, 2 et 6 m du poumon, c.-à-d. Dans les départements où se trouvent habituellement les foyers de tuberculose.

1.3.2 Tuberculose Focale

La tuberculose focale est définie par la localisation de la tuberculose dans deux sites non contigus. C'est une forme grave puisque la mortalité atteint 16 à 25 %, elle est surtout l'apanage des sujets immunodéprimés. La localisation préalable et sa survenue dans un contexte de tuberculose disséminée chez une patiente non infectée par le VIH sont exceptionnelles.

1.3.3 Granulome tuberculeux

En médecine, un granulome est un terme assez général désignant de petites papules érythémateuses ou tumeurs vasculaires inflammatoires ou diverses formes d'amas de cellules épithélioïdes entourés de lymphocytes, qui apparaissent sur la peau, des muqueuses ou des organes internes. Il peut ou non être source de prurit ou contenir du pus. On considère qu'ils sont une réaction inflammatoire chronique provoquée par des agents infectieux ou non infectieux. Pour le dermatologue, le mot peut avoir deux sens différents :

- petit agrégat nodulaire constitué de cellules mononucléées inflammatoires ;
- agrégat de macrophages modifiés ressemblant à des cellules épithéliales ;

Ce type de granulome est souvent bordé de lymphocytes porteurs de cellules géantes multinucléées. Son nom évoque sa forme la plus fréquente en forme de grain, mais il prend parfois l'apparence de plaques. Dans certains cas, il peut y avoir présence de pus. Certains granulomes renferment des éosinophiles ainsi que des cellules plasmiques. Une fibrose borde souvent la périphérie de la lésion [2].

1.3.4 Tuberculose miliaire

La tuberculose miliaire est une forme rare de la maladie. Généralement, la bactérie responsable de l'infection attaque un seul organe (poumon, cerveau) ou une seule partie du corps (ganglions, os, appareil urinaire). La tuberculose miliaire se caractérise, quant à elle, par la diffusion des lésions dans l'organisme.

La bactérie est transportée par le sang. Cette complication est peu fréquente et survient souvent à la suite d'une tuberculose pulmonaire. Elle concerne surtout les personnes âgées et celles qui présentent une maladie affaiblissant leur système immunitaire (sida, diabète).

Les signes cliniques de cette forme de tuberculose ne sont pas spécifiques. La toux, la fièvre et le gonflement des ganglions (adénopathie) font partie des symptômes les plus fréquemment rapportés.

La tuberculose miliaire peut également entraîner une augmentation du volume du foie (hépatomégalie) et de la rate (splénomégalie), une inflammation du pancréas (pancréatite) et une insuffisance surrénalienne [9].

1.3.5 Tuberculose Fibro-cavernous

La tuberculose pulmonaire caverneuse fibreuse connue comme une maladie chronique dans laquelle les poumons ont formé une cavité, est la principale caractéristique qui distingue la tuberculose pulmonaire caverneuse de

la tuberculose fibreuse. Typiquement la tuberculose pulmonaire caverneuse fibreuse est la prochaine étape après une tuberculose délicate et infiltrante. Derrière lui pourrait venir la dernière étape tuberculeuse l'étape finale de la maladie du poumon, où de vastes zones peuvent être facilement remplacés par du tissu fibreux, ce qui peut entraîner de graves complications [1].

1.4 Conclusion

Nous avons présenté dans cette partie la tuberculose pulmonaire et ses types. Afin de diagnostiquer cette maladie nous pouvons faire appel à l'imagerie médicale afin d'automatiser l'analyse des images de patients, typiquement des images de radiologique. Nous aborderons cette problématique dans en détails dans le chapitre 3. Dans le chapitre 2, nous présenterons les notions de base permettant de réaliser une telle tâche.

Chapitre 2

Apprentissage machine et imagerie médicale

2.1 Introduction

Dans ce chapitre, nous allons présenter le domaine de l'imagerie médicale ainsi que les méthodes d'apprentissage machine utilisées dans ce domaine dans le but d'effectuer des diagnostics médicaux ou des analyses sémantiques des images médicales.

Nous rappelons que notre objectif est l'aide au diagnostic de la maladie de tuberculose de manière automatique. Une solution pour automatiser cette tâche est d'exploiter des images de radiologie ou de scanner, dans lesquelles la maladie peut être détectée. Pour cela, il est essentiel de produire des données utiles et immédiatement exploitables en intégrant l'intelligence artificielle aux outils existants de reconnaissance et d'analyse sémantique d'images médicales.

2.2 Imagerie médicale

L'imagerie médicale regroupe les moyens d'acquisition et de restitution d'images du corps humain à partir de différents phénomènes physiques tels que l'absorption des rayons X, la résonance magnétique nucléaire, la réflexion d'ondes ultrasons la radioactivité auxquels on associe parfois des techniques d'imagerie optique comme l'endoscopie [5].

Le but de l'imagerie médicale est de créer une représentation visuelle intelligible d'une information à caractère médical. Cette problématique s'inscrit plus globalement dans le cadre de l'image scientifique et technique : l'objectif est en effet de pouvoir représenter sous un format relativement simple une grande quantité d'informations issues d'une multitude de mesures acquises selon un mode bien défini.

2.2.1 Formats de fichiers d'images médicales

Les formats de fichiers d'images médicales peuvent être divisés en deux catégories. La première concerne les formats destinés à normaliser les images générées par des modalités de diagnostic, comme par exemple le format Dicom. La seconde concerne les formats conçus dans le but de faciliter et de renforcer l'analyse post-traitement, par exemple Analyze, Nifti et Minc.

Analyze

Analyze 7.5 a été créé à la fin des années 1980 en tant que format utilisé par le logiciel commercial Analyze développé à la Mayo Clinic de Rochester, dans le Minnesota, aux États-Unis. Pendant plus d'une décennie, le format était la norme «de facto» pour le post-traitement de l'imagerie médicale [24].

L'avantage du format Analyze est qu'il a été conçu pour des données multidimensionnelles (volume). En effet, il est possible de stocker dans un fichier des données 3D ou 4D (la quatrième dimension étant typiquement l'information temporelle). Un volume Analyze 7.5 est constitué de deux fichiers binaires : un fichier image avec l'extension «.img» contenant les données brutes du voxel et un fichier d'en-tête portant l'extension «.hdr» contenant les métadonnées, telles que le nombre de pixels dans les pixels x, y, et les directions z, la taille du voxel et le type de données [24].

Nifti

Nifti est un format de fichier créé au début des années 2000 par un comité des National Institutes of Health dans le but de créer un format de neuroimagerie préservant les avantages du format Analyze, mais résolvant

les faiblesses. Le Nifti peut en fait être considéré comme un format d'analyse révisé. Le format remplit certains des champs inutilisés / peu utilisés présents dans l'en-tête Analyze 7.5 pour stocker de nouvelles informations telles que l'orientation de l'image dans le but d'éviter l'ambiguïté gauche-droite dans les études sur le cerveau. De plus, Nifti inclut la prise en charge du type de données non prévu dans le format d'Analyse comme le 16 bits non signé. Bien que le format permette également le stockage des données d'en-tête et de pixels dans des fichiers séparés, les images sont généralement enregistrées dans un fichier «.nii» unique dans lequel l'en-tête et les données de pixels sont fusionnés. L'en-tête a une taille de 348 octets dans le cas de stockage de données «.hdr» et «.img», et une taille de 352 octets dans le cas d'un fichier «.nii» unique pour la présence de quatre octets supplémentaires à la fin, essentiellement pour faire de la taille un multiple de 16, et aussi pour fournir un moyen de stocker des métadonnées supplémentaires, auquel cas ces 4 octets sont non nuls [24].

Le format Nifti a rapidement remplacé le format Analyze dans le domaine de "neuroimaging" en tant que format par défaut adopté par certains des progiciels du domaine public les plus répandus, tels que FSL, SPM et AFNI.

Une version de mise à jour de la norme, le Nifti-2, développée pour gérer un ensemble de données plus volumineux, a été définie en 2011. Cette nouvelle version code chacune des dimensions d'une matrice d'image avec un entier de 64 bits au lieu d'un fichier de 16 bits. dans le Nifti-1, en éliminant la limitation de la taille limite de 32 767. Cette version mise à jour conserve presque toutes les caractéristiques du Nifti-1 [24].

Minc

Le format de fichier Minc a été développé à l'Institut neurologique de Montréal (INM) à partir de 1992 pour fournir un format de données flexible pour l'imagerie médicale. La première version du format Minc (Minc1) était basée sur le format standard NetCDF (Network Common Data Format). Par la suite, pour dépasser la limite de prise en charge des fichiers de données volumineux et proposer de nouvelles fonctionnalités, l'équipe de développement de Minc a choisi de passer de NetCDF à Hierarchical Data Format version 5 (HDF5). Cette nouvelle version, incompatible avec la précédente,

s'appelait Minc2 [24]. Le format est principalement utilisé par les outils logiciels développés par le Centre d'imagerie cérébrale de l'INM, c'est-à-dire un visualiseur et une bibliothèque de logiciels de traitement. Un ensemble d'utilitaires permettant la conversion vers et à partir des formats Dicom et Nifti, ainsi qu'entre Minc1 et Minc2, ont été mis à disposition par le même groupe.

Dicom

La norme Dicom a été établie par l'American College of Radiology et la National Electric Manufacturers Association. Malgré sa date de naissance en 1993, la norme Dicom est réellement introduite dans les services d'imagerie à la fin des années 90. Aujourd'hui, le standard Dicom est la colonne vertébrale de tous les départements d'imagerie médicale. La valeur ajoutée de son adoption en termes d'accès, d'échange et de facilité d'utilisation des images médicales à des fins de diagnostic est en général énorme [24].

2.3 Apprentissage automatique

L'apprentissage automatique est un champ d'étude de l'intelligence artificielle qui se base sur des approches statistiques pour donner aux ordinateurs la capacité d'apprendre à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans explicitement programmés pour chacune [4]. Plus largement, cela concerne la conception, l'analyse, le développement et l'implémentation de telles méthodes. Dans l'apprentissage automatique on distingue trois types : supervisé, semi-supervisé et non supervisé

2.3.1 Apprentissage supervisé

Si les classes sont prédéterminées et les exemples connus, le système apprend à classer selon un modèle de classification ou de classement ; on parle alors d'apprentissage supervisé . Un expert doit préalablement étiqueter des exemples. Le processus se passe en deux phases. Lors de la première phase dite d'apprentissage, il s'agit de déterminer un modèle à partir des données étiquetées. La seconde phase en ligne, dite de test consiste à prédire l'étiquette d'une nouvelle donnée, connaissant le modèle préalablement appris [16].

Quelques algorithmes d'apprentissage supervisé

Réseaux de neurones : un réseau de neurones est un système dont la conception est à l'origine schématiquement inspirée du fonctionnement des neurones biologiques, et qui par la suite s'est rapproché des méthodes statistiques.

Les réseaux de neurones sont généralement optimisés par des méthodes d'apprentissage de type probabiliste. Ils sont placés d'une part dans la famille des applications statistiques, qu'ils enrichissent avec un ensemble de paradigmes permettant de créer des classifications rapides (réseaux de Kohonen en particulier), et d'autre part, dans la famille des méthodes de l'intelligence artificielle auxquelles ils fournissent un mécanisme perceptif indépendant des idées propres à celui qui l'implémente, et fournissant des informations d'entrée au raisonnement logique forme [11].

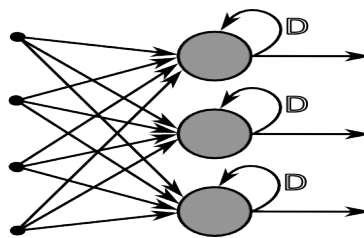


FIGURE 2.1 – Schéma d'un neurone

Machines à vecteurs de support(SVM) : Support Vector Machines, souvent traduit par l'appellation de Séparateurs à Vaste Marge (SVM) sont une classe d'algorithmes d'apprentissage initialement définis pour la discrimination c'est-à-dire la prévision d'une variable qualitative initialement binaire. Ils ont été ensuite généralisés à la prévision d'une variable quantitative. Dans le cas de la discrimination d'une variable dichotomique, ils sont basés sur la recherche de l'hyperplan de marge optimale qui, lorsque c'est possible, classe ou sépare correctement les données tout en étant le plus éloigné possible de toutes les observations. Le principe est donc de trouver un classificateur, ou une fonction de discrimination, dont la capacité de généralisation et la qualité de prévision sont les plus grandes possible [15].

2.3.2 Apprentissage par ensemble

Lorsqu'il faut prendre une décision importante, il vaut souvent mieux recueillir plusieurs avis que de se fier à un seul. Utiliser un modèle de machine learning pour prédire. Nous parlons donc de méthodes ensemblistes, qui sont bagging, boosting et Staking.

Boosting

Le boosting est une technique qui a des origines théoriques ancrées dans le modèle PAC. Dans ce cadre, un résultat fort a été montré : tout algorithme d'apprentissage produisant des hypothèses faisant mieux qu'une procédure de choix aléatoire peut être boostée jusqu'à atteindre une erreur aussi faible que voulue et ce, avec une probabilité aussi grande que désiré et en un temps polynomial [17].

Bagging

Cette technique concerne les apprenants dont l'erreur est due en plus grande partie à variance, c'est-à-dire des apprenants qui réagissent à de petites modifications de l'ensemble d'apprentissage en produisant des hypothèses différentes à chaque fois [17].

La solution proposée est justement d'utiliser plusieurs fois l'apprenant pour obtenir autant d'hypothèses et les combiner au sein d'un vote. Pour obtenir des hypothèses différentes, le bagging propose de construire de nouveaux échantillons à partir des données initiales.

Forêts aléatoires (Random Forests) Les forêts d'arbres décisionnels ou forêts aléatoires ont été formellement proposées en 2001 par Leo Breiman et Adèle Cutler. Elles font partie des techniques d'apprentissage automatique. Cet algorithme combine les concepts de sous-espaces aléatoires et de bagging. L'algorithme des forêts d'arbres décisionnels effectue un apprentissage sur de multiples arbres de décision entraînés sur des sous-ensembles de données légèrement différents.[26], et les directions z , la taille du voxel et le type de données [24].

Staking

De manière générale, le terme *stacking* (Wolpert, 1992) fait référence à toute méthode ensembliste consistant à combiner un ensemble d'hypothèses hétérogènes, en utilisant leurs réponses pour un exemple e_i . Ces réponses sont vues comme les attributs d'une nouvelle représentation de e_i , qui est utilisée par une autre méthode d'apprentissage pour produire une hypothèse finale [17].

2.3.3 Apprentissage profond

L'apprentissage profond est un sous-domaine de l'apprentissage automatique qui s'intéresse aux algorithmes inspirés par la structure et la fonction du cerveau, appelés réseaux de neurones artificiels. Elle implique un type particulier de modèle mathématique qui peut être considéré comme une composition de blocs simples d'un certain type, dans une structure multicouche et où certains de ces blocs peuvent être ajustés pour mieux prédire le résultat final.

Dans l'apprentissage en profondeur, un modèle informatique apprend à effectuer des tâches de classification directement à partir d'images, de texte ou de son. Les modèles d'apprentissage en profondeur peuvent atteindre une précision de pointe, dépassant parfois les performances d'un niveau humain. Les modèles sont formés en utilisant une grande masse de données étiquetées et des architectures de réseau neuronal qui contiennent de nombreuses couches.

Architectures de réseaux profonds

Il existe différents types d'architectures de Réseau de neurones convolutionnels tels que GoogleNet [28], Alexnet [22], ResNet [19], VGGNet [27], etc

AlexNet [22] : AlexNet est un réseau de neurones à convolution profonde conçu pour traiter les tâches complexes de classification de scènes sur des données Imagenet. Il s'agit de classer l'entrée donnée dans l'une des 1000 classes. Alexnet dispose de huit couches. Les cinq premières sont convolutives et les trois dernières sont des couches entièrement connectées. Entre les deux, nous avons également des «couches» appelées Pooling et Activation.

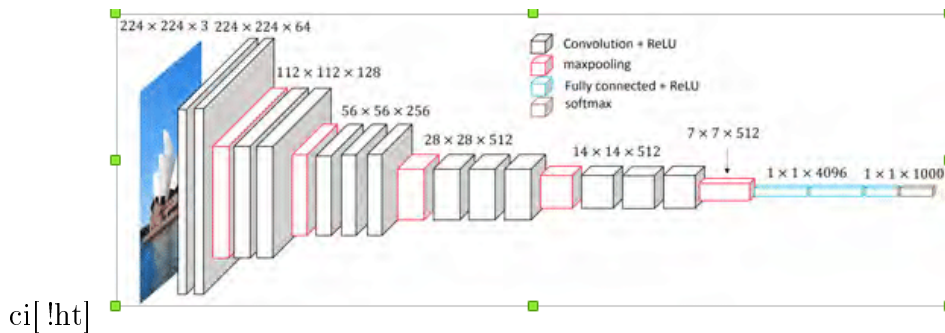


FIGURE 2.2 – L'architecture VGGNet

VGGNet [27] : VGGNet est un réseau de neurones qui a très bien fonctionné dans le cadre du défi de reconnaissance visuelle à grande échelle Image Net (ILSVRC) en 2014. Il a obtenu la première place pour la tâche de localisation des images et la deuxième place pour la tâche de classification des images. VGG-Net contient des couches convolutives ultérieures suivies par des couches de regroupement. Les couches de regroupement sont responsables de rendre les couches plus étroites. L'architecture de VGG-Net est illustrée dans la figure 2.2

Inception [28] : Inception est une architecture conçue par des chercheurs de Google. Elle contient 22 couches par rapport à VGG qui dispose de 19 couches. Les chercheurs ont également développé une nouvelle approche appelée le module Inception (figure 2.3). L'architecture finale contient plusieurs modules d'Inception, comme le montre la figure 2.4

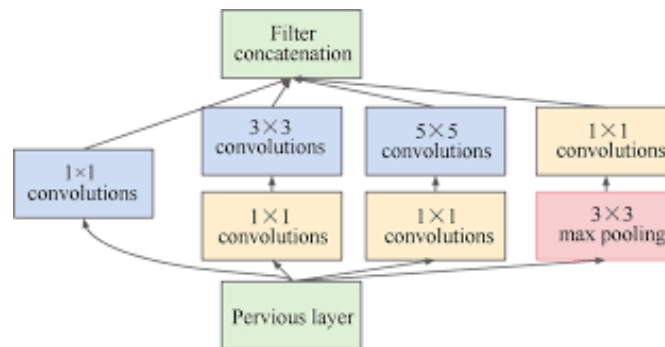


FIGURE 2.3 – Le module Inception

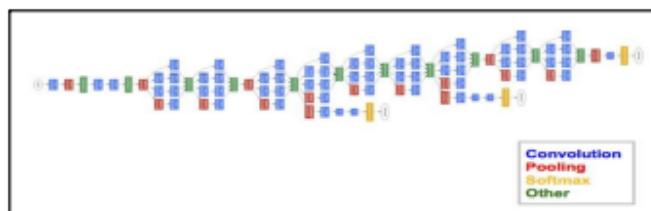


FIGURE 2.4 – L’architecture GoogleNet

ResNet [19] : ResNet est l’abréviation du nom composé RESidual NETwork (Réseaux résiduels). ResNet est l’une des architectures qui définissent véritablement la profondeur d’une architecture d’apprentissage en profondeur. Les ResNets sont constitués de plusieurs modules résiduels subséquents, qui constituent le bloc de construction de base de l’architecture ResNet. Ces modules résiduels sont empilés les uns sur les autres pour former un réseau complet. La figure 2.5 présente une représentation de l’architecture ResNet.

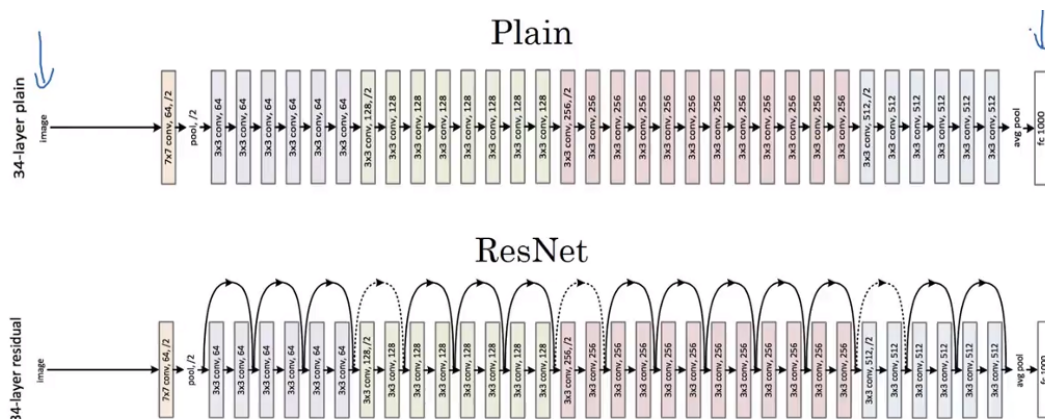


FIGURE 2.5 – l’architecture ResNet

2.3.4 Apprentissage par transfert

L’apprentissage par transfert est une méthode d’apprentissage automatique où un modèle développé pour une tâche est réutilisé comme point de départ pour un modèle dans une deuxième étape d’apprentissage.

L'apprentissage par transfert est une solution prometteuse au difficile problème de l'apprentissage des réseaux profonds au moyen de bases d'apprentissage de petite taille en présence d'une grande variabilité visuelle intra-classe.

L'apprentissage par transfert peut être vu comme la capacité d'un système à reconnaître et appliquer des connaissances et des compétences, apprises à partir de tâches antérieures, sur de nouvelles tâches ou domaines partageant des similitudes.

2.3.5 Mesures d'évaluation

Matrice de confusion

La matrice de confusion Donne un aperçu claire sur la qualité d'un système d'apprentissage/classification en précisant le nombre de décisions dans chaque classe. Chaque colonne de la matrice représente le nombre d'occurrences d'une classe prédite. Chaque ligne représente le nombre d'occurrences d'une classe réelle (ou de référence). Où :

		Classe prédite	
		Classe 0	Classe 1
Classe réelle	Classe 0	VN	FN
	Classe 1	FP	VP

FIGURE 2.6 – Matrice de confusion

- Vrai positif VP : Elément de la classe 1 correctement prédit ;
- Vrai négatif VN : Elément de la classe 0 correctement prédit ;
- Faux positif FP : Elément de la classe 1 mal prédit ;
- Faux négatif FN : Elément de la classe 0 mal prédit ;

Précision (Precision)

La précision est la proportion des exemples positifs (vrais classes) qui sont prédits positifs. C'est la proportion des VP par rapport aux exemples positifs dans l'ensemble de test.c'est une précision par rapport aux prédictions,et on peut le calculer à partir de la matrice de confusion par la formule :

$$Pr = \frac{\#VP}{\#VP + \#FP}$$

où : $\#VP$ et $\#FP$, dénotent le nombre des vrais positifs et des faux positifs, respectivement.

Rappel (Recall)

Le rappel est la proportion des exemples prédits correctement positifs par rapport aux exemples qui sont réellement positifs (vraies classes) dans le corpus de test. Mesure la capacité du modèle à extraire du corpus de test les exemples positifs. on peut le calculer à partir de la matrice de confusion par la formule :

$$Rp = \frac{\#VP}{\#VP + \#FN}$$

où : $\#VP$ et $\#FN$ dénotent le nombre des vrais positifs et des faux négatifs, respectivement.

Kappa de Cohen

kappa de Cohen est un coefficient destiné à mesurer l'accord entre deux variables qualitatives ayant les mêmes modalités. Le calcul de cette mesure se fait de la manière suivante :

$$K = \frac{P_a - P_e}{1 - P_e}$$

Où :

- K : la valeur de la mesure Kappa de Cohen ;
- P_a : l'accord relatif entre codeurs

$$P_a = \frac{1}{N}[\#VP + \#VN]$$

où : $\#VP$ et $\#VN$ dénotent le nombre des vrais positifs et des vrais négatifs, respectivement.

- P_e : la probabilité d'un accord aléatoire

$$P_e = \frac{1}{N^2}[(\#VP + \#FP) * (\#VP + \#FN) + (\#VN + \#FN) * (\#FN + \#FP)]$$

où : $\#VP$, $\#FP$, $\#VN$ et $\#FN$ dénotent le nombre des vrais positifs, des faux positifs, des vrais négatifs et des faux négatifs, respectivement.

- N : nombre d'unités statistiques (nombre d'échantillons) ;

F-mesure (F-measure)

F-mesure combine la précision et le rappel

$$F - mesure = \frac{2 * Pr * Rp}{Pr + Rp}$$

Erreur quadratique moyenne (Root Mean Square error)

L'erreur quadratique moyenne mesure l'ampleur de l'erreur entre deux ensembles de données. En d'autres termes, elle compare une valeur prédite et une valeur observée ou connue

$$RMSE = \sqrt{\frac{\sum_1^n (p_i - o_i)}{Kappa\ de\ Cohenn}}$$

Courbe ROC [8]

Une courbe ROC (Receiver Operating Characteristic) est un graphique représentant les performances d'un modèle de classification pour tous les seuils de classification. Cette courbe trace le taux de vrais positifs en fonction du taux de faux positifs :

- Le taux de vrais positifs (TVP) est l'équivalent du rappel ;
- Le taux de faux positifs (TFP) est défini comme suit :

$$TFP = \frac{\#FP}{\#FP + \#VN}$$

où $\#VN$ et $\#FP$, dénotent le nombre des vrais négatifs et des faux positifs, respectivement ;

Une courbe ROC trace les valeurs TVP et TFP pour différents seuils de classification. Diminuer la valeur du seuil de classification permet de classer plus d'éléments comme positifs, ce qui augmente le nombre de faux positifs et de vrais positifs.

L'aire sous la courbe ROC (AUC) [8]

AUC signifie "l'Aire sous la courbe ROC". Cette valeur mesure l'intégralité de l'aire à deux dimensions située sous l'ensemble de la courbe ROC (par calculs d'intégrales) de (0,0) à (1,1).

2.4 Utilisation de l'apprentissage machine pour l'analyse d'images

L'apprentissage machine permet la réalisation automatique de tâches essentielles en imagerie médicale comme par exemple : la détection des anomalies, la caractérisation des anomalies détectées, la segmentation d'organes, le suivi de l'évolution de certaines maladies (cancer par exemple). Nous présentons dans ce qui suit l'utilisation de l'apprentissage profond dans le domaine de l'imagerie médicale.

2.4.1 Apprentissage profond pour l'analyse d'images médicales

Les chercheurs estiment que les images médicales représentent actuellement au moins 90% de toutes les données médicales, ce qui en fait la plus importante source de données dans l'industrie des soins de santé. Cela devient une quantité écrasante à l'échelle humaine, quand on considère que les radiologues dans les salles d'urgence de certains hôpitaux reçoivent quotidiennement des milliers d'images. De nouvelles méthodes sont donc nécessaires pour extraire et représenter les données de ces images plus efficacement.

Bien que l'une des applications les plus courantes d'apprentissage machine dans le domaine médicale ait été en réalité l'imagerie médicale, ce n'est que récemment que des algorithmes d'apprentissage profond ont été introduits.

NiftyNet

NiftyNet est une plate-forme de réseaux de neurones convolutionnels (CNN) open source basée sur Tensorflow pour la recherche en analyse d'images médicales et en thérapie guidée par l'image. La structure modulaire de NiftyNet est conçue pour le partage de réseaux et de modèles pré-formés. En utilisant cette structure modulaire, vous pouvez commencer avec les réseaux pré-formés établis en utilisant des outils intégrés, adapter les réseaux existants à vos données d'imagerie, construisez rapidement de nouvelles solutions à vos propres problèmes d'analyse d'images. Niftynet a été appliqué dans

l'état de l'art, surtout dans le domaine de la segmentation des images médicales et a donné des résultats très intéressants.

2.5 Conclusion

Nous avons survolé dans ce chapitre les domaines de l'imagerie médicale et de l'apprentissage machine. Nous présenterons dans le prochain chapitre quelques travaux ayant exploité l'apprentissage automatique pour l'analyse sémantique des images médicales dans le cadre du diagnostic de la maladie de tuberculose pulmonaire.

Chapitre 3

Prédiction automatique du degré de sévérité de tuberculose pulmonaire

3.1 Introduction

Nous présentons dans ce chapitre les difficultés et les défis de l'analyse des images médicales ainsi que quelques travaux réalisés pour la prédiction automatique du degré de sévérité de la tuberculose pulmonaire dans le cadre de la campagne internationale ImageCLEF.

3.2 Difficulté et défis de l'analyse des images médicales

L'analyse sémantique des images médicales via des techniques d'apprentissage automatique présente plusieurs difficultés et défis. On peut en citer : la taille des images qui est très grande par rapport aux autres types d'images. De plus, comme les images médicales sont généralement en 3D, cela nécessite la projection sur les différentes dimensions. Par conséquent, la taille des données se verra encore plus grande. Rajoutons à cela, les différents problèmes connus dans le domaine de l'apprentissage automatique, surtout la difficulté de se procurer suffisamment de données afin de bien entraîner les modèles. Une autre difficulté concernent les caractéristiques des images médicales. En

effet, dans une image médicale, les informations décrivant la maladie sont restreintes par rapport à l'ensemble d'informations contenues dans l'image, ce qui entraîne beaucoup de bruits et cela affecte la performance des modèles. Ces difficultés sont confirmées lors des différents travaux réalisés dans plusieurs tâches, typiquement les tâches d'analyses d'images adoptées dans le cadre des campagnes d'évaluation, comme nous le verrons par la suite pour la campagne internationale ImageCLEF.

3.3 Compagne d'évaluation ImageCLEF

ImageCLEF¹ est une campagne lancée en 2003 dans le cadre du Forum d'évaluation multilingue (CLEF), a pour objectif de fournir une aide à l'évaluation de :

1. méthodes indépendantes de la langue pour l'annotation automatique d'images avec des concepts ;
2. méthodes de recherche d'informations multimodales combinaison de caractéristiques visuelles et textuelles ;
3. méthodes de récupération d'images multilingues, afin de comparer l'effet de la récupération des annotations d'images et des formulations de requêtes dans plusieurs langues

ImageCLEF a pour objectif de fournir un forum d'évaluation pour l'annotation et la récupération d'images multilingues. Motivé par la nécessité de prendre en charge les utilisateurs multilingues de la communauté mondiale accédant au corpus croissant d'informations visuelles, l'objectif principal d'ImageCLEF est de soutenir l'avancement du domaine de l'analyse, de l'indexation, de la classification et de la récupération des médias visuels, en développant les éléments nécessaires. infrastructure pour l'évaluation de systèmes de récupération d'informations visuelles fonctionnant dans des contextes monolingue, multilingue et indépendant du langage. ImageCLEF vise à fournir des ressources réutilisables à ces fins d'analyse comparative.

Cette organisation présente de nombreux défis, et ces défis sont limités dans des tâches proposés par l'organisation ImageCLEF 2017 [12], ImageCLEF 2018 [21] et ImageCLEF 2019 [3].

1. <https://www.imageclef.org/>

3.3.1 ImageCLEF 2017

Pour l'édition 2017, ImageCLEF a organisé trois tâches principales avec un objectif global d'analyse comparative, de la récupération et de la synthèse des processus de prise de vie, de la prédiction du type de tuberculose à partir d'images tomodensitométriques et de la prédiction de légendes d'images biomédicales ; et une tâche pilote d'analyse d'images de télédétection.

3.3.2 ImageCLEF 2018

En 2018, ImageCLEF a organisé trois tâches principales avec un objectif global d'analyse comparative de synthèse et de récupération de liens de vie, la détection de concepts dans des images biomédicales, la prédiction de légendes et la prédiction du score de gravité de la tuberculose à partir d'images tomodensitométriques ; et une tâche pilote de réponse à une question visuelle médicale.

3.3.3 ImageCLEF 2019 [20]

Pour l'édition 2019, ImageCLEF organise quatre tâches principales avec un objectif global de promotion de l'évaluation des technologies pour l'annotation, l'indexation et la récupération de données visuelles dans le but de fournir un accès à l'information dans de grandes collections d'images dans différents scénarios d'utilisation et domaines : médecine, nature et sécurité.

3.4 Tâche Tuberculosis

3.4.1 Organisation de la tâche

Dans cette tâche pour l'année 2019 il ya 2 sous tâches : CTR - CT report, SVR - Severity scoring. Les deux sous-tâches (SVR et CTR) utilisent le même ensemble de données contenant 335 scanners thoraciques de patients atteints de tuberculose, ainsi qu'un ensemble de métadonnées pertinentes sur le plan clinique. 218 patients sont utilisés pour la formation et 117 pour le test.

Pour tous les patients, il existe des images de tomodensitométrie 3D avec une taille d'image par tranche de 512 * 512 pixels et un nombre de tranches variant d'environ 50 à 400. Toutes les images de tomographie sont stockées au format de fichier NIFTI avec l'extension de fichier .nii.gz (g- fichiers .nii compressés).

SVR - Prédiction du score de gravité (SVR - Severity scoring) [3]

Cette sous-tâche vise à évaluer le score de gravité de la tuberculose. Le score de gravité est un score cumulatif de gravité du cas de TBC attribué par un médecin. À l'origine, le score variait de 1 ("critique / très mauvais") à 5 ("très bien").

Le but de cette sous-tâche est d'évaluer la gravité en fonction de l'image tomodensitométrique et de quelques méta-données supplémentaires, notamment l'incapacité, la rechute, la comorbidité, la bacillothérapie et le tabagisme. Le score de gravité initial est inclus dans les méta-données d'entraînement, mais le score final que les participants doivent évaluer est réduit à une catégorie binaire : "LOW" (scores 4 et 5) et "HIGH" (scores 1, 2 et 3).

3.4.2 Données

ImageCLEF fournit des images tomodensitométriques 3D avec une taille d'image par tranche de 512 * 512 pixels et un nombre de tranches compris entre 50 et 400 environ. Toutes les images tomodensitométriques sont stockées au format de fichier NIFTI avec l'extension de fichier .nii.gz (g- fichiers .nii compressés).

Les deux sous-tâches (SVR et CTR) utilisent le même ensemble de données contenant 335 tomodensitogrammes thoraciques de patients atteints de tuberculose, ainsi qu'un ensemble de métadonnées pertinentes sur le plan clinique. 218 patients sont utilisés pour la formation et 117 pour le test.

3.4.3 Mesures d'évaluation

Les résultats sont évalués en considérant cette sous-tâche comme un problème de classification binaire et un problème de régression. Le problème de

classification est évalué à l'aide de courbes ROC produites à partir des probabilités fournies par les participants. Pour le problème de régression, l'erreur quadratique moyenne est utilisée [13].

3.5 Quelques travaux de prédiction du degré de sévérité de la tuberculose pulmonaire

3.5.1 Travail de UIIP _ BioMed [25]

Dans ce travail, les lésions tuberculeuses ont été étiquetées manuellement sur un total de 198 tomodensitogrammes 3D. La deuxième étape visait à corriger le marquage initial des lésions et à effectuer une segmentation plus précise des lésions. Les deux étapes de l'étiquetage ont été réalisées à l'aide d'un logiciel auxiliaire conçu par l'équipe.

L'équipe a été classée première dans la tâche avec une valeur de RMSE égale à 0.7840 et une valeur de AUC égale à 0.7025 (voir le tableau 3.1).

3.5.2 Travail de MEDGIFT [14]

Les auteurs de ce travail ont d'abord créé les images 3D et les masques pulmonaires isométriques. Après avoir analysé les résolutions multiples et les distances inter-coupes dans le jeu de données, ils ont opté pour une taille de voxel de 1 mm pour capturer un maximum des informations disponibles.

L'équipe a été classé deuxième dans la tâche avec une valeur de RMSE égale à 0.8513 et une valeur de AUC estimée à 0.7162 (voir le tableau 3.1).

3.5.3 Travail de MostaganemFSEI [18]

L'équipe MostaganemFSEI a utilisé une approche qui consiste à extraire un seul descripteur sémantique unique pour chaque patient au lieu de considérer toutes les coupes comme des échantillons séparés.

3.6 Résultats de prédiction du degré de sévérité de la tuberculose pulmonaire

Le tableau suivant montre les résultats obtenus pour la sous-tâche de classification de degré de sévérité de tuberculose pulmonaire pour l'année 2018. L'évaluation a été réalisée sur l'ensemble d'images de Test en utilisant l'Erreur quadratique moyenne comme mesure d'évaluation .

GroupName	Run	RMSE	Rank_RMSE	AUC	Rank_AUC
UIIP_BioMed	SVR_run_TBdescs2_zparts3_thrprob50_rf100.csv	0.7840	1	0.7025	6
MedGIFT	SVR_HOG_std_euclidean_TST.csv	0.8513	2	0.7162	5
VISTA@UEvora	SVR-Run-07-Mohan-MLP-6FTT100.txt	0.8883	3	0.6239	21
MostaganemFSEI	SVR_mostaganemFSEI_run3.txt	0.9721	12	0.5987	25
VISTA@UEvora	SVR-Run-03-Mohan-MLP.txt	0.9768	13	0.6620	13

TABLE 3.1 – Résultats de la prédiction du degré de sévérité de tuberculose pulmonaire dans la tâche Tuberculosis de ImageCLEF 2018

Chapitre 4

Contribution et résultats

4.1 Introduction

Dans cette section, nous présentons notre contribution en détaillant l'approche utilisée. Nous décrivons tout d'abord la tâche "Score de gravité de la tuberculose (SVR)" du défi international Imageclef 2019 [20]. Nous détaillerons également les expérimentations réalisées ainsi qu'une discussion des résultats d'évaluation obtenus.

4.2 Contribution

Dans notre travail nous avons suivi plusieurs étapes pour atteindre notre but qui est la prédiction automatique du degré de sévérité de tuberculose pulmonaire. Nous proposons un système qui inclut les étapes suivantes que nous détaillons dans les sections suivantes :

1. extraction des slices
2. filtrage des données
3. apprentissage d'un modèle profond de classification

Le système est décrit dans la figure 4.1.

4.2.1 Extractions des slices

Dans cette partie nous extrayons les images de format JPG à partir des images 3D (format nii) des 3 dimensions.

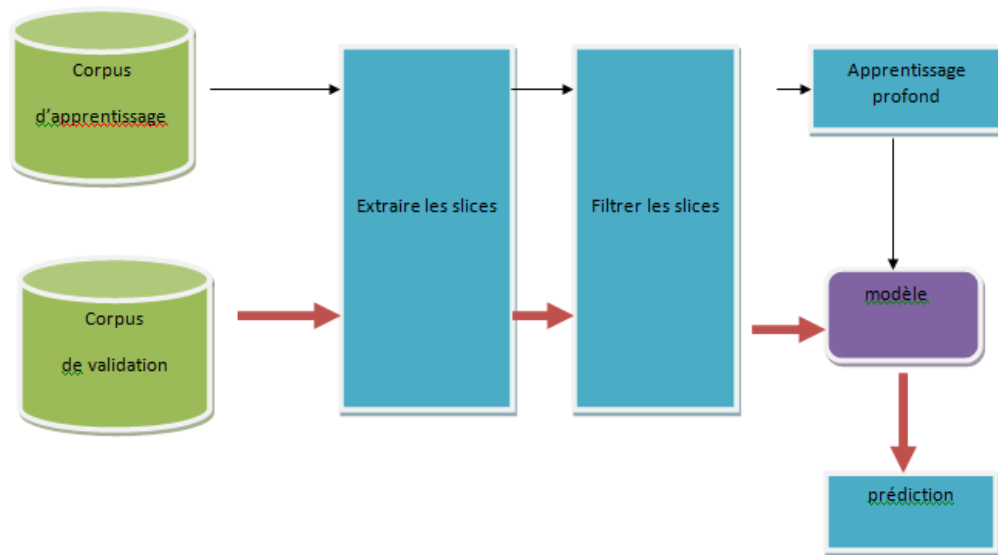


FIGURE 4.1 – Système développé

4.2.2 Filtrage

Dans cette partie nous avons créé un groupe des slices qui sont susceptibles de contenir l'information pertinente relative à la présence de la tuberculose. Cela est réalisé via l'intervention d'un expert médical.

Ce groupe est après exploité pour filtrer les slices des patients. En effet, pour chaque patient on propose de ne garder que les N images les plus similaires au groupe de filtrage. Pour cela, on calcule une distance moyenne entre chaque image et l'ensemble des slices du groupe, pour ne garder à la fin que les N images relatives aux N plus petites distances, n'importe quelle mesure de distance entre images peut être utilisée.

Dans notre travail, nous avons utilisé la distance "ssim" (Structural Similarity Index) et $N = 50$.

4.2.3 Apprentissage

Nous avons choisi d’entraîner un modèle profond exploitant notre corpus de données afin de prédire le degré de sévérité des cas tuberculeux étudiés. Nous avons testé l’architecture “*InceptionV3*” (voir la section 2.3.3).

4.3 Expérimentation et résultats

Nous avons utilisé dans nos expérimentations la collection de données fournies par le comité d’organisation de imageClef 2019 qu’on a subdivisé en deux parties : une pour l’apprentissage et une autre pour la validation. Le tableau 4.1 décrit cet ensemble de données.

Apprentissage	Validation	Totale
109	109	218

TABLE 4.1 – Composition des collections d’apprentissage et validation de corpus

Premièrement, nous avons décompressé les fichiers (images format nifty) et nous avons extrait les slices. Le résultat de cette opération a donné trois ensembles de slices correspondant aux trois dimensions de l’image 3D. Le nombre de slices pour chaque image Nifty est d’environ 50 – 400 images. Nous avons choisi de travailler sur les slices de la dimension Z , parce que des expérimentations nous ont montré que c’est la dimension qui donne les meilleures résultats avec notre approche. Ensuite, nous avons filtré les slices comme décrit dans la section 4.2.2.

Nous avons choisi de générer un modèle profond qui utilise notre ensemble de données pour prédire la gravité des cas de tuberculose. Nous rappelons que le problème est modélisé sous forme d’une classification binaire où les classes sont “High” et “Low”. Nous avons entraîné après un modèle profond en choisissant l’architecture “*InceptionV3*”.

Nous avons évalué notre modèle sur la collection de validation. Nous avons obtenu la matrice de confusion décrite dans le tableau 4.2.

Le tableau 4.3 montre les résultats en termes de précision moyenne (Accuracy) et l'aire sous la courbe ROC (AUC) obtenus par notre modèle entraîné et l'architecture "Resnet50".

Le modèle "*InceptionV3*" atteint une précision moyenne d'environ 53.59% contre 63% pour "Resnet50". Bien que le modèle *resnet50* est plus performant en termes de précision moyenne, l'architecture *InceptionV3* a obtenu de meilleures performances en termes de AUC comme montré dans le tableau 4.3.

	low	high
low	71	7
high	64	11

TABLE 4.2 – Matrice de confusion obtenue par le modèle *InceptionV3* sur la collection de validation.

	Accuracy	AUC
<i>Resnet50</i>	0.6300	0.3400
<i>InceptionV3</i>	0.5359	0.4933

TABLE 4.3 – Résultats en termes de précision moyenne(Accuracy) et de l'aire sous la courbe ROC (AUC) obtenus sur la collection de validation.

4.4 Conclusion

Dans ce chapitre, nous avons décrit notre contribution dans le cadre de ce projet. Nous avons proposé d'entraîner un réseau de neurones profond en exploitant un ensemble de slices filtrés de la dimension Z . Nous avons testé et comparé la performance de l'architecture *InceptionV3* avec *Resnet50*. Les résultats ont montré que ce dernier est plus performant en termes de précision moyenne alors que l'architecture *InceptionV3* est meilleure en termes de l'aire sous la courbe ROC.

Conclusion générale

Ce rapport présente aborde la problématique de l'analyse sémantique et automatique des images médicales de type tomodensitométrie dans le but de prédire automatiquement le degré de sévérité de la maladie de tuberculose pulmonaire.

Dans un premier temps, nous avons mené une étude bibliographique sur la tuberculose pour avoir une idée claire sur cette maladie et les différences entre les cinq types de la tuberculose pulmonaire. Ensuite, nous avons présenté le concept de l'apprentissage profond et ses architectures. Nous avons également survolé le domaine de l'imagerie médicale en mettant l'accent sur quelques applications de l'apprentissage profond dans le domaine de l'imagerie médicale. Nous avons exploré les travaux qui ont été effectués dans le cadre du défi international ImageCLEF2019 [20] sur la prédiction automatique du degré de sévérité de la maladie de tuberculose pulmonaire.

Nous avons proposé d'entraîner un réseau de neurones profond comme modèle de prédiction du degré de sévérité des cas tuberculeux. Nous avons testé l'architecture *InceptionV3*.

Comme perspectives, nous projetons à tester d'autres architectures. Sélectionner des données et faire des efforts dans le prétraitement des images avant l'apprentissage constitue également une bonne piste à explorer. L'utilisation des masques pour se focaliser uniquement les parties des images relatives aux deux poumons fait également partie de nos perspectives.

Bibliographie

- [1] tuberculose-miliaire <https://womenhealthnetfr.netlify.com/les-maladies-infecti21/fibreuse-tuberculose4071>.
- [2] Educalingo. <https://educalingo.com/fr/dic-fr/granulome>.
- [3] Imageclef 2019. <https://www.imageclef.org/2019>.
- [4] Inserm. <https://www.fr.zimple3d.com/flat/algorithm/>.
- [5] Inserm domaine de recherche(imagerie medicale). <https://histoire.inserm.fr/les-domaines-de-recherche/imagerie-medicale>.
- [6] Larousse encyclopedie. <https://www.larousse.fr/encyclopedie/medical/tuberculose/16743>.
- [7] Lien etudier : (tuberculose pulmonaire). <https://www.etudier.com/dissertations/La-Tuberculose/49134468.html>.
- [8] Machine learning crash-course-classification. <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=fr>.
- [9] Medisite tuberculose-miliaire definition. <https://www.medisite.fr/maladies-tuberculose-miliaire-la-definition.1147433.42.html>.
- [10] passeport santé tuberculose. https://www.passeportsante.net/fr/Maux/Problemes/Fiche.aspx?doc=tuberculose_pm.
- [11] rosemees ltd reseau de neurones artificiels definition. <https://rosemees.com/reseau-de-neurones-artificiels-definition/>.
- [12] Yashin Dicente Cid, Alexander Kalinovsky, Vitali Liauchuk, Vassili Kovalev, , and Henning Müller. Overview of ImageCLEFtuberculosis 2017 - predicting tuberculosis type and drug resistances. In *CLEF2017 Working Notes*, CEUR Workshop Proceedings, Dublin, Ireland, September 11-14 2017. CEUR-WS.org <<http://ceur-ws.org>>.

- [13] Yashin Dicente Cid, Vitali Liauchuk, Vassili Kovalev, , and Henning Müller. Overview of ImageCLEF tuberculosis 2018 - detecting multi-drug resistance, classifying tuberculosis type, and assessing severity score. In *CLEF2018 Working Notes*, CEUR Workshop Proceedings, Avignon, France, September 10-14 2018. CEUR-WS.org <<http://ceur-ws.org>>.
- [14] Yashin Dicente Cid and Henning Müller. Texture-based graph model of the lungs for drug resistance detection, tuberculosis type classification, and severity scoring : Participation in the imageclef 2018 tuberculosis task. 09 2018.
- [15] Theodoros Evgeniou and Massimiliano Pontil. Support vector machines : Theory and applications. volume 2049, pages 249–257, 01 2001.
- [16] Emmanuel Grolleau. Introduction à l'apprentissage machine. *Emmanuel Grolleau Observatoire de Paris – LESIA – Service d'Informatique Scientifique*.
- [17] Aurlien Gron. *Hands-On Machine Learning with Scikit-Learn and TensorFlow : Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Inc., 1st edition, 2017.
- [18] Abdelkader Hamadi and Yagoub Djamel Eddine. Imageclef 2018 : Semantic descriptors for tuberculosis et image classification. 09 2018.
- [19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *arXiv preprint arXiv :1512.03385*, 2015.
- [20] Bogdan Ionescu, Henning Müller, Renaud Péteri, Yashin Dicente Cid, Vitali Liauchuk, Vassili Kovalev, Dzmitri Klimuk, Aleh Tarasau, Asma Ben Abacha, Sadid A. Hasan, Vivek Datla, Joey Liu, Dina Demner-Fushman, Duc-Tien Dang-Nguyen, Luca Piras, Michael Riegler, Minh-Triet Tran, Mathias Lux, Cathal Gurrin, Obioma Pelka, Christoph M. Friedrich, Alba García Seco de Herrera, Narciso Garcia, Ergina Kavallieratou, Carlos Roberto del Blanco, Carlos Cuevas Rodríguez, Nikos Vasilopoulos, Konstantinos Karampidis, Jon Chamberlain, Adrian Clark, and Antonio Campello. ImageCLEF 2019 : Multimedia retrieval in medicine, lifelogging, security and nature. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction*, Proceedings of the 10th International Conference of the CLEF Association (CLEF 2019), Lugano, Switzerland, September 9-12 2019. LNCS Lecture Notes in Computer Science, Springer.

- [21] Bogdan Ionescu, Henning Müller, Mauricio Villegas, Alba García Seco de Herrera, Carsten Eickhoff, Vincent Andrearczyk, Yashin Dicente Cid, Vitali Liauchuk, Vassili Kovalev, Sadid A. Hasan, Yuan Ling, Oladimeji Farri, Joey Liu, Matthew Lungren, Duc-Tien Dang-Nguyen, Luca Piras, Michael Riegler, Liting Zhou, Mathias Lux, and Cathal Gurrin. Overview of ImageCLEF 2018 : Challenges, datasets and evaluation. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction*, Proceedings of the Ninth International Conference of the CLEF Association (CLEF 2018), Avignon, France, September 10-14 2018. LNCS Lecture Notes in Computer Science, Springer.
- [22] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Commun. ACM*, 60(6) :84–90, May 2017.
- [23] Paras Lakhani and Baskaran Sundaram. Deep learning at chest radiography : automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, 284, no 2 :574–582, 2017.
- [24] Michele Larobina and Loredana Murino. Medical image file formats. *J Digit Imaging*, 27(2) :200–206, Apr 2014. 24338090[pmid].
- [25] Vitali Liauchuk, Aleh Tarasau, Eduard Snezhko, and Vassili Kovalev. Imageclef 2018 : Lesion-based tb-descriptor for ct image analysis. 09 2018.
- [26] Robert Nisbet, John Elder, and Garreseau-de-neurones-artificiels-definitiony Miner. *Handbook of Statistical Analysis and Data Mining Applications*. Academic Press, Inc., Orlando, FL, USA, 2009.
- [27] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.
- [28] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Computer Vision and Pattern Recognition (CVPR)*, 2015.