



Faculté des Sciences
Département d'Informatique

Domaine : Mathématiques Informatique
Filière : Informatique

Mémoire de fin d'études en vu de l'obtention du
Diplôme de Master Académique

**Identification automatique de la
modalité des images médicales et
d'organes anatomiques par
apprentissage profond par transfert**

Présenté par :
M. AMMAR AOUCHE Chakib
M. BELLOUCHE Youcef

Soutenu le 08/07/2023, Devant le jury composé de :

Pr. BOULIF Menouer - Président
Dr. BENZENATI Tayeb - Examinateur
Pr. GACEB Djamel - Encadreur

Dédicace

“

Je dédie ce travail à :

À ma chère mère Fatiha,

Et à mon cher père Zine el Abdine,

Qui ont été à mes côtés à chaque étape, Et m'ont motivé à atteindre de nombreux objectifs.

À mon chers frère Adem et à ma chère sœur Maria, Pour leurs conseils tout au long de mon parcours.

À tous mes professeurs qui m'ont accompagné au fil de mes années universitaires, Et pour tout leur soutien apporté.

À tous mes chers amis et à mon binôme Youcef, Pour leur aide et leur soutien.

”

- Chakib

“

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

Je dédie ce travail :

À la mémoire de mon défunt père.

À ma chère mère

Qui ont été à mes côtés à chaque étape, Et m'ont motivé à atteindre de nombreux objectifs.

À mes chers frères et ma seule sœur , Pour leurs conseils tout au long de mon parcours, leurs encouragements permanents, et leur soutien moral.

À tous mes chers amis et à mon binôme Chakib, Pour leur aide et leur soutien.

À Monsieur Gaceb Djamel pour pour m'avoir montré le chemin pour apprendre et découvrir de nouvelles choses dans la vie.

”

- Youcef

Remerciements

Tout d'abord, nous tenons à remercier Dieu de nous avoir donné le courage et la patience nécessaires pour mener à bien ce travail.

Nous souhaitons exprimer nos sincères remerciements à nos chers parents pour leur soutien moral et leurs encouragements.

Nous adressons nos sincères remerciements au Pr. Djamel Gaceb qui nous a confié ce sujet et encadré durant ce stage. L'intérêt qu'il a manifesté pour notre travail, sa bienveillance, sa rigueur scientifique et sa grandes qualités humaines ont été d'une aide précieuse et nous a permis de mener à bien ce travail.

Enfin, nous remercions Pr. Boulif et Dr. Benzenati d'avoir accepté de faire partie de notre jury, ainsi que toutes les personnes qui ont participé, de près ou de loin, au développement de ce projet de fin d'études.

Résumé

L'objectif de ce travail de notre projet de fin d'études est de développer un modèle d'apprentissage profond par transfert, plus précisément un CNN, en comparant différentes architectures existantes pour la classification des modalités d'images médicales et d'organes anatomiques. La diversification de ces modalités joue un rôle essentiel dans le diagnostic, le suivi des traitements et la planification des interventions chirurgicales. Cependant, cette diversification de modalité associée à la diversité des organes anatomiques ciblés, complique la tâche d'automatisation de l'analyse avec un traitement adapté et précis sur des grands volumes d'images. Il est donc essentiel de faciliter cette tâche en faisant appel à l'intelligence artificielle en général, et à l'apprentissage profond en particulier .

Dans ce contexte, nous avons entraîné sur notre base d'images médicales six modèles CNN les plus populaires avec une performance déjà prouvée dans le domaine médical. Plusieurs approches basées sur l'apprentissage profond par transfert ont été élaborée pour la catégorisation des images médicales selon leur modalité d'acquisition ou le type d'organe anatomique ciblé. Ceci vise également à réduire la grande charge et à faciliter le travail des spécialistes de la santé dans leurs tâches quotidiennes en leur faisons gagner du temps précieux. Ceci complète et enrichi la liste des rares travaux de recherche déjà investis dans cette directions. Les résultats expérimentaux obtenus sont encourageants et intéressants pour les travaux d'un projet PRFU en cours qui cible le développement de nouvelles technologies d'IA pour l'aide au diagnostic des maladies dangereuses par imagerie médicale.

Mots clés : Intelligence artificiel, Apprentissage automatique, Apprentissage profond, Vision par ordinateur, Classification, Imagerie médical, modalité d'imagerie médical, Organe anatomique.

Abstract

The objective of this work of our graduation project is to develop a model of deep learning by transfer, more precisely a CNN, by comparing different existing architectures for classification modalities of medical images and anatomical organs. The diversification of these methods plays a role essential in the diagnosis, the follow-up of treatments and the planning of surgical interventions. However, this modality diversification associated with the diversity of targeted anatomical organs, complicates the task of automating analysis with adapted and precise processing on large volumes of images. It is therefore essential to facilitate this task by using artificial intelligence in general, and deep learning in particular.

In this context, we trained on our medical image database six most popular CNN models with already proven performance in the medical field. Several approaches based on deep transfer learning have been developed for the categorization of medical images according to their acquisition modality or the type of anatomical organ targeted. This also aims to reduce the heavy load and facilitate the work of healthcare professionals in their daily tasks by saving them valuable time. This completes and enriches the list of rare research work already invested in this direction. The experimental results obtained are encouraging and interesting for the work of an ongoing PRFU project which targets the development of new AI technologies to aid in the diagnosis of dangerous diseases by medical imaging.

Table des matières

Dédicace	I
Remerciements	III
Résumé	IV
Introduction générale	1
1 Notions de base en imagerie médicale et intelligence artificielle	4
1.1 Introduction	5
1.2 Image	5
1.3 Les Types de formats d'image	5
1.4 Modes de représentation d'images	7
1.5 Caractéristiques d'une image	7
1.6 Imagerie médicale	8
1.6.1 Acquisition d'images médicales	8
1.7 Intelligence artificielle et Imagerie médicale	11
1.7.1 Apprentissage automatique	11
1.7.2 Apprentissage profond	14
1.7.3 L'apprentissage profond par transfert	14
1.7.4 Réseau neuronal convolutif (CNN)	15
1.7.5 Les architectures CNN les plus populaires	19
1.7.6 Base d'images (Dataset)	21
1.8 Conclusion	22
2 Détection de modalités et d'organes anatomique médicale par apprentissage profond	23
2.1 Introduction	24
2.2 Application de l'apprentissage profond en imagerie médicale	24
2.3 Application de l'apprentissage profond pour la classification d'organe anatomique	25
2.4 Comparatif des approches existantes de classification d'organe anatomique	26
2.5 Application de l'apprentissage profond pour la classification de modalités d'imagerie médicale	26
2.6 Comparatif des approches existantes de classification des modalités d'imagerie médicale	28
2.7 Conclusion	29
3 Nos Approches	31

Table des matières

3.1	Introduction	32
3.2	Méthodologie	32
3.3	Approches utilisés	32
3.3.1	Classification des modalités d'imagerie médicale	33
3.3.2	Classification des organes anatomiques	34
3.4	Augmentation des données	34
3.5	Conclusion	35
4	Expérimentation et Résultats	37
4.1	Introduction	38
4.2	Outils utilisés	38
4.2.1	Python	38
4.2.2	TensorFlow	39
4.2.3	Keras	39
4.2.4	Environnements de développement	40
4.3	Métriques d'évaluation des modèles de classification	41
4.3.1	Matrice de confusion	41
4.3.2	Accuracy	42
4.3.3	Précision	42
4.3.4	Rappel	42
4.3.5	F1-score	43
4.4	Base d'images utilisées	43
4.5	Expérimentations et résultats	48
4.5.1	Expérimentations et résultats de la classification des modalité d'imagerie médicale	48
4.5.2	Expérimentations et résultats de la classification des organes anatomique	54
4.6	Discussion des résultats	54
4.7	Conclusion	55
	Conclusion générale	57
	Bibliographie	59

Table des figures

1.1	Image médicale neurologique acquis à partir d'IRM [8]	8
1.2	Image médicale acquis à partir d'une Échographie [8]	9
1.3	Radiographie pulmonaire [16]	9
1.4	Image médicale microscopique de la maladie de Huntington (une pathologie neurodégénérative rare et héréditaire. [15])	10
1.5	Image médicale acquis avec la médecine nucléaire. [17]	11
1.6	Les champs de l'intelligence artificiel [8]	11
1.7	Réseau de neurones [8]	13
1.8	Architecture d'un CNN	16
1.9	L'opération de convolution	16
1.10	L'opération de Max pooling	17
1.11	L'opération de flattening	18
2.1	Les hypers paramètre du modèle proposé [21]	25
2.2	Les résultats des différents CNN [23]	26
2.3	Les résultats de l'étude de Yuhai et al [24]	27
2.4	Architecture de Φ -Net [28]	28
3.1	Schéma synoptique de l'approche 3 basée sur la combinaison des CNNs	34
3.2	Exemples d'images transformées par les méthodes d'augmentation de données [41]	35
4.1	Logo de python	38
4.2	Logo de TensorFlow	39
4.3	Logo de Keras	40
4.4	Matrice de confusion	41
4.5	Image médicale du cerveau acquis par IRM de dataset utilisée [30]	44
4.6	Image médicale du torse acquis par Rayon X de dataset utilisée [33]	45
4.7	Image médicale de main acquis acquis par Rayon X de dataset utilisée [35]	46
4.8	Image médicale de poitrine de femme par Ultrasound de dataset utilisée [36]	46
4.9	Image médicale de la tête de fœtus par Ultrasound de dataset utilisée [37]	47
4.10	Image médicale du torse acquis par CT scan [38]	47
4.11	Image médicale du rein acquis par CT scan [39]	48
4.12	Organigramme du modèle VGG 16 basé sur les différents niveaux de Fine-tuning	50
4.13	Organigramme du modèle VGG 19 basé sur les différents niveaux de Fine-tuning	51
4.14	Organigramme du modèle Resnet 50 basé sur le niveau de Fine-tuning	52
4.15	Organigramme du modèle Inception 3	52

Liste des tableaux

2.1	Tableau comparatif des différentes approches pour la classification d'organes anatomique	26
2.2	Tableau comparatif des différentes approches pour la classification des modalités d'imagerie médicale	29
4.1	Tableau de dataset RSNA-MICCAI Brain Tumor Radiogenomic Classification	45
4.2	Tableau comparatif des résultats des différents modèles en mode extraction pour la classification des modalités	49
4.3	Tableau comparatif des résultats des différents modèles en mode fine tuning pour la classification des modalités	53
4.4	Tableau comparatif des résultats des différents modèles en mode fine tuning pour la classification des modalités et combinaison des modèles	53
4.5	Tableau comparatif des résultats des différents modèles en mode extraction pour la classification des organes anatomiques	54

Liste des sigles et acronymes

JPEG *Joint Photographic Experts Group*

PNG *Portable Network Graphics*

GIF *Graphics Interchange Format*

TIF *Tagged Image File*

SVG *Scalable Vector Graphics*

IA *Intelligence artificiel*

ML *Machine learning*

DL *Deep learning*

TL *Transfer learning*

CNN *Convolutional neural network*

DT *Dataset*

FC *Fully-Connected*

SVM *Support Vector Machine*

RF *Random Forest*

IRM *Imagerie par résonance magnétique*

Introduction générale

Introduction générale

La médecine est une science qui a pour but de guérir et de préserver le fonctionnement normal du corps humain, des organes qui le constituent ainsi que de la santé physique et psychique des individus, et cela grâce à la bonne pratique du diagnostic, du traitement et de la prévention contre les maladies.

Le domaine de l'imagerie médicale a connu une évolution spectaculaire ces dernières décennies, offrant des opportunités sans précédent pour le diagnostic précoce et précis de diverses pathologies. Cependant, l'interprétation manuelle des images médicales et la localisation des organes anatomiques demeurent des tâches complexes et chronophages pour les radiologues et les cliniciens. Dans ce contexte, l'utilisation de techniques d'apprentissage automatique, et en particulier de l'apprentissage profond, s'est avérée prometteuse pour améliorer l'efficacité et la précision de ces processus.

Aujourd'hui, la communauté scientifique utilise des algorithmes d'apprentissage profond pour améliorer et aider les médecins dans le travail. Ces algorithmes offrent une caractérisation automatique plus pertinente et sont capables d'apprendre en développant de larges connaissances sur une base d'images de grand volume. La possibilité de transférer l'apprentissage d'une manière incrémentale et évolutive offre un grand avantage à ces algorithmes à traiter des tâches de reconnaissance, de prédiction ou de classification avec des meilleures précisions. Ces propriétés, sont particulièrement intéressant dans le domaine médicale qui est très exigeant en terme de précision sur des datasets, qui sont souvent limités. En exploitant des modèles déjà entraînés sur des tâches de classification d'images, nous pouvons capitaliser sur les caractéristiques visuelles apprises pour aider à l'identification automatique des modalités d'images médicales et des organes anatomiques.

Un grand nombre de méthodes d'apprentissage profond utilise des réseaux de neurones profonds à convolution (CNN). Elles sont appliquées en analyse d'images médicales avec succès donnant des résultats prometteurs. Le domaine d'application couvre tout le spectre de l'analyse d'images médicales, y compris la détection, la segmentation, la classification et le diagnostic assisté par ordinateur.

La classification des modalités d'imagerie médicale et d'organes anatomiques facilite le travail de recherche qui est fait par les médecins et professionnels du secteur médical, en connaissant la modalité d'image, les cliniciens peuvent interpréter correctement l'image et prendre des décisions médicales précises. Par exemple, elle permet aux cliniciens de trouver rapidement les images médicales dont ils ont besoin pour effectuer des diagnostics, planifier des traitements et suivre la progression des maladies.

C'est dans cette optique, que cette étude a été proposée ou nous comparons des modèles CNN d'apprentissage profond par transfert pour la classification et l'identification des modalités d'imagerie médicale et des organes anatomiques sur une base d'images créées à partir de plusieurs sources. Pour cela, différents modèles CNN pré-entraînés sont testés, plus particulièrement les réseaux VGG 16, VGG 19, Resnet 50, XCption , Inception et NASNet. La combinaison des ces modèles a été élaborée également.

Nous avons étudié aussi la capacité de transfert des connaissances et des caractéristiques apprises dans un premier lieu sur la base d'images ImageNet contenant plus de 14 millions images non médicales aux images médicales constituant un ensemble d'images de taille relativement petite. De ce fait, nous avons choisi d'explorer différents niveaux de transfert d'apprentissage qui peuvent être définis comme un transfert progressif consti-

Introduction générale

tuant un transfert des caractéristiques apprises sur la base ImageNet aux images de notre base de données. Ce type de transfert peut être utile lorsque les réseaux de neurones profonds sont pré-entraînés sur des ensembles de données très différents du domaine en question et lorsque nous disposons d'une base d'images de taille faible.

En somme, ce mémoire vise à apporter une contribution significative à l'identification automatique des modalités d'images médicales et d'organes anatomiques.

Ce mémoire est organisé en quatre chapitres :

- **Le premier chapitre** présente quelques notions de base sur les modalités d'imagerie médicale, l'apprentissage automatique et l'apprentissage profond et leur application en reconnaissance de formes et classification d'images.
- **Le deuxième chapitre** présente les travaux existants sur l'application de l'apprentissage profond dans le traitement des images médicales, notamment la classification d'images médicales. Il aborde ensuite en détails la détection de modalités et d'organes anatomiques médicaux par apprentissage profond, en expliquant les techniques de base et les approches les plus récentes utilisées pour améliorer la précision et la fiabilité des résultats.
- **Le troisième chapitre** présente et détaille nos approches proposées pour la classification des modalités et d'organes anatomiques basée sur l'apprentissage profond par transfert.
- **Le quatrième chapitre** est consacré à l'expérimentation des approches proposées et à l'analyse et discussion des résultats obtenus.
- **A la fin de ce mémoire**, nous clôturons notre mémoire par une conclusion et quelques perspectives.

Chapitre 1

Notions de base en imagerie médicale et intelligence artificielle

1.1 Introduction

Dans ce chapitre, nous allons explorer les définitions clés liées à l'imagerie médicale et à l'intelligence artificielle. Tout d'abord, nous examinerons ce qu'est une image. Nous explorerons également les différents types d'imagerie médicale, tels que la radiographie, l'IRM et la tomodensitométrie, ainsi que leurs avantages.

Enfin, nous aborderons l'intelligence artificielle et ses différentes applications dans le domaine médical, en particulier dans l'analyse d'images médicales. Nous discuterons plus spécifiquement du l'apprentissage profond (Deep learning) et du l'apprentissage par transfert (transfert learning), qui sont deux techniques d'apprentissage automatique très utilisées pour l'analyse d'images médicales. Ce chapitre fournira une introduction aux concepts clés nécessaires pour comprendre les applications de l'imagerie médicale et de l'intelligence artificielle dans le domaine de la santé.

1.2 Image

Une image numérique est une représentation visuelle de données sous forme d'un ensemble de pixels. Chaque pixel est un élément discret de l'image qui contient des informations sur la couleur et l'intensité lumineuse d'un point spécifique de l'image. Les images numériques sont généralement stockées et manipulées sous forme de fichiers informatiques, où chaque pixel est représenté par des valeurs numériques. Ces valeurs peuvent être basées sur différents modèles de couleur, tels que RVB (rouge, vert, bleu) ou niveau de gris.

Cette dernière peut être capturée à l'aide d'un appareil photo numérique, d'un scanner ou générée par des logiciels de création graphique. Elle peut représenter différents types de contenu visuel, tels que des photographies, des illustrations, des graphiques, des schémas, des diagrammes, etc.

L'avantage des images numériques réside dans leur flexibilité et leur facilité de manipulation. Elles peuvent être modifiées, retouchées, redimensionnées, compressées, transformées et combinées à l'aide de logiciels de traitement d'images pour répondre aux besoins spécifiques de l'utilisateur. Cependant, il est important de noter que lors de ces manipulations, des artefacts indésirables peuvent parfois apparaître, ce qui peut altérer la qualité de l'image [1].

1.3 Les Types de formats d'image

Il existe plusieurs types d'images utilisées dans différents contextes. Voici quelques exemples courants :

BitMap : Ce sont des images composées de pixels individuels, organisés en une grille. Ils sont couramment utilisés pour les photographies, les images numériques et les images sur le Web. Les formats les plus populaires du Type BitMap :

- **JPEG ou JPG** : Le format d'image le plus courant sur le web est le .jpeg ou .jpg.

Ce format est largement utilisé, par exemple pour les papiers à en-tête dans les documents Word d'entreprise. Les images dans ce format sont codées et compressées, permettant une lecture universelle par une multitude de logiciels et d'applications sur différents types d'appareils électroniques tels que des ordinateurs, tablettes ou téléphones portables. Le format .jpeg est également léger et peut être rapidement téléchargé même avec une connexion internet bas débit grâce à sa compression maximale. Cependant, ce format est connu pour sa compression "avec perte" qui réduit la qualité de l'image à mesure que le fichier diminue en taille, et cette dégradation est irréversible. Il n'est donc pas recommandé pour la retouche photographique, car chaque enregistrement ultérieur diminue encore la qualité. En outre, ce format ne permet pas de régler la balance des couleurs ou les styles d'images qui sont appliqués de manière permanente [11].

- **PNG** : Le format d'image graphique .png est recommandé pour les documents interactifs tels que les pages web, mais n'est pas adapté pour l'impression. Il est non destructif, ce qui signifie que l'image peut être modifiée sans perte de qualité. Cependant, le format .png a une faible résolution et n'est pas optimal pour les images de haute qualité. Le format .png est également utile si vous souhaitez stocker plusieurs versions de la même image sans compromettre la qualité. Cependant, le principal inconvénient de ce format est sa taille de fichier, qui est significativement plus grande que celle du format .jpeg et peut prendre plus de temps à télécharger. De plus, le format .png ne prend pas en charge les animations ni le stockage de plusieurs images dans un seul fichier [11].
- **GIF** : Les fichiers d'images .gif sont fréquemment utilisés sous forme d'animations. Le format .gif peut inclure jusqu'à 256 couleurs dans le spectre RVB et peut être utilisé pour créer des animations avec plusieurs images qui peuvent être rendues transparentes pour une intégration harmonieuse dans les pages web. Ce format est adapté pour la plupart des plateformes en ligne, telles que les réseaux sociaux et les sites internet. Cependant, les images en format .gif sont limitées à une profondeur de couleurs de 8 bits (256), ce qui peut entraîner une apparence granuleuse ou trop compressée. En outre, le format .gif ne permet pas de reproduire les nuances, telles que les lumières, les ombres et les dégradés [11].

Image vectoriel : Contrairement aux images BitMap, les images vectorielles sont constituées de formes géométriques définies par des courbes, des lignes et des points. Elles sont basées sur des équations mathématiques et peuvent être redimensionnées sans perte de qualité. Le format d'images vectorielles le plus populaire est le SVG.

- **SVG** : Le format de données .svg est utilisé pour décrire des ensembles de graphiques vectoriels en utilisant le langage XML. Il est largement utilisé pour la création d'icônes pour les sites web. Contrairement aux formats .jpeg ou .png, l'image .svg peut être agrandie et compressée sans perte de qualité d'image, et elle ne devient pas floue sur les appareils à haute densité de pixels. Les éléments image .svg ont une petite taille, occupant donc moins d'espace. Le principal avantage du format d'image vectorielle .svg est qu'il permet de redimensionner une image sans perte de

qualité, contrairement aux autres formats. Les formes géométriques de base, comme les rectangles ou les ellipses, sont déjà créées pour faciliter la tâche de l'utilisateur. Le contenu .svg peut être visualisé dans n'importe quel navigateur tel que Chrome, Firefox, Safari [11].

1.4 Modes de représentation d'images

Les modes de représentation d'image sont des techniques qui définissent comment les pixels de l'image sont stockés et organisés.

Il existe plusieurs modes de représentation en traitement d'image, mais les plus couramment utilisés sont :

- **Le mode RVB (Rouge, Vert, Bleu)** : C'est le mode d'image le plus couramment utilisé en photographie numérique et en affichage d'images sur des écrans. Dans ce mode, chaque pixel est représenté par une combinaison de trois valeurs de couleur (rouge, vert et bleu) sur une échelle de 0 à 255 pour chaque canal de couleur, ce qui permet de créer un large éventail de couleurs.
- **Le mode de niveaux de gris** : Dans ce mode, chaque pixel est représenté par une seule valeur de niveau de gris, allant de 0 (noir) à 255 (blanc), ce qui permet de représenter des images en noir et blanc.
- **Le mode binaire** : Dans ce mode, chaque pixel n'a que deux valeurs possibles : 0 ou 1. Les images binaires sont couramment utilisées pour représenter des images en noir et blanc avec des contours clairement définis, telles que les images de textes, les logos, etc.

1.5 Caractéristiques d'une image

Une image a plusieurs caractéristiques :

- **Résolution** : c'est le nombre de pixels de l'image par pouce. Plus la résolution est élevée, plus l'image est nette et détaillée.
- **Taille** : la taille d'une image dépend de sa résolution et de son format. Plus une image est grande, plus elle prendra de l'espace de stockage et de temps de chargement.
- **Profondeur de couleur** : il s'agit du nombre de couleurs qui composent l'image. Plus la profondeur de couleur est élevée, plus l'image est riche en nuances.
- **Compression** : les images peuvent être compressées pour réduire leur taille de stockage et de chargement. Cependant, la compression peut également entraîner une perte de qualité d'image (ex. format jpeg)..
- **Transparence** : certaines images peuvent être rendues transparentes pour permettre une intégration harmonieuse dans d'autres éléments visuels.

1.6 Imagerie médicale

L'imagerie médicale fait référence à plusieurs technologies auxquelles il est fait appel pour étudier le corps humain à des fins de diagnostic, de contrôle ou de traitement des maladies. Ces technologies permettent d'obtenir différentes informations concernant la partie du corps explorée ou traitée pour une éventuelle affection ou lésion, ou de vérifier l'efficacité d'un traitement médicale.

1.6.1 Acquisition d'images médicales

Dans l'imagerie médicale, il existe plusieurs moyens d'acquisition d'images médicales :

- **L'imagerie par résonance magnétique** (IRM) produit des images volumétriques à haute résolution spatiale principalement des noyaux d'hydrogène, en utilisant un champ magnétique appliqué de l'extérieur en conjonction avec des impulsions de radio-fréquence (RF) non ionisantes. L'IRM est couramment utilisé dans de nombreuses applications, y compris l'imagerie musculo-squelettique, cardio-vasculaire et neurologique (Figure 4.5) avec un superbe contraste des tissus mous [8]. L'IRM a deux types, les séquences T1 et T2 font référence aux temps de relaxation des protons dans les tissus du corps qui sont mesurés par l'IRM.

Les séquences T1 sont utilisées pour produire des images de haute qualité des tissus mous du corps, comme les muscles et les organes internes.

Les séquences T2 sont utilisées pour produire des images de haute qualité des fluides du corps, comme le liquide céphalo-rachidien et les tissus oedémateux.

Les IRMs n'utilisent pas de rayonnement ionisant, ce qui les rend plus sûres que les radiographies et les scanners CT. Elles sont particulièrement utiles pour visualiser les tissus mous, tels que les organes internes, les articulations et le cerveau. Cependant, elles sont plus coûteuses que les autres modalités et peuvent prendre plus de temps.



FIG. 1.1 : Image médicale neurologique acquis à partir d'IRM [8]

- **L'imagerie par ultrasons** utilise des impulsions dans la plage de 1 à 10 MHz pour imager les tissus de manière non invasive et relativement peu coûteuse. L'effet de

rétro-diffusion de l'impulsion acoustique interagissant avec les structures internes est utilisé pour mesurer l'écho pour produire l'image (Figure 4.15). L'imagerie par ultrasons est rapide, permettant, par exemple, l'imagerie en temps réel du flux sanguin dans les artères grâce au décalage Doppler [8].

Les images par ultrasons sont rapides, non invasives et n'utilisent pas de rayonnement ionisant. Elles sont souvent utilisées pour visualiser les organes internes, les vaisseaux sanguins et les fœtus pendant la grossesse. Cependant, elles ont des limites en termes de profondeur de pénétration et de qualité d'image, ce qui peut rendre difficile la visualisation de certaines parties du corps.



FIG. 1.2 : Image médicale acquise à partir d'une Échographie [8]

- **Les rayons X** sont des rayonnements électromagnétiques ayant la faculté de traverser certains tissus du corps humain. Ils restituent ainsi des images des os, des muscles et des organes. Les examens utilisant des rayons X sont les radiographies (Figure 1.3), la radioscopie et les scanners (tomodensitométrie, CT scan) [16]. Les radiographies sont rapides, bon marché et largement disponibles. Elles sont souvent utilisées pour diagnostiquer les fractures osseuses, mais peuvent également être utilisées pour visualiser les poumons, les organes abdominaux et d'autres parties du corps. Cependant, elles exposent les patients à une petite quantité de rayonnement ionisant, ce qui peut être dangereux à haute dose.

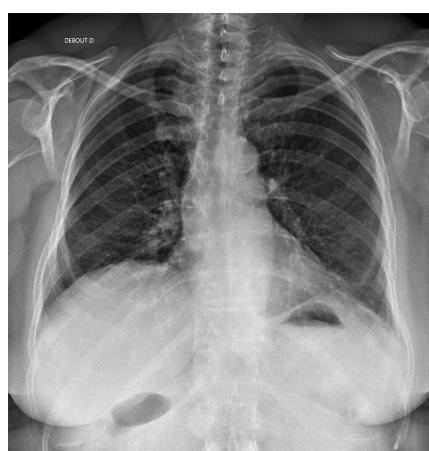


FIG. 1.3 : Radiographie pulmonaire [16]

- **La microscopie** s'agit de l'observation de petits objets, trop petits pour être observés à l'oeil nu. La microscopie optique repose sur l'utilisation de la lumière visible comme source d'illumination, généralement étendue du proche ultraviolet au proche infrarouge (Figure 1.4). Les microscopes sont largement présents dans les services et laboratoires d'anatomo-pathologie où les observations se font sur des coupes fines de tissus issues de biopsies [15].

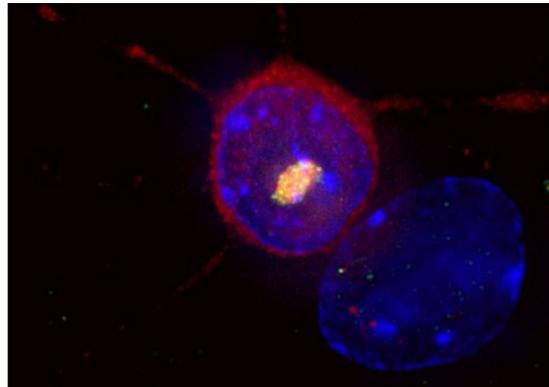


FIG. 1.4 : Image médicale microscopique de la maladie de Huntington (une pathologie neurodégénérative rare et héréditaire. [15])

- **La médecine nucléaire** est la spécialité médicale qui utilise les propriétés de la radioactivité à des fins médicales. Les éléments radioactifs, ou radionucléides, ont la particularité d'émettre différents types de rayonnements. Ces rayonnements traversent différemment les organes selon leur densité. Cette propriété est exploitée pour observer les organes durant leur fonctionnement, permettant ainsi l'étude, le diagnostic et le suivi de très nombreuses maladies (Figure 1.5). La médecine nucléaire, donnant des images en 2 puis maintenant en 3 dimensions, apporte un complément à la radiologie, à l'échographie et à l'IRM [17].

La scintigraphie est une technique d'imagerie nucléaire qui utilise des radiopharmaceutiques pour visualiser la fonction des organes et des tissus. Elle est souvent utilisée pour diagnostiquer les maladies cardiaques, la thyroïde et les tumeurs osseuses. Cependant, elle expose les patients à une petite quantité de rayonnement ionisant.

La PET-CT est une technique d'imagerie nucléaire qui combine une tomographie par émission de positrons (PET) et une tomographie par ordinateur (CT). Elle est souvent utilisée pour diagnostiquer les tumeurs, la maladie d'Alzheimer et d'autres troubles neurologiques. Cependant, elle expose les patients à une quantité importante de rayonnement ionisant.

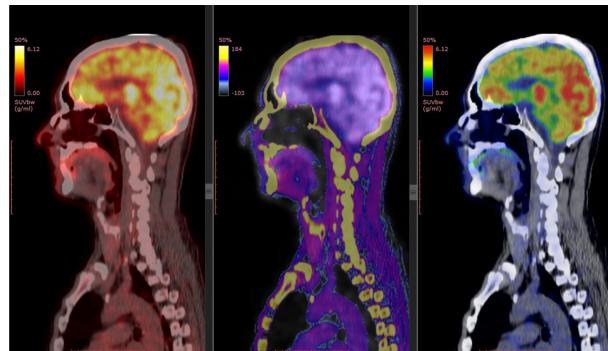


FIG. 1.5 : Image médicale acquis avec la médecine nucléaire. [17]

1.7 Intelligence artificielle et Imagerie médicale

L'intelligence artificielle (IA, ou AI en anglais pour Artificial Intelligence) consiste à mettre en œuvre un certain nombre de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle. L'IA se retrouve implémentée dans un nombre grandissant de domaines d'application [8].

1.7.1 Apprentissage automatique

L'apprentissage automatique (Machine learning) est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'apprendre à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches (classification, reconnaissance, prédiction, détection d'anomalie ...) sans être explicitement programmés pour chacune [12].

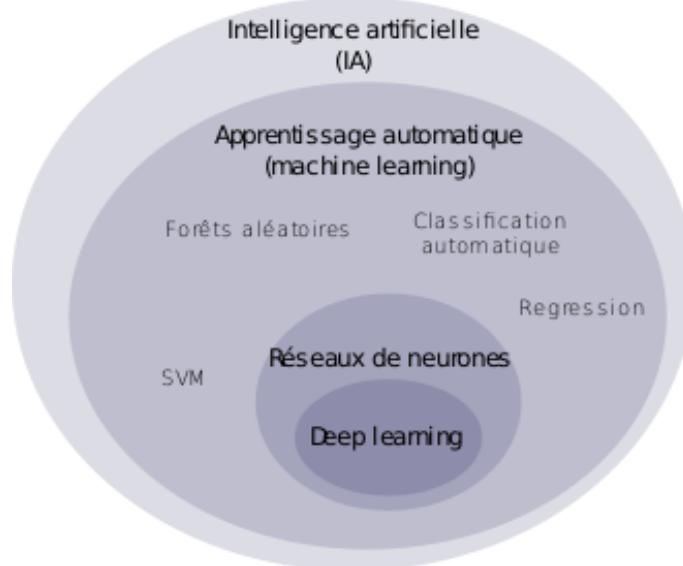


FIG. 1.6 : Les champs de l'intelligence artificiel [8]

A ce niveau, différents classificateurs ont été développés, les plus fréquents : KPPV(KNN), K-means, les réseaux de neurones (ANN), SVM, arbre de décision, forêt d'arbre aléatoire (RF). Ces classificateurs sont toujours d'actualité, car ils peuvent être intégrés dans des architectures d'apprentissage profond (voir la section 1.4.2).

KNN

KPPV ou k- Plus proches voisins (k-Nearest Neighbors), est un algorithme d'apprentissage utilisé dans le domaine de l'apprentissage automatique. Il est largement utilisé pour la classification et la régression dans les tâches de prédiction. L'idée de base derrière KNN est de trouver les k échantillons d'entraînement les plus proches d'un nouvel échantillon dont nous souhaitons prédire l'étiquette.

K-Means

est un algorithme de clustering qui permet de regrouper des données en fonction de leur similarité. L'objectif de l'algorithme est de diviser un ensemble de données en K clusters distincts, où chaque cluster représente un groupe de données ayant des caractéristiques similaires.

ANN

Réseau de Neurones Artificiels (Artificial Neural Network) C'est un type de modèle d'apprentissage automatique qui est inspiré du fonctionnement du cerveau humain. Les réseaux de neurones artificiels sont constitués de couches de neurones interconnectés qui sont capables d'apprendre à partir de données d'entrée et de produire des prédictions en fonction de ces données. Le modèle le plus utilisé dans cette catégorie est le perceptron multi-couches (MLP). Un réseau de neurones artificiels, ou réseau neuronal artificiel, 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, en particulier bayésien. 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 de l'implémenter, et des informations d'entrée au raisonnement logique formel

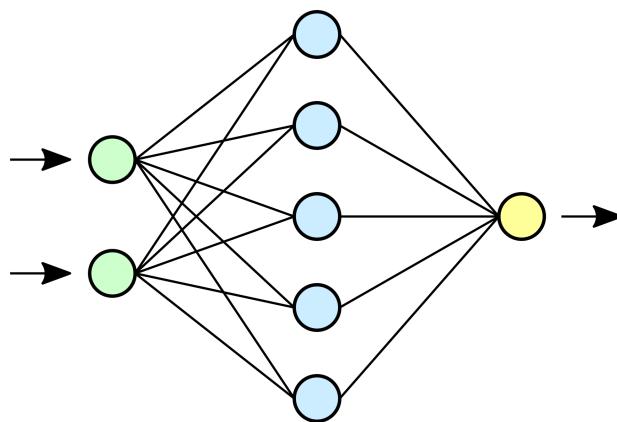


FIG. 1.7 : Réseau de neurones [8]

Arbre de décision

est une méthode d'analyse prédictive utilisée en apprentissage automatique pour prendre des décisions en fonction de multiples conditions. Il s'agit d'un modèle de type arborescent où chaque noeud représente une condition sur une variable d'entrée. L'un des avantages de l'arbre de décision est sa facilité d'interprétation, car il est possible de comprendre les décisions prises en suivant simplement le chemin de l'arbre. Cependant, il peut être sensible aux données aberrantes et aux choix initiaux de l'algorithme, ce qui peut conduire à un arbre non optimal.

Forêt d'arbres aléatoires

La forêt d'arbres aléatoires (Random Forest en anglais) est un algorithme d'apprentissage automatique qui est utilisé pour la classification, la régression et d'autres tâches de prédiction. Il fonctionne en construisant un grand nombre d'arbres de décision, chacun étant entraîné sur un sous-ensemble aléatoire des données d'entraînement. Les prédictions sont ensuite obtenues en moyennant les prédictions de chaque arbre individuel.

Cependant, l'inconvénient de la forêt d'arbres aléatoires est qu'elle peut être sujette à l'overfitting, en particulier si le nombre d'arbres dans la forêt est trop grand. En outre, elle peut être plus lente à s'exécuter que certains autres algorithmes plus simples, tels que la régression logistique.

Pour obtenir les meilleurs résultats avec la forêt d'arbres aléatoires, il est important de choisir les paramètres optimaux, tels que le nombre d'arbres dans la forêt, la profondeur de chaque arbre, et les critères utilisés pour diviser les noeuds de l'arbre.

SVM

Le SVM (Support Vector Machine) est un algorithme d'apprentissage supervisé utilisé pour la classification et la régression. L'objectif du SVM est de trouver l'hyperplan qui

sépare le mieux les données en classes distinctes. L'hyperplan est choisi de manière à maximiser la marge, c'est-à-dire la distance entre l'hyperplan et les points les plus proches de chaque classe. L'avantage du SVM est sa capacité à bien fonctionner dans des espaces de grande dimension, même lorsque le nombre de dimensions est supérieur au nombre d'échantillons. Cependant, le SVM peut être sensible au choix de la fonction de noyau et de ses paramètres, et le temps de calcul peut être important pour de grands ensembles de données.

L'apprentissage automatique présente un problème d'extraction des caractéristiques qui sont régies dans l'apprentissage profond

1.7.2 Apprentissage profond

L'apprentissage profond (Deep learning) est un type d'apprentissage automatique ou machine learning qui repose sur l'apprentissage de représentations de données par opposition à des algorithmes, qui sont spécifiques à une tâche. L'apprentissage profond utilise des couches cachées de réseaux de neurones artificiels combinés avec des classificateurs. Les algorithmes d'apprentissage profond s'opposent aux algorithmes d'apprentissage peu profonds du fait du nombre de transformations réalisées sur les données entre la couche d'entrée et la couche de sortie, où une transformation correspond à une unité de traitement définie par des poids et des seuils. En d'autres termes, l'algorithme apprend à partir d'exemples de données afin d'être en mesure de produire des résultats précis lorsqu'on lui présente de nouvelles données [13].

L'apprentissage profond présente des problèmes de limitation des données et il faut beaucoup de temps pour apprendre, pour cette raison l'apprentissage par transfert a été créé pour résoudre ces problèmes

1.7.3 L'apprentissage profond par transfert

L'apprentissage par transfert (transfer learning en anglais) est l'un des champs de recherche de l'apprentissage automatique qui vise à transférer des connaissances d'une ou plusieurs tâches sources vers une ou plusieurs tâches cibles. Il 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.

En vision par ordinateur, le transfert d'apprentissage est souvent utilisé pour les réseaux de neurones convolutionnels (CNN).

Le TF est utile dans les situations où il est difficile ou coûteux d'obtenir suffisamment de données pour entraîner un nouveau modèle de zéro. En utilisant un modèle pré-entraîné, on peut profiter de l'apprentissage qui a déjà été réalisé sur une tâche similaire et ainsi accélérer le processus d'entraînement.

Il existe deux mécanismes de TL : Mode extracteur de caractéristique et mode fine tuning.

- **Mode Extracteur de caractéristiques**

Dans cette méthode, on utilise un modèle pré-entraîné pour extraire des caractéristiques à partir d'images (par exemple, des motifs de lignes, des formes, des couleurs, etc.). Ensuite, on entraîne un nouveau modèle de classification en utilisant les caractéristiques extraites à partir des images d'entraînement. Cette méthode est souvent utilisée lorsque le jeu de données d'entraînement est petit et que le modèle pré-entraîné a fait son entrainement sur des images très similaires à celles de la base cible, donc il peut être utilisé pour extraire des caractéristiques pertinentes sur la nouvelle base.

- **Mode de fine tuning**

En apprentissage automatique, le réglage fin est une approche de l'apprentissage par transfert dans laquelle les poids d'un modèle pré-entraîné sont formés sur de nouvelles données. Le réglage fin peut être effectué sur l'ensemble du réseau neuronal ou sur un sous-ensemble de ses couches, auquel cas les couches qui ne font pas l'objet d'un réglage fin sont "gelées" (non mises à jour pendant l'étape de rétropropagation). Dans certaines architectures, comme les réseaux CNN, il est courant de geler les premières couches (les plus proches de la couche d'entrée) parce qu'elles capturent des caractéristiques de niveau inférieur (descripteurs d'images génériques), tandis que les couches ultérieures discernent souvent des caractéristiques de haut niveau qui peuvent être davantage liées à la tâche sur laquelle le modèle est entraîné. Les modèles pré-entraînés sur des corpus larges et généraux sont généralement affinés en réutilisant les paramètres du modèle comme point de départ et en ajoutant une couche spécifique à la tâche entraînée à partir de zéro. L'affinement du modèle complet est également courant et donne souvent de meilleurs résultats sur des datasets larges, mais il est plus coûteux en temps de calcul.

1.7.4 Réseau neuronal convolutif (CNN)

Les CNN (Convolutional Neural Networks) ou réseaux de neurones convolutifs en français, sont un type d'architecture de réseau de neurones artificiels, qui sont souvent utilisés pour la reconnaissance d'images, la classification et la segmentation d'images, ainsi que d'autres tâches de vision par ordinateur.

Les CNN sont constitués de couches de neurones qui effectuent des opérations de convolution sur les entrées de données, généralement des images, afin de les transformer en une représentation plus abstraite et compréhensible par le réseau. Les opérations de convolution permettent de capturer les caractéristiques importantes de l'image, telles que les contours, les textures et les formes, en utilisant des filtres qui sont appris par le réseau lors de la phase d'apprentissage.

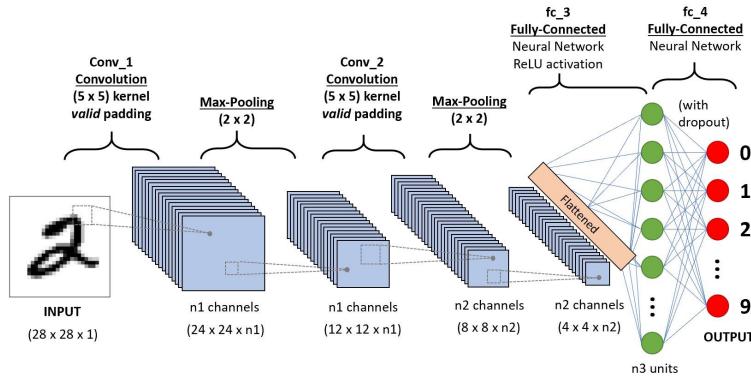


FIG. 1.8 : Architecture d'un CNN

L'opération de convolution

La convolution est une opération mathématique qui consiste à appliquer un filtre (aussi appelé noyau ou kernel) à une image ou une représentation de données en entrée.

Le filtre est une petite matrice de nombres qui est appliquée à chaque position de l'image ou de la représentation de données. La convolution permet de détecter des motifs locaux dans l'image ou la représentation de données, comme des bords, des coins, ou des textures spécifiques.

La sortie de la convolution est une nouvelle représentation de données, appelée carte de caractéristiques ou feature map, qui met en évidence les motifs détectés par le filtre. Cette carte de caractéristiques peut ensuite être utilisée comme entrée pour d'autres couches de réseau de neurones, afin de réaliser des tâches plus complexes telles que la classification d'images, la segmentation sémantique ou la détection d'objets.

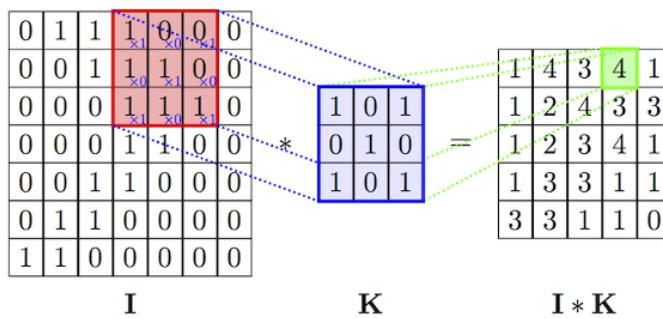


FIG. 1.9 : L'opération de convolution

L'opération de max pooling

Le max pooling est une opération qui permet de réduire la taille d'une carte de caractéristiques (ou feature map) tout en préservant les informations les plus importantes.

Le max pooling consiste à diviser la carte de caractéristiques en zones non-overlapping de taille fixe (par exemple 2x2) et à remplacer chaque zone par la valeur maximale de cette zone.

Cette opération permet de réduire la dimensionnalité de la carte de caractéristiques tout en conservant les caractéristiques les plus saillantes. En effet, en prenant la valeur maximale de chaque zone, on conserve les zones les plus activées, c'est-à-dire celles qui contiennent les informations les plus importantes.

Le max pooling est souvent utilisé après une convolution, pour réduire la taille de la carte de caractéristiques avant de la passer à une couche suivante du réseau de neurones. Cette opération peut également aider à réduire le surapprentissage (overfitting) en introduisant une certaine invariance aux petites translations dans l'image d'entrée.

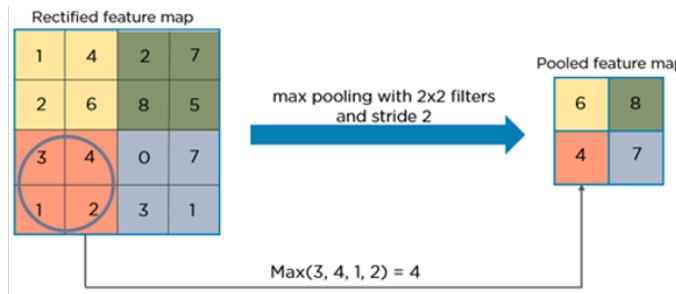


FIG. 1.10 : L'opération de Max pooling

Flattened layers

Les flattened layers ou couches aplatises en français, sont une étape finale qui est souvent utilisée dans les réseaux de neurones convolutionnels (CNN) pour transformer une carte de caractéristiques (ou feature map) en un vecteur unidimensionnel.

Après une série de couches de convolution et de pooling, la sortie du réseau de neurones est souvent une carte de caractéristiques qui est une représentation en 2 ou 3 dimensions des caractéristiques de l'image d'entrée. Pour passer de cette représentation à un vecteur unidimensionnel, on applique une opération de "flattening" (aplatissement), qui consiste à transformer chaque pixel de la carte de caractéristiques en une seule valeur.

Pour ce faire, on prend tous les pixels de la carte de caractéristiques et on les concatène en un seul vecteur unidimensionnel. Ainsi, chaque élément du vecteur correspond à une caractéristique particulière extraite de l'image d'entrée.

Cette transformation en vecteur unidimensionnel permet ensuite d'utiliser des couches de neurones entièrement connectées pour réaliser des tâches telles que la classification d'images ou la prédiction de valeurs numériques.

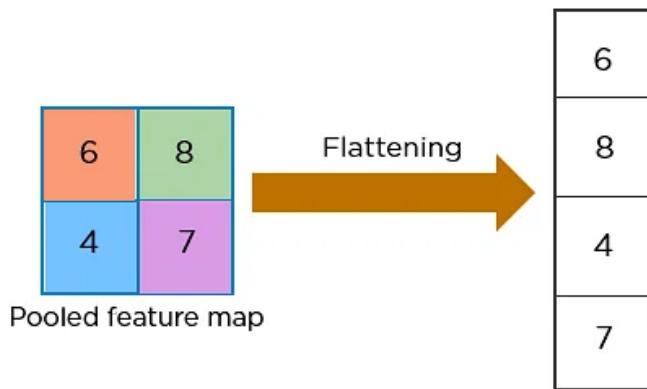


FIG. 1.11 : L'opération de flattening

Couche fully-connected (FC)

La couche FC applique successivement une combinaison linéaire puis une fonction d'activation afin de classifier l'image d'entrée, elle renvoie un vecteur de taille N, où N est le nombre de classes correspondantes au problème étudié.

Les fonctions d'activation

Un fonction d'activation est une fonction mathématique non-linéaire appliquée à la sortie d'un neurone ou d'une couche de neurones d'un réseau de neurones artificiels, elles sont souvent utilisées dans les CNN. Les fonctions d'activation non-linéaires permettent au réseau de neurones de modéliser des relations non-linéaires entre les entrées et les sorties. Il existe plusieurs fonctions d'activation couramment utilisées en deep learning pour la vision par ordinateur, notamment :

- Sigmoïde : cette fonction prend en entrée un nombre réel quelconque et renvoie une valeur comprise entre 0 et 1.
- Tanh : cette fonction est similaire à la sigmoïde, mais elle renvoie une valeur comprise entre -1 et 1.
- ReLU (Rectified Linear Unit) : cette fonction est devenue l'une des plus populaires en vision par ordinateur. Elle renvoie la valeur d'entrée si elle est positive, sinon elle renvoie 0. Elle permet une convergence plus rapide de l'algorithme de rétropropagation.
- ELU (Exponential Linear Unit) : cette fonction est une autre alternative à la ReLU. Elle est similaire à la fonction Leaky ReLU, mais elle permet à l'entrée négative de passer et utilise une fonction exponentielle pour les valeurs négatives.
- Softmax : cette fonction est souvent utilisée pour la classification multiclasse. Elle prend un vecteur de scores en entrée et renvoie un vecteur de probabilités normalisé qui somme à 1.

Les fonctions d'optimisation

Il existe plusieurs fonctions d'optimisation couramment utilisées en intelligence artificielle pour entraîner les modèles. Voici quelques-unes des fonctions d'optimisation les plus courantes :

- Descente de gradient : c'est une méthode d'optimisation itérative qui ajuste les paramètres du modèle pour minimiser une fonction de perte. Elle est souvent utilisée dans les réseaux de neurones pour entraîner des modèles de classification ou de régression.
- Optimisation par essaims de particules : cette méthode d'optimisation est inspirée de la nature et utilise un ensemble de particules pour explorer l'espace de recherche. Elle est souvent utilisée pour optimiser des fonctions non convexes.
- Optimisation bayésienne : cette méthode utilise des modèles probabilistes pour trouver les paramètres optimaux d'un modèle. Elle est souvent utilisée pour l'optimisation de modèles de machine learning à coût élevé.
- Optimisation par algorithme génétique : cette méthode est inspirée de l'évolution biologique et utilise des stratégies de sélection naturelle pour améliorer les performances du modèle. Elle est souvent utilisée pour l'optimisation de structures de réseaux de neurones.
- Optimisation Adam : est utilisé pour la formation de modèles d'apprentissage profond. Il s'agit d'une extension de la descente de gradient stochastique. Dans cet algorithme d'optimisation, les moyennes courantes des gradients et des seconds moments des gradients sont utilisées. Il est utilisé pour calculer les taux d'apprentissage adaptatifs pour chaque paramètre.

Les hyperparamètres

sont des paramètres externes qui sont utilisés pour contrôler le processus d'apprentissage dans les modèles d'apprentissage automatique. Contrairement aux paramètres du modèle, qui sont appris à partir des données d'entraînement, les hyperparamètres sont définis par l'utilisateur avant le début de l'apprentissage.

Les hyperparamètres peuvent varier selon le type de modèle, mais ils peuvent inclure des éléments tels que le taux d'apprentissage, le nombre d'itérations, la taille du lot, la régularisation, la complexité du modèle et le choix de la fonction d'activation. Ces hyperparamètres peuvent avoir un impact significatif sur la performance et la précision du modèle, et donc leur réglage est une étape cruciale dans le processus d'apprentissage.

1.7.5 Les architectures CNN les plus populaires

LENet-5

Yann LeCun a développé LeNet dans les années 1990, qui est considéré comme le premier réseau de neurones convolutionnels [3]. LeNet a été amélioré en 1998 et est devenu

l'une des architectures les plus efficaces et les plus connue [2], notamment utilisée pour la lecture des codes postaux et des chiffres. Cette architecture utilise quatre couches convolutives et de mise en commun alternées, suivies de trois couches entièrement connectées. LeNet a été le premier CNN capable d'apprendre automatiquement les caractéristiques des pixels bruts, tout en réduisant le nombre de paramètres.

AlexNet

AlexNet est l'architecture CNN l'une des plus célèbres en vision par ordinateur, popularisée par Alex Krizhevsky, Ilya Sutskever et Geoff Hinton en 2012 [4]. AlexNet a participé au défi ImageNet ILSVRC en 2012 et a considérablement surpassé les performances des autres finalistes. Cette architecture est composée de 5 couches convolutives utilisant des unités linéaires rectifiées (ReLU) comme fonctions d'activation, 3 couches de max pooling et 3 couches entièrement connectées (FC).

VGG

L'efficacité des CNN dans les tâches de reconnaissance d'image a accéléré les recherches en matière de conception architecturale. Simonyan et al on proposé une méthode simple et efficace pour la conception d'architectures CNN. Cette architecture, appelée le Groupe visuel Géométrie (VGG) de l'Université d'Oxford [5], a remplacé les filtres 11x11 et 5x5 par une pile de couches de filtres de 3x3. L'utilisation de filtres plus petits permet également de réduire le nombre de paramètres, ce qui réduit la complexité de calcul. VGG a ainsi contribué à améliorer la performance et la complexité des CNN. on a plusieure variété des modèles : VGG-11, VGG-16 et VGG-19, le nombre implique simplement l'architecture (le nombre des couches utiliser)

ResNet

Kaiming He et son équipe ont développé l'architecture CNN ResNet [7], qui utilise des connexions de saut uniques et la normalisation par lots (Batch Normalization) pour améliorer les performances du réseau. Contrairement aux architectures précédentes, ResNet ne comporte pas de couches entièrement connectées à la fin du réseau. Cependant, cette architecture est très coûteuse à évaluer en raison de la grande quantité de paramètres. Malgré cela, ResNet est considéré comme l'un des modèles de réseau neuronal convolutif les plus avancés à ce jour et est largement utilisé dans la pratique. En effet, RestNet a remporté le défi ILSVRC 2015. on a plusieure variété des modèles : ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-110, ResNet-152, ResNet-164, ResNet-1202, ect..., Le nom ResNet suivi d'un nombre à deux chiffres ou plus implique simplement l'architecture ResNet avec un certain nombre de couches de réseau neuronal.

GoogleNet (Inception-V1)

GoogleNet, également connu sous le nom d'Inception-V1, a remporté le concours ILSVRC de 2014 et a été développé par une équipe de Google dirigée par Christian Szegedy

et d'autres chercheurs. Cette architecture CNN utilise des modules Inception, qui sont des blocs contenant des filtres de différentes tailles (1x1, 3x3 et 5x5) pour capturer des informations spatiales à différentes échelles. Ces filtres sont suivis d'un filtre Concat qui permet de concaténer les résultats des filtres. De plus, GoogleNet utilise une mise en commun moyenne globale à la dernière couche au lieu d'une couche entièrement connectée, ce qui réduit considérablement le nombre de paramètres de 60 millions à 4 millions [6].

NASNet

Google a introduit NASNet, qui a posé le problème de trouver la meilleure architecture CNN comme un problème d'apprentissage par renforcement. Essentiellement, l'idée était de rechercher la meilleure combinaison de paramètres de l'espace de recherche donné, tels que les tailles de filtre, les canaux de sortie, les pas, le nombre de couches, etc. Dans ce cadre d'apprentissage par renforcement, la récompense après chaque action de recherche était la précision de l'architecture recherchée sur l'ensemble de données donné. NASNet a obtenu des résultats de pointe lors de la compétition ImageNet. Cependant, la puissance de calcul requise pour NASNet était si importante que seules quelques entreprises ont pu utiliser la même méthodologie.

1.7.6 Base d'images (Dataset)

Une base d'images ou dataset (ensemble de données) est une collection d'images (ou d'autres données) qui est utilisée pour entraîner, valider ou tester un réseau de neurones convolutionnel (CNN).

Les datasets contiennent généralement des images annotées avec des étiquettes (labels) qui indiquent la classe ou la catégorie à laquelle appartient chaque image. Par exemple, un dataset de classification d'images de fruits pourrait contenir des images de pommes, de bananes, d'oranges, etc. avec des étiquettes correspondantes.

Les datasets sont utilisés pour entraîner des réseaux de neurones convolutionnels en leur fournissant un grand nombre d'exemples d'images et de leurs étiquettes correspondantes. Lors de l'entraînement, le réseau de neurones apprend à reconnaître les caractéristiques qui permettent de distinguer les différentes classes d'images.

Une fois le réseau entraîné, il est testé sur un autre ensemble de données, appelé ensemble de test, pour évaluer ses performances de classification. Les datasets sont donc un élément clé de l'apprentissage profond et sont utilisés pour entraîner des modèles de classification, de détection d'objets, de segmentation sémantique, entre autres applications.

dans l'étape d'apprentissage c'est dataset est petit on peut l'augmenter, Cela implique d'appliquer des transformations aux images, telles que le retournement, la rotation, la mise à l'échelle, le recadrage ou le réglage de la luminosité/du contraste. Cela peut aider à améliorer la robustesse des modèles de classification d'images.

1.8 Conclusion

En conclusion, nous avons exploré dans ce chapitre les notions de base relatives à l'image et à l'imagerie médicale, ainsi que l'intelligence artificielle et plus particulièrement le deep learning et le transfert learning. Nous avons vu que l'image est une représentation visuelle d'un objet ou d'une scène . En imagerie médicale, l'image est utilisée pour diagnostiquer et traiter les maladies, et peut être obtenue à partir de différentes modalités telles que la radiologie, l'IRM, la tomodensitométrie, etc.

L'intelligence artificielle a révolutionné le domaine de l'imagerie médicale en permettant une analyse plus rapide et plus précise des images. Le deep learning est une technique d'apprentissage automatique basée sur des réseaux de neurones artificiels profonds, qui peut être utilisée pour la segmentation d'images, la classification de tumeurs, la détection de maladies, etc. Le transfert learning, quant à lui, permet d'utiliser les connaissances acquises lors de l'apprentissage sur un ensemble de données pour en faciliter l'apprentissage sur un autre ensemble de données.

En somme, l'imagerie médicale et l'intelligence artificielle sont des domaines en constante évolution qui offrent de nouvelles perspectives pour le diagnostic et le traitement des maladies. En comprenant les notions de base de l'image et de l'apprentissage automatique, nous pouvons mieux comprendre comment ces technologies fonctionnent et comment elles peuvent être utilisées pour améliorer la santé et le bien-être des patients.

Chapitre 2

Détection de modalités et d'organes anatomique médicale par apprentissage profond

2.1 Introduction

Le domaine de l'imagerie médicale est en constante évolution et l'utilisation de techniques avancées d'apprentissage automatique est en train de révolutionner la façon dont les images médicales sont analysées et interprétées. Dans ce chapitre, nous allons nous concentrer sur la détection de modalités et d'organes anatomiques médicaux par apprentissage profond. Nous commencerons par présenter quelques applications existantes de l'apprentissage profond dans le traitement d'image médicale, notamment la classification d'images médicales. Nous aborderons ensuite en détail la détection de modalités et d'organes anatomiques médicaux par apprentissage profond, en expliquant les techniques de base et les approches les plus récentes utilisées pour améliorer la précision et la fiabilité des résultats.

2.2 Application de l'apprentissage profond en imagerie médicale

L'intelligence artificielle et l'apprentissage automatique et profond en particulier ont changé la relation qu'a l'imagerie médicale avec l'informatique, et a révolutionné le domaine et ouvert plusieurs portes et sujets de recherche. L'apprentissage profond peut être utilisé pour extraire des informations utiles à partir d'images médicales complexes. Les réseaux de neurones profonds peuvent être entraînés à reconnaître des motifs et des structures spécifiques dans des images médicales, et à effectuer des tâches telles que la détection de maladie ou la classification de tumeurs.

L'utilisation de l'apprentissage profond, et plus particulièrement des réseaux de neurones convolutifs (CNN), pour créer des systèmes de prédiction et de diagnostic médicale, est devenue très populaire ces dernières années. Les CNN offrent en effet de bonnes performances pour l'analyse d'images médicales. De nombreux travaux ont ainsi été réalisés pour différentes pathologies, notamment les pathologies rétiniennes, cardiaques, abdominales et musculo-squelettiques, avec des résultats très intéressants.

Google a également travaillé à la lutte contre le cancer du poumon en développant un algorithme capable de détecter le cancer de manière plus précise qu'un humain. En comparant leurs résultats à ceux des meilleurs oncologues américains, ils ont constaté que leur technologie a permis de détecter 5% de cancers supplémentaires. En outre, l'algorithme a permis de réduire de 11% les faux positifs [18].

Des chercheurs se concentrent sur d'autres sujets de recherche utilisant des algorithmes d'apprentissage profond en raison de leur excellente performance dans l'analyse et la compréhension d'images, notamment la segmentation des métastases hépatiques par tomodensitométrie (TDM) [19], la segmentation des tumeurs cérébrales sur des images et la classification des maladies pulmonaires interstitielles avec une tomodensitométrie thoracique à haute résolution [20]. Ces tâches nécessitent une analyse précise de l'image et les algorithmes d'apprentissage profond peuvent offrir des résultats plus fiables.

2.3 Application de l'apprentissage profond pour la classification d'organe anatomique

La classification automatisée d'organes anatomiques est une étape importante et une condition préalable pour de nombreux systèmes de diagnostic médicale. La complexité spatiale et la variabilité de l'anatomie dans l'ensemble du corps humain rendent la classification difficile, 8 chercheurs de l'institut national de santé des USA, ont entraîné un modèle à identifier des images de tomographie axiale, en utilisant des images clés acquises à partir d'archive PACS hospitalier en les spécifiant en 5 classes anatomique : le cou, les poumons, le foie, le bassin et les jambes. Le modèle utilisé est un CNN qui a atteint une précision de 0,998. [21].

HyperParameter	Layer 1	Layer 2	Layer 3	Layer 4
Number of filters	96	256	384	256
kernel size	11×11	5×5	3×3	3×3
stride	4	2	2	2
Learning rate	0.01			
Momentum	0.9			
Weight Decay	0.0005			
Training epochs	30-60			
Number of units in fully connected layer				4096

FIG. 2.1 : Les hypers paramètre du modèle proposé [21]

En 2017, des chercheurs japonais ont travaillé sur la classification d'images médicales prises par endoscopie œsophagogastrroduodénale pour reconnaître les emplacements anatomiques. Les scientifiques ont utilisé 27 335 images d'endoscopie gastro-œsophagienne (EGD) d'hôpitaux japonais, pour entraîner leur modèle de diagnostic, Les images ont été catégorisées en quatre emplacements anatomiques et trois emplacements supplémentaires de l'estomac, permettant une classification anatomique précise des images EGD. Les classes utilisées étaient le larynx, l'œsophage, l'estomac et le duodénum, ainsi que les subdivisions de l'estomac supérieur, moyen et inférieur. Le modèle utilisé a été GoogleNet qui a eu une précision de 97% [22].

Un article apparu en 2017 aussi, présente le travail de trois chercheurs de Malaisie pour la classification automatique d'anatomie dans les images médicales. Ils ont évalué trois architectures CNN, LeNet, AlexNet et GoogLeNet, et proposent une architecture CNN. Les chercheurs ont utilisé un ensemble de données de 37 198 images de diverses organes anatomiques obtenues à partir de différentes modalités d'imagerie médicale, notamment CT, MRI, PET, échographie et rayons X. La classification a été effectuée en utilisant des

classes représentant différentes parties du corps humain, notamment la tête et le cou, le thorax, l'abdomen et le bassin, les membres supérieurs et les membres inférieurs [23].

Model	Runtime in seconds	Validation Loss	Validation Accuracy (%)	Test Accuracy (%)
LeNet[14]	1655	1.3	58	59
AlexNet[15]	33466	1.39	65	74
GoogLeNet[17]	52470	1.2	55	45
Proposed CNN	16728	0.67	76.6	81

FIG. 2.2 : Les résultats des différents CNN [23]

2.4 Comparatif des approches existantes de classification d'organe anatomique

Voici un tableau comparatif des différents travaux de la communauté scientifique dans la classification d'organe anatomique.

Méthode	Modèle	Accuracy	Base de données	Remarque
Holger el al [21]	CNN personnelle	99.8	Archive de PACS hospitalier	Classification des images de tomographie
Takiyama et al [22]	GoogleNet	97	Hôpital japonais	Classification anatomique précise des images EGD
Khan et al [23]	LeNet AlexNet GoogleNet modèle proposé	59 74 45 81	Source inconnus	Classification automatique d'anatomie dans les images médicale

TAB. 2.1 : Tableau comparatif des différentes approches pour la classification d'organes anatomique

2.5 Application de l'apprentissage profond pour la classification de modalités d'imagerie médicale

La classification d'image médicale par rapport à la modalité d'acquisition est une étape primaire et importante, car elle facilite le travail de recherche fait par les médecins et pro-

fessionnels du secteur médicale, mais en connaissant la modalité d'image, les cliniciens peuvent interpréter correctement l'image et prendre des décisions médicales précises. Par exemple, elle permet aux cliniciens de trouver rapidement les images médicales dont ils ont besoin pour effectuer des diagnostics, planifier des traitements et suivre la progression des maladies.

Parmi les travaux ayant utilisé les technique d'apprentissage profond pour la détection de modalités, nous trouvons le travail de **Yu et al**, qui ont utilisé l'apprentissage profond par transfert pour pré-entraîner plusieurs CNN sur un ensemble de données plus important (ImageNet) et le réajuster sur l'ensemble de données médicales (ImageCLEF2015 et ImageCLEF2016). Les modèles utilisés dans ce travail sont CNN-6 , VGGNet-16 et ResNet-50. Les modèles VGGNet-16 et ResNet-50 sont conçus pour 1000 classes, donc ils ont remplacé la dernière couche entièrement connectée avec 30 neurones pour produire 30 probabilités postérieures. Les résultats de leurs travaux sont dans la figure ci-dessous [24].

Methods	10FCV		Evaluation	
	DS_Original	DS_Aug1	DS_Aug1	DS_Aug2
Baseline_2015 [10,14]	-	-	60.91	-
CNN-6	59.95	57.09	58.33	66.13
VGGNet-16	87.27	67.83	70.50	71.61
ResNet-50	89.34	72.10	75.75	76.78
Our proposed model	90.22	72.42	76.07	76.87

Table 2. Accuracy of visual methods in ImageCLEF2016.

Methods	10FCV		Evaluation	
	DS_Original	DS_Aug1	DS_Aug1	DS_Aug2
Baseline_2016 [23]	-	-	85.38	-
CNN-6	75.87	70.67	74.70	81.86
VGGNet-16	85.13	78.99	81.73	83.54
ResNet-50	87.47	82.51	85.25	86.92
Our proposed model	88.40	82.61	86.07	87.37

FIG. 2.3 : Les résultats de l'étude de Yuhai et al [24]

Dans des travaux plus récente, on peut trouver le travail de **Kim et al** [25], qui ont utilisé la base de données the Open Access Biomedical Image Search engine de la librairie national de médecine des état unis (NLM), mais aussi la base de donnée ImageCLEF2013. Pour leur recherche, ils ont utilisé une nouvel méthode appelée Class-selective Relevance Mapping (CRM), pour localiser et visualiser les régions d'intérêt dans une image médicale afin d'améliorer l'explication des prédictions des modèles CNN pour la classification d'images médicales. La méthode CRM mesure les contributions positives et négatives de chaque élément spatial dans les cartes de caractéristiques produites à partir de la dernière couche de convolution, ce qui permet une classification correcte de l'image d'entrée. Ainsi avec le modèle qu'ils ont utilisé un VGG16 pré-entraîné pour classer sept types différents modalités d'image, ils ont eu une précision de 0.98 .

Dans une autre étude, de 4 chercheurs américains **Remedios et al** [26], ont présenté une technique d'apprentissage en profondeur pour identifier des images médicales acquises par IRM dans le but de classifier les images IRB en catégories spécifiques pour une récupération efficace. Avec un total de 3418 images médicales par IRM acquis à partir de plusieurs sources, les images ont été divisées en plusieurs classes par rapport aux diffé-

rents types de scanner IRM (T1, T2, FLAIR), ainsi que les sous-classes T1 pré vs post et FLAIR pré vs post. L'architecture utilisé dans cette étude est un CNN appelé Φ -Net . Ce modèle a été créé à partir de combinaison CNNs avec le concept d'apprentissage résiduel [27]. Les résultats de l'étude ont que Φ -Net a eu une précision moyenne de 97.57% pour la classification de T1 vs T2 vs FLAIR et ils ont comparé Φ -Net avec ResNet qui a eu une précision de 95.47%.

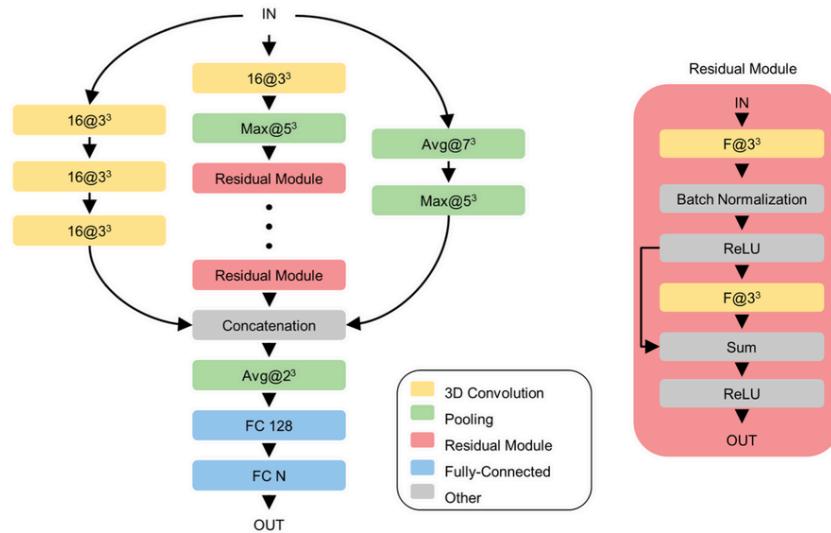


FIG. 2.4 : Architecture de Φ -Net [28]

En 2021, l'approche proposée par **Chiang et al** [29], porte sur un modèle CNN pour la classification de 4 classes d'imagerie médicale (CT de l'abdomen, CT du cerveau, IRM du cerveau et IRM de la colonne vertébrale). Les auteurs, ont utilisé des images médicales au format JPEG, acquis auprès de l'hôpital taiwanais Shuang-Ho. Le nombre d'images de chaque classe est d'environ 700, le model utilisé par ces derniers, un CNN modifié à partir d'une architecture CNN fournie par Matlab 2020a (The MathWorks, Inc., Natick, MA) pour la classification des données manuscrites. Avec ce modèle, ils ont une précision de 99.5% et un F-score de 99%.

2.6 Comparatif des approches existantes de classification des modalités d'imagerie médicale

Chaque approche présentée en amont, a ses avantages et ses inconvénients. Ainsi pour bien étudier toutes les approches nous avons fait un tableau comparatif des différents travaux en présentant leurs résultats et spécifités.

Méthode	Modèle	Accuracy	Base de données	Remarque
Yu et al [24]	CNN-6 VGGNet16 ResNet50	81.86 83.54 86.92	ImageCLEF 2015 ImageCLEF 2016	Classification de differents modalité
Kim et al [25]	VGG 16	98	ImageCLEF 2013	Classification des différentes modalité en utilisant un CRM
Remedios et al [26]	Φ -Net	99.27	Source inconnus	Classification des différentes types d'IRM
Chen-Hua et al [29]	Modèle person- nelle	99.5	Base de données de l'hôpital de Taiwan	Classification des différentes modalité (CT de l'abdomen, CT du cer- veau, IRM du cerveau et IRM de la colonne vertébrale)

TAB. 2.2 : Tableau comparatif des différentes approches pour la classification des modalités d'imagerie médicale

2.7 Conclusion

En conclusion de ce chapitre, ce domaine présente des défis spécifiques qu'il faut prendre en compte lors de l'élaboration d'une nouvelle approche. Les principales difficultés incluent la variabilité des images médicales, la complexité des structures anatomiques et la diversité des modalités d'imagerie utilisées.

Les approches basées sur l'apprentissage profond ont montré de grandes promesses dans la classification des modalités d'images médicales et des organes anatomiques. Cependant, elles nécessitent souvent de grandes quantités de données , ce qui peut être difficile à obtenir dans le domaine médical.

C'est dans ce contexte que notre approche basée sur l'apprentissage profond par transfert se justifie. En utilisant des modèles de réseau neuronal pré-entraînés sur des bases d'images génériques, nous pouvons bénéficier des représentations de haut niveau apprises à partir de données massives et diversifiées. Notre démarche, présentée en détail dans le chapitre 3, vise à exploiter cette technique pour résoudre les problèmes de classification des modalités d'images médicales et des organes anatomiques de manière plus efficace et

précise.

Ainsi, en utilisant l'apprentissage profond par transfert, nous cherchons à surmonter les limitations des approches existantes et à fournir une solution capable d'identifier automatiquement la modalité des images médicales et les organes anatomiques. Dans le chapitre suivant, nous détaillerons notre méthodologie et les étapes clés de notre approche, en mettant en évidence les avantages qu'elle offre par rapport aux approches traditionnelles et en justifiant les choix effectués pour notre étude spécifique.

Chapitre 3

Nos Approches

3.1 Introduction

Dans ce chapitre, nous abordons en détail notre approche pour la classification d'imagerie médicale par rapport aux modalités et aux organes anatomiques. Nous commencerons par présenter et expliquer notre approche, ensuite nous parlerons des différentes datasets utilisées, finalement nous discuterons des techniques utilisées pour améliorer nos travaux.

3.2 Méthodologie

Dans ce travail, nous explorons les différents modèles d'apprentissage par transfert pour la classification d'un ensemble de données d'imagerie médicale par rapport aux modalités d'acquisition d'imagerie et aux organes anatomiques, à l'aide des différentes architectures CNN.

Nous étudions dans un premier lieu la capacité de transfert des caractéristiques apprises sur la base d'images ImageNet contenant plus de 14 millions d'images non médicales aux images médicales constituant un ensemble de données de petite taille. Par conséquent, nous avons choisi d'explorer les niveaux de transfert qui peuvent être définis comme un transfert progressif constitué d'un transfert des caractéristiques apprises sur la base d'images ImageNet aux images médicales. Ce type de transfert peut être utile lorsque les réseaux de neurones profonds sont pré-entraînés sur des ensembles de données très différents du domaine en question et lorsque nous disposons d'une base de données assez petite en raison du manque de données médicales pour entraîner un algorithme d'apprentissage profond à partir de zéro. L'exécution d'un algorithme d'apprentissage profond avec des millions de paramètres et moins de données entrave la formation avec un sur-ajustement. Le modèle a donc tendance à échouer lors de la généralisation de l'apprentissage, ce qui donne une faible précision sur l'ensemble de données de test. La régularisation peut réduire la variance élevée dans une certaine mesure, mais la formation d'un cadre d'apprentissage profond à partir de zéro reste interdit. L'augmentation des données peut être utilisée pour accroître davantage la taille de l'ensemble de données (Section 3.4).

Pour atteindre notre objectif nous avons utilisé six architectures CNN : VGG 16, VGG 19, ResNet 50, Inception 3, Xception et NASNet. Tous ces modèles sont déjà pré-entraînés sur les jeux de données ImageNet avant de les entraîner par TL sur notre ensemble de données d'imagerie médicale (voir section 3.4), par la suite les évaluer en utilisant des métriques basées sur la précision, le rappel, le F1-score, l'accuracy.

3.3 Approches utilisées

L'apprentissage profond nécessite un volume de données très important, qui sont difficiles à obtenir avec l'imagerie médicale actuelle. Pour cette raison, nous avons utilisé la technique d'apprentissage par transfert expliquée précédemment. C'est pourquoi nous

avons étudié la classifications de modalité et d'organe anatomique avec six modèle CNN les plus populaires et performants dans l'imagerie médicale (VGG 16, VGG 19, Resnet 50, Inception, XCeption 3 et Nasnet), qui sont pré-entraînés sur la base d'image populaire ImageNet.

3.3.1 Classification des modalités d'imagerie médicale

Pour la classification des modalités d'imagerie médicale, nous avons exploré trois approches pour déterminer la meilleure conduite à suivre et qui nous donnerait des meilleurs résultats, les deux premières approches porte sur l'élaboration de plusieurs modèles CNNs selon deux mode d'apprentissage par transfert (mode extracteur de caractéristique et mode de fine tuning). La troisième approche consiste à combiner différents modèles CNNs pour une meilleure classification.

- **Approche 1:** Pour la première approche, nous utilisons l'apprentissage par transfert en mode extracteur de caractéristiques pour la classification des modalités. Dans ce mode les couches du modèles et les hyperparamètres sont figés et les caractéristiques apprises lors de l'entraînement sur le datasets initial sont transférées directement à la nouvelle tâche. Ce mode est souvent utilisé lorsque les données d'entraînement pour le nouveau problème sont très similaires. Ainsi nous ajustons la partie classification (classifieur Softmax) de nos modèles déjà pré-entraînés avec notre base d'images médicales (voir section 3.4). La méthode proposée porte sur deux étapes, la première consiste à extraire les caractéristiques des images médicales, telles que ces dernières soient passées comme entrées des réseaux neuronaux, pour en sortir un vecteur de caractéristiques qui dépend de ces images. La seconde étape consiste à passer ce vecteur comme entrée du classifieur softmax. En résumé, cette stratégie consiste à passer en entrée pour chaque architecture des images, puis générer des prédictions.
- **Approche 2:** Pour notre deuxième approche, nous avons opté pour le mode de Fine tuning de l'apprentissage par transfert pour la classification des modalités. Dans ce mode, un modèle pré-entraîné est utilisé comme point de départ, mais contrairement au mode extracteur de caractéristiqueion, plusieurs couches du modèle sont réajustées pour être entraînées et ajuster avec de nouvelles données spécifiques à l'imagerie médicale. Les couches qui ne sont pas affinées sont gelées ceci concerne les couches les plus à gauche) pour préserver certaines connaissances de modèle CNN pré entraîné qui a déjà appris sur un large dataset pour extraire des caractéristiques de bas niveaux (générique) qui concerne toute sorte d'image. Cette méthode est composée de trois étapes, la première l'affinage des paramètres des couches non figées, deuxièmement extraction des caractéristiques des images, enfin la dernière étape, consiste en la génération de prédictions pour la classification grâce au classifieur Softmax.
- **Approche 3:** Pour notre troisième approche, nous avons opté pour le mode combinaison (stacking) de plusieurs modèles CNN. Dans ce mode, plusieurs modèles pré-entraînés sont utilisé comme point de départ, nous formons les modèles sur une

base de données qui ne contient que des images d'IRM de différents types (Flair, T1w, T1wCE et T2w), nous obtiendrons ainsi des modèles entraînés et les combineraissons avec un mode combinaison stacking.

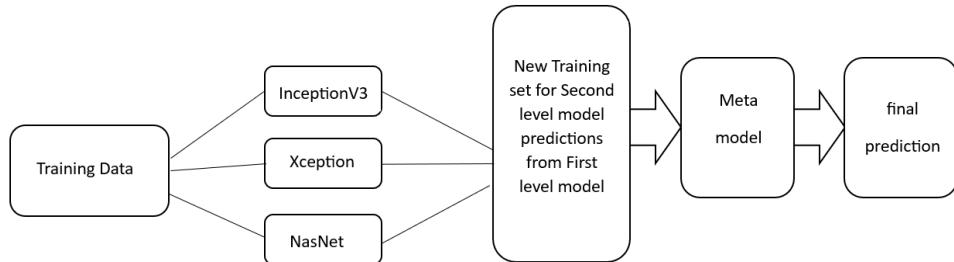


FIG. 3.1 : Schéma synoptique de l'approche 3 basée sur la combinaison des CNNs

Le stacking consiste à entraîner plusieurs modèles de réseaux de neurones indépendamment les uns des autres sur un même jeu de données. Chaque modèle peut avoir une architecture différente avec des paramètres différents. Une fois que les modèles individuels sont entraînés, une couche de fusion est ajoutée au sommet de ces modèles. Cette couche de fusion prend les prédictions de chaque modèle individuel en entrée et apprend à combiner ces prédictions pour produire une prédition finale. La couche de fusion peut être un simple modèle linéaire ou un autre modèle d'apprentissage automatique plus complexe qu'on appelle méta modèle. Le stacking permet de tirer parti de la diversité des modèles individuels en combinant leurs forces et en atténuant leurs faiblesses.

3.3.2 Classification des organes anatomiques

Pour la classification des organes anatomiques nous allons utiliser le mode apprentissage par transfert en mode extracteur de caractéristique, ainsi nous entraînons le classifieur Softmax des modèles CNN pré-entraînés sur notre base d'images médicales (voir section 3.4), pour détecter les différents organes anatomiques à partir du vecteur de caractéristiques extrait des couches convolutives.

Pour résumer, on utilise pour la classification des organes anatomiques la même approche que la première approche de la classification des modalités d'imagerie médicale.

3.4 Augmentation des données

Il existe une approche utilisée pour améliorer l'apprentissage profond sur des petits datasets, elle s'appelle l'augmentation des données ou data augmentation en anglais. Elle consiste à créer de nouvelles données d'apprentissage en appliquant des transformations simples sur les données existantes, ce qui permet d'augmenter la taille de l'ensemble de données disponible pour améliorer l'entraînement du modèle. Elle présente plusieurs avantages pour l'apprentissage profond, elle permet la généralisation des modèles, en exposant

le modèle à une variété de transformations, il devient plus résistant aux variations telles que la rotation, le zoom, le décalage, le changement d'échelle, etc. Les transformations que nous avons utilisées pour l'augmentation des données sont la rotation, le redimensionnement et le recadrage. En utilisant l'augmentation des données, il est ainsi possible d'améliorer les performances des modèles d'apprentissage profond de manière extraordinaire.

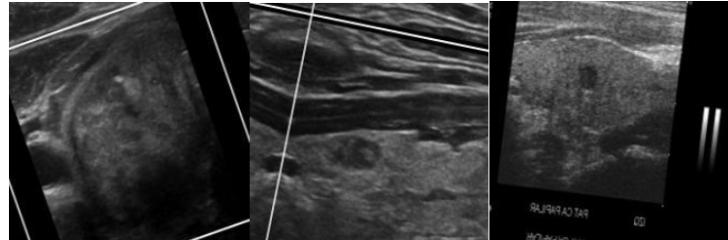


FIG. 3.2 : Exemples d'images transformées par les méthodes d'augmentation de données [41]

3.5 Conclusion

Dans ce chapitre, nous avons présenté nos approches pour la classification des modalités d'images médicales et des organes anatomiques en utilisant l'apprentissage profond par transfert. Nous avons exploré trois différentes approches pour résoudre ces problématiques spécifiques.

La première approche consiste à utiliser le mode d'extraction du transfert d'apprentissage pour la classification des modalités d'images médicales. Nous avons exploité des modèles de réseau neuronal pré-entraînés, tels que VGG 16 et ResNet, pour extraire des caractéristiques de haut niveau à partir des images médicales. Ensuite, ces caractéristiques ont été utilisées pour entraîner des classificateurs spécifiques aux modalités. Cette approche nous permet de bénéficier des connaissances apprises à partir de grandes bases de données génériques.

La deuxième approche que nous avons adoptée est l'utilisation du mode fine-tuning du transfert d'apprentissage pour la classification des modalités d'images médicales. Nous avons commencé par initialiser les modèles de réseau pré-entraînés sur des données génériques, puis nous les avons adaptés et ajustés en utilisant des données médicales spécifiques. Ce processus de fine-tuning nous permet d'ajuster les poids du réseau de manière à mieux représenter les caractéristiques propres aux modalités d'imagerie médicale.

Enfin, notre troisième approche concerne la classification des organes anatomiques en utilisant le mode d'extraction du transfert d'apprentissage. Nous avons appliqué les mêmes principes que pour la classification des modalités, mais cette fois en nous concentrant sur l'identification des organes anatomiques. Les modèles de réseau pré-entraînés ont été utilisés pour extraire des caractéristiques spécifiques aux organes anatomiques, et des classificateurs ont été entraînés en conséquence.

Cette approche présente plusieurs avantages, tels que la capacité à traiter des données médicales variées et complexes, la réduction du nombre d'images conséquent à l'entraî-

Chapitre 3. Nos Approches

nement d'un CNN de zéro et l'amélioration des performances de classification. Elle offre également une plus grande flexibilité pour l'ajout de nouvelles modalités d'imagerie ou de nouveaux organes anatomiques.

Dans le chapitre suivant, nous présenterons les résultats expérimentaux de notre approche, en évaluant les performances de chaque architectures et méthodes utilisé.

Chapitre 4

Expérimentation et Résultats

4.1 Introduction

Dans ce chapitre, nous présenterons en détail les expérimentations que nous avons menées dans le cadre de notre étude, ainsi que les résultats obtenus. Tout d'abord, nous présenterons les différents outils. Ensuite nous présenterons les résultats de nos expérimentations et les analyserons avec des métrique adéquat à l'évaluation de modèles de deep learning.

4.2 Outils utilisés

Pour pouvoir avancer sur notre travail et mener des expérimentations, nous avons besoin d'utiliser plusieurs outils, qui nous facilitent le travail, que ce soit un langage de programmation, des bibliothèques, des technologies de deep learning mais aussi un environnement de développement virtuel.

4.2.1 Python

Python est un langage de programmation interprété. Il a été créé par Guido van Rossum et sa première version a été publiée en 1991. Depuis lors, Python a gagné en popularité et est devenu l'un des langages de programmation les plus utilisés dans divers domaines, tels que le développement web, l'analyse de données, l'intelligence artificielle et l'automatisation de tâches. Les caractéristiques clés de Python sont sa syntaxe simple et lisible, qui favorise la lisibilité et la facilité de compréhension du code. Il encourage également une approche modulaire de la programmation, ce qui facilite la réutilisation de code et la collaboration entre développeurs. Python offre une vaste bibliothèque standard qui couvre un large éventail de fonctionnalités, allant des opérations de base telles que la manipulation de chaînes de caractères et la gestion des fichiers, jusqu'aux fonctionnalités avancées comme l'analyse de données, le développement web et l'intelligence artificielle. De plus, la communauté Python est très active et propose de nombreuses bibliothèques tierces spécialisées, telles que NumPy, pandas, scikit-learn, TensorFlow et PyTorch, qui étendent encore les capacités de Python dans différents domaines.

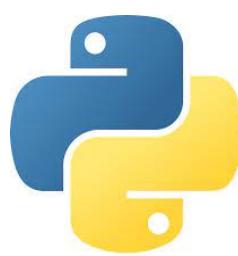


FIG. 4.1 : Logo de python

4.2.2 TensorFlow

TensorFlow est une bibliothèque open source d'apprentissage automatique développée par Google. Elle est spécialement conçue pour créer, former et déployer des modèles d'apprentissage automatique. TensorFlow est devenu l'un des outils les plus populaires dans le domaine de l'intelligence artificielle en raison de sa flexibilité, de sa puissance de calcul et de sa facilité d'utilisation. TensorFlow offre une large gamme d'outils et de fonctionnalités pour développer des modèles d'apprentissage automatique. Il prend en charge différents types de réseaux neuronaux, y compris les réseaux de neurones convolutifs (CNN). TensorFlow fournit également des modules pour la préparation des données, la visualisation des modèles, l'optimisation des hyperparamètres et l'évaluation des performances des modèles. En outre, TensorFlow offre des interfaces compatibles avec différents langages de programmation, notamment Python et C++. Il est largement utilisé dans l'industrie et la recherche pour résoudre des problèmes d'apprentissage automatique dans divers domaines tels que la vision par ordinateur, le traitement du langage naturel, la reconnaissance vocale, la prédiction et l'analyse de données.



FIG. 4.2 : Logo de TensorFlow

4.2.3 Keras

Keras est une bibliothèque open source d'apprentissage automatique (machine learning) écrite en Python. Elle fournit une interface conviviale et haut niveau pour développer et entraîner des modèles d'apprentissage automatique. Keras est conçu pour être simple, intuitif et extensible, ce qui en fait un outil populaire pour les débutants et les experts en apprentissage automatique. L'objectif principal de Keras est de permettre aux utilisateurs de définir et de configurer rapidement des modèles d'apprentissage automatique, en mettant l'accent sur la facilité d'utilisation. Il offre une abstraction de haut niveau pour définir des réseaux neuronaux et des couches de neurones, en permettant aux utilisateurs de créer des architectures de modèles complexes avec un minimum de code. Keras prend en charge différents types de modèles, y compris les réseaux de neurones profonds (DNN), les réseaux de neurones récurrents (RNN) et les réseaux de neurones convolutifs (CNN). Une caractéristique clé de Keras est sa compatibilité avec plusieurs bibliothèques d'apprentissage automatique sous-jacentes, y compris TensorFlow.. Cela permet aux utilisateurs de profiter de la puissance et de la flexibilité de ces bibliothèques tout en bénéficiant de la simplicité et de la convivialité de Keras. TensorFlow est la bibliothèque d'apprentissage automatique la plus couramment utilisée avec Keras, et les deux sont souvent utilisés

ensemble pour développer des modèles d'apprentissage automatique. En plus de faciliter la création de modèles, Keras offre également des fonctionnalités pour l'entraînement des modèles, l'évaluation des performances, la régularisation, l'optimisation des hyperparamètres et la visualisation des résultats. Il est également extensible, ce qui signifie que les utilisateurs peuvent ajouter leurs propres couches de neurones, fonctions d'activation ou fonctions de coût personnalisées pour répondre à leurs besoins spécifiques. Keras est largement utilisé dans la communauté de l'apprentissage automatique en raison de sa simplicité, de sa flexibilité et de sa puissance. Que ce soit pour des projets académiques, des applications industrielles ou des projets de recherche, Keras offre un cadre convivial pour développer rapidement des modèles d'apprentissage automatique performants.



FIG. 4.3 : Logo de Keras

4.2.4 Environnements de développement

Notebook jupyter

Les notebooks Jupyter sont des environnements de développement interactifs qui permettent de créer et d'exécuter du code en utilisant des cellules, combinant du code, des textes explicatifs et des visualisations. Ils tirent leur nom des trois langages de programmation qu'ils prennent en charge : Julia, Python et R (d'où le nom "Ju"- "py"- "ter"). Les notebooks Jupyter sont basés sur le projet open source Jupyter, autrefois connu sous le nom d'IPython Notebook. Ils sont largement utilisés dans les domaines de la science des données, de l'apprentissage automatique et de la recherche, mais peuvent également être utilisés pour des tâches de développement général. Les notebooks Jupyter sont accessibles via un navigateur web et permettent aux utilisateurs de créer des cellules de code et d'y exécuter du code Python (ou d'autres langages supportés). Chaque cellule peut être exécutée indépendamment, ce qui permet de tester et de modifier le code de manière interactive. Les résultats des cellules de code, tels que les sorties, les graphiques ou les tableaux, sont affichés directement sous la cellule correspondante.

Google Colab

Google Colab, abréviation de Google Collaboratory, est un environnement de développement et de partage de code en ligne basé sur le cloud. Il s'agit d'un service gratuit proposé par Google qui permet d'exécuter du code Python, notamment des notebooks Jupyter, directement dans le navigateur web, sans nécessiter d'installation ni de configuration complexe. L'un des principaux avantages de Google Colab est qu'il fournit un accès gratuit à des ressources de calcul puissantes, y compris des unités de traitement graphique (GPU) et des unités de traitement tensoriel (TPU). Cela permet aux utilisateurs de bénéficier de la puissance de calcul nécessaire pour exécuter des tâches intensives

en termes de calcul, telles que l’entraînement de modèles d’apprentissage automatique ou l’analyse de données volumineuses. Google Colab offre également une intégration étroite avec d’autres services de Google, tels que Google Drive, ce qui facilite le chargement et le partage de données et de fichiers. Les notebooks créés dans Colab sont automatiquement sauvegardés dans Google Drive, ce qui permet d’y accéder facilement à partir de différents appareils et de les partager avec d’autres utilisateurs, et offre des fonctionnalités avancées telles que l’exécution de commandes système, la visualisation de données, l’utilisation de bibliothèques populaires d’apprentissage automatique telles que TensorFlow et PyTorch, ainsi que la collaboration en temps réel avec d’autres utilisateurs. Google Colab est devenu un outil populaire parmi les développeurs, les chercheurs et les étudiants en apprentissage automatique en raison de sa simplicité d’utilisation, de sa flexibilité et de sa disponibilité gratuite. Il permet de travailler sur des projets de data science et d’apprentissage automatique sans avoir à se soucier de la configuration matérielle ou des contraintes de ressources, tout en favorisant la collaboration et le partage de connaissances.

4.3 Métriques d’évaluation des modèles de classification

Il existe différentes approches d’évaluation des modèles d’apprentissage profond, permettant de vérifier la fiabilité et la pertinence des modèles. Les métriques d’évaluation les plus utilisés sont : Matrice de confusion, Accuracy, Precision (Précision), Rappel (Recall) et F1-score

4.3.1 Matrice de confusion

La matrice de confusion ou matrice d’erreur est l’un des concepts clés lorsque nous parlons de problèmes de classification. Cette matrice est un tableau à deux dimensions (« réelle » et « prédictive ») et des ensembles de « classes » dans les deux dimensions. Nos classifications réelles sont des colonnes et celles prédictives sont des lignes comme le montre le tableau ci-dessous :

		Réelle	
		Positives (1)	Négatives (0)
Prédite	Positives (1)	TP	FP
	Négatives (0)	FN	TN

FIG. 4.4 : Matrice de confusion

Presque toutes les mesures de performance sont basées sur la matrice de confusion et

les chiffres qu'elle contient.

- **Vrai positifs (TP)** : les vrais positifs sont les cas où la classe réelle de la donnée était 1 (True) et la prédiction est également 1 (True).
- **Vrai négatif (TN)** : les vrais négatifs sont les cas où la classe réelle de la donnée était 0 (Faux) et la prédiction est également 0 (Faux).
- **Faux positifs (FP)** : les faux positifs sont les cas où la classe réelle de la donnée était 0 (False) et la prédiction est 1 (True). « Faux » est dû au fait que le modèle a prédit de manière incorrecte et « positive » parce que la classe prédite était positive (1).
- **Faux négatif (FN)** : les faux négatifs sont les cas où la classe réelle de donnée était à 1 (Vrai) et la classe prédite est à 0 (Faux). « Faux » est dû au fait que le modèle a prédit de manière incorrecte et « négative » parce que la classe prédite était négative (0).

4.3.2 Accuracy

L'Accuracy est une mesure d'évaluation courante pour les problèmes de classification. Elle est calculée comme un nombre de prédictions correctes divisé par le nombre total d'échantillons. Lorsque nous le multiplions par 100, nous obtenons une précision en pourcentage. L'Accuracy est une bonne mesure lorsque les classes de variables cibles dans les données sont presque équilibrées.

$$Accuracy = \frac{\text{Nombre de predictions correctes}(TP + TN)}{\text{Nombre total d'échantillons}(TP + FP + FN + TN)} \quad (4.1)$$

4.3.3 Précision

La précision est une métrique très utile et elle contient plus d'informations que l'Accuracy. Essentiellement, avec précision, nous répondons à la question : « Lorsque le modèle prédit positif, à quelle fréquence est-il correct ? ». Elle est calculée pour chaque classe séparément avec la formule :

$$P = \frac{TP}{TP + FP} \quad (4.2)$$

4.3.4 Rappel

Rappel (Recall) peut être décrit comme la capacité du classificateur à trouver tous les échantillons positifs. Avec cette métrique, nous essayons de répondre à la question : "Quelle proportion de positifs réels a été correctement identifiée ?" Il est défini comme la

fraction d'échantillons d'une classe spécifique qui sont correctement prédits par le modèle ou mathématiquement :

$$R = \frac{TP}{TP + FN} \quad (4.3)$$

Différence entre Précision et Recall dans le cadre de classification : Recall nous donne des informations sur les performances d'un classifieur par rapport aux faux négatifs (combien en avons-nous manqué), tandis que la précision nous donne des informations sur ses performances par rapport aux faux positifs (combien en avons-nous attrapés).

4.3.5 F1-score

Moyenne harmonique de la précision et du rappel. Mesure la capacité du système à donner toutes les solutions pertinentes et à refuser les autres.

$$FScore = \frac{2PR}{P + R} \quad (4.4)$$

4.4 Base d'images utilisées

L'une des composantes essentielles pour la création et l'entraînement de modèle d'apprentissage profond, c'est une base d'images. Pour cela, le choix et l'élaboration d'un Datasets est une étape importante dans notre travail de recherche d'identification des modalités d'imagerie médicale.

Ainsi pour nous nous sommes reposés sur le site Kaggle, une plateforme en ligne populaire dédiée à la science des données et aux compétitions de machine learning. Elle a été fondée en 2010 et a été acquise par Google en 2017. Kaggle offre une communauté de scientifiques des données et propose diverses ressources, notamment des ensembles de données, des noyaux (notebooks) de code, des tutoriels et des compétitions de machine learning et de deep learning.

Nous avons créé notre base d'images à partir de neuf bases de données disponibles en public sur le site Kaggle. D'abord on a traité nos jeux de données, en éliminant les images inutilisables ou au format indésirable mais aussi en enlevant les données en extra non nécessaires à notre travail. Ensuite nous avons classé nos bases d'images en fonction de l'objectif voulu, pour la classification des modalités, nous avons classé les bases par rapport aux modalités, par exemple, toutes les bases d'imagerie médicale par IRM dans un même fichier. Et pour la classification des organes anatomiques, nous avons fait la même chose par rapport aux classes voulues et choisies.



FIG. 4.5 : Image médicale du cerveau acquis par IRM de dataset utilisée [30]

Base d'images médicales par IRM

La base de données utilisée pour les images par IRM, est une base d'images publique disponible sur Kaggle au nom de Brain Tumor MRI Dataset, elle contient 7023 images médicales du cerveau acquises par IRM [30].

On a utilisé une autre base spécifiquement pour l'approche de combinaison des modèles (troisième approche), le dataset RSNA-MICCAI Brain Tumor Radiogenomic Classification[32], qui se compose de données DICOM converties au format PNG. La taille de l'ensemble de données d'origine était de 128 Go, mais pour faciliter l'analyse, un prétraitement a été effectué pour réduire la taille à 5,2 Go. La conversion de fichiers DICOM au format PNG peut demander beaucoup de temps et de calculs, en particulier compte tenu de la grande taille de l'ensemble de données. Pour résoudre ce problème, une machine virtuelle (VM) Google Cloud Platform (GCP) a été utilisée pour convertir les données DICOM en PNG tout en préservant la taille des images originales.

Le jeu de données est divisé en trois groupes : entraînement, validation (public) et test (privée). Dans ces dossiers on trouve quatre sous-dossiers, chacun correspondant à un type spécifique d'analyse structurelle multi paramétrée au format DICOM (mpMRI). Les IRM incluses sont les suivantes : récupération par inversion atténuée par fluide (FLAIR), pré-contraste pondéré en T1 (T1w), post-contraste pondéré en T1 (T1wCE) et pondéré en T2 (T2w).

Base de données	Division	Classe	Nombre d'images	Taille des images redimensionnées
RSNA MICCAI PNG	Train	FLAIR	50682	224x224x3
		T1w	55440	
		T1wCE	68012	
		T2w	67722	
		Total	241856	
	Valid	FLAIR	3000	224x224x3
		T1w	3000	
		T1wCE	3000	
		T2w	300	
		Total	12000	
	Test	FLAIR	7926	224x224x3
		T1w	8791	
		T1wCE	10482	
		T2w	9863	
		Total	37067	
	Total image		290923	

TAB. 4.1 : Tableau de dataset RSNA-MICCAI Brain Tumor Radiogenomic Classification

Base d'images médicales par Rayon X

Pour les images de rayon x, nous avons utilisé trois base d'images disponible aussi sur kaggle, la première s'appelle Chest X-ray Images, elle contient 5856 images [33], la deuxième base d'images s'appelle Chest Xray Masks and Labels, qui contient 1600 images médicales de torse humain [34]. la dernière base s'appelle figure-detection, qui contient 5567 images médicales de main [35].

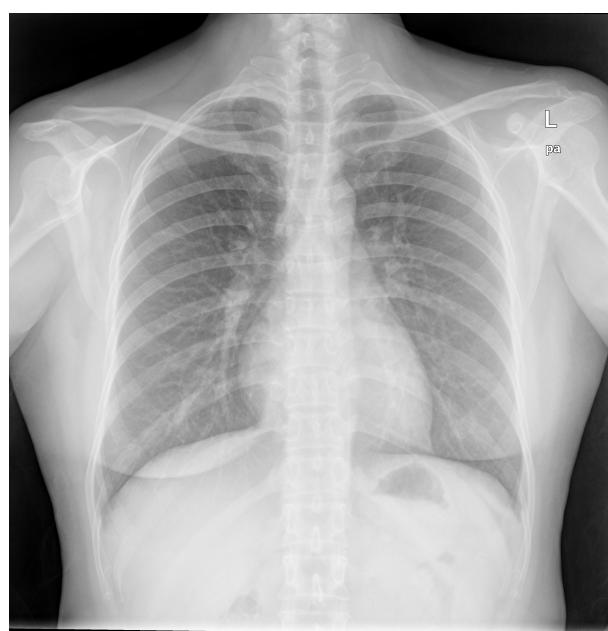


FIG. 4.6 : Image médicale du torse acquis par Rayon X de dataset utilisée [33]



FIG. 4.7 : Image médicale de main acquis par Rayon X de dataset utilisée [35]

Base d'images médicales par Ultrasound

Nous avons utilisé deux base d'images par ultrasound en libre utilisation, l'une contient 780 images médicales de poitrine de femmes [36], l'autre 1336 images d'échographie de la tête de fœtus [37]



FIG. 4.8 : Image médicale de poitrine de femme par Ultrasound de dataset utilisée [36]



FIG. 4.9 : Image médicale de la tête de fœtus par Ultrasound de dataset utilisée [37]

Base d'images médicales par CT-Scan

Finalement pour les images médicales acquises par ct scan, nous avons téléchargé deux base de données, la première contient 988 images du torse [38], et l'autre plus de 12000 images médicales des reins [39].

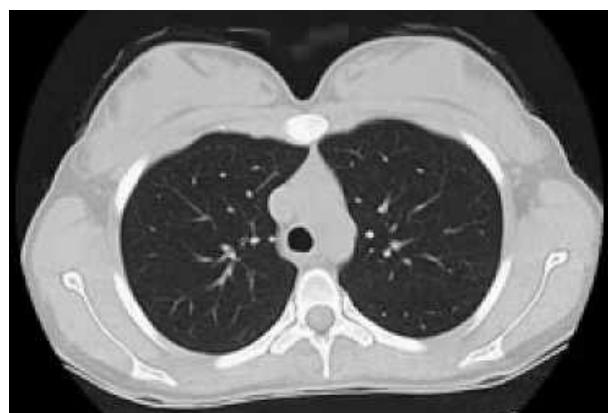


FIG. 4.10 : Image médicale du torse acquis par CT scan [38]



FIG. 4.11 : Image médicale du rein acquis par CT scan [39]

4.5 Expérimentations et résultats

4.5.1 Expérimentations et résultats de la classification des modalités d'imagerie médicale

Pour étudier la pertinence de l'apprentissage profond pour la classification des modalités d'acquisition d'imagerie médicale, nous avons comparé plusieurs scénarios, différentes approches avec plusieurs modèles CNN en utilisant le transfert d'apprentissage sur notre base d'images.

Résultats des modèles CNN en mode extracteur de caractéristiques

Nous avons entraîné six modèles CNN (VGG 16, VGG 19, Resnet 50, Inception 3, Xception, NASNet) déjà pré-entraîné sur la base d'images ImageNet en mode extraction du transfert d'apprentissage sur notre Base d'images médicales avec augmentation d'image, qui contient quatre classes (IRM, ultrasound, Ct-scan et rayon x).

Modèle	Accuracy	F1 score	Précision	Recall
VGG 16	0.96	0.96	0.96	0.96
VGG 19	0.94	0.95	0.95	0.94
Resnet 50	0.81	0.81	0.82	0.80
Inception 3	0.97	0.97	0.97	0.97
XCEPTION	0.976	0.976	0.976	0.976
NASNet	0.99	0.99	0.99	0.99

TAB. 4.2 : Tableau comparatif des résultats des différents modèles en mode extraction pour la classification des modalités

Résultats des modèles CNN en mode Fine Tuning

Nous avons entraîné cinq modèles CNN (VGG 16, VGG 19, Resnet 50, Inception 3, XCEPTION) déjà pré-entraînés sur la base d'images ImageNet en mode fine tuning du transfert d'apprentissage sur notre Base d'images médicales, qui contient quatre classes (IRM, ultrasound, Ct-Scan et rayon x). Nous avons entraîné nos modèles sur 5 époques avec un batch size de 100. Pour chaque modèle nous avons figé un certain pourcentage de couches (qui se trouve le plus à gauche, les plus proches à l'image initiale) :

- **VGG 16** : contient 5 blocs de convolution, nous avons ainsi générée deux modèles à partir de ce dernier, l'un on l'a figé à 80% (4 blocs) des ses blocs et l'autre à 60% (3 blocs).
- **VGG 19** : contient 5 blocs, nous avons ainsi générer deux modèles à partir de ce dernier, l'un on l'a figé à 80% (4 blocs) de ses blocks et l'autre à 60% (3 blocs).
- **Resnet 50** : contient 50 couches, nous avons ainsi générer un modèle à partir de ce dernier, avec un pourcentage de 80% du modèle figé (40 couches).
- **Inception 3** : contient 48 couches, nous avons ainsi générer deux modèles à partir de ce dernier, l'un on l'a figé 90% (43 couches) des ses couches et l'autre à 80% (38 couches).
- **Xception** : contient 71 couches, nous avons ainsi générer deux modèles à partir de ce dernier, l'un on l'a figé 90% (64 couches) des ses couches et l'autre à 80% (57 couches).

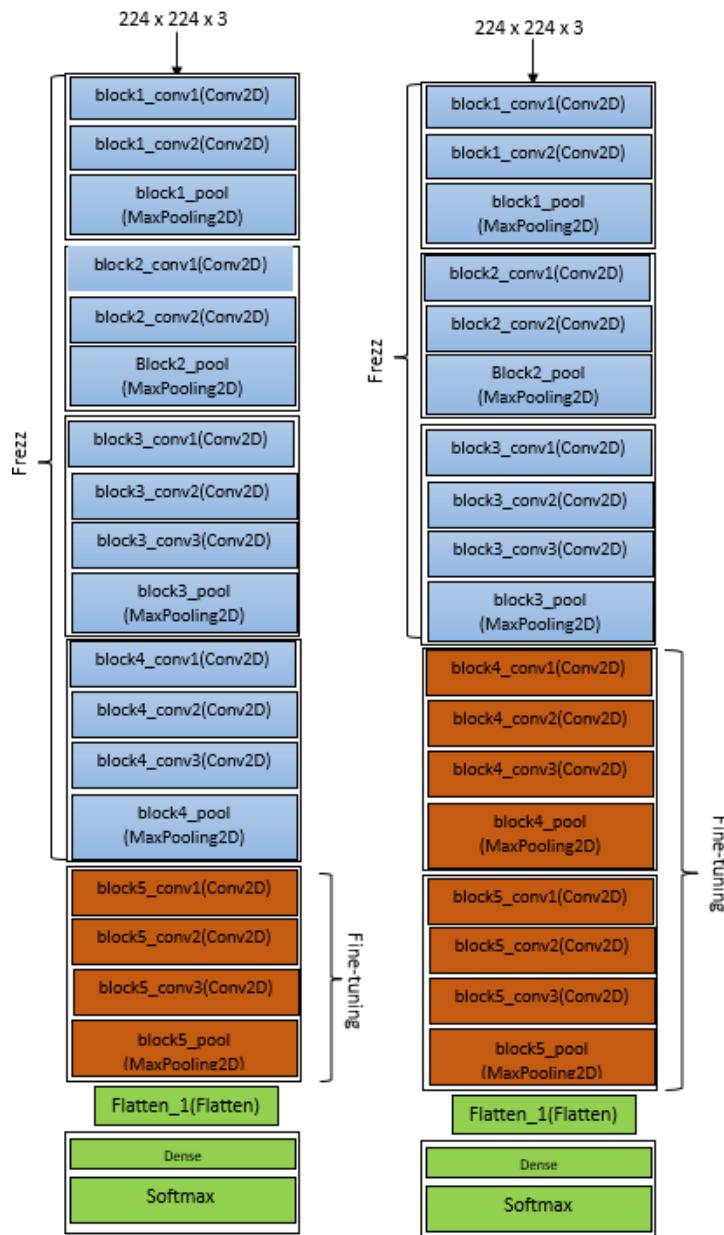


FIG. 4.12 : Organigramme du modèle VGG 16 basé sur les différents niveaux de Fine-tuning

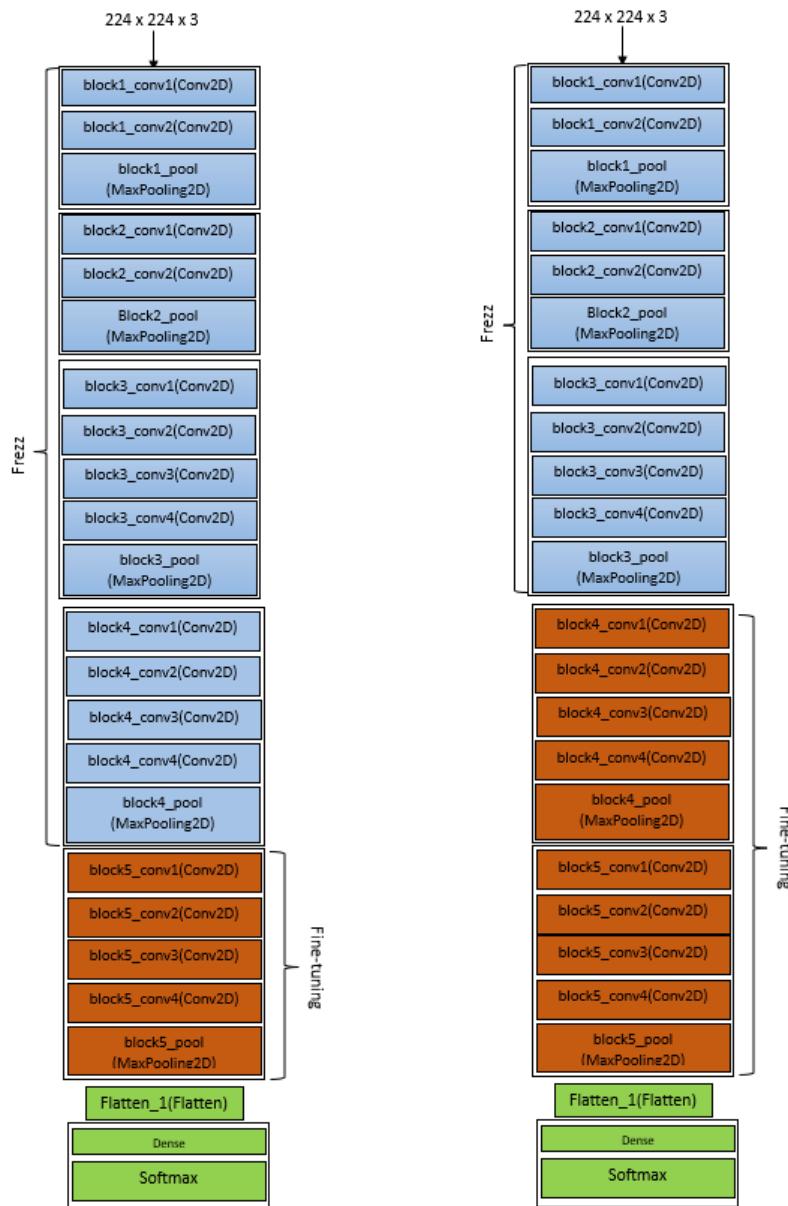


FIG. 4.13 : Organigramme du modèle VGG 19 basé sur les différents niveaux de Fine-tuning

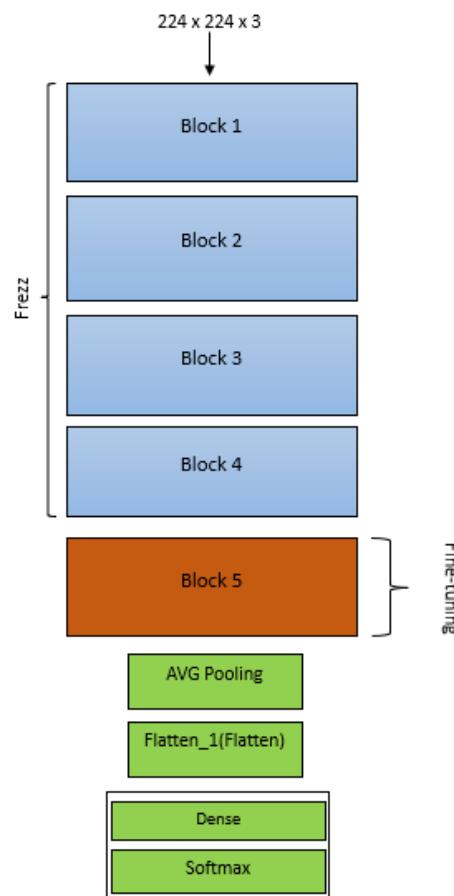


FIG. 4.14 : Organigramme du modèle Resnet 50 basé sur le niveau de Fine-tuning

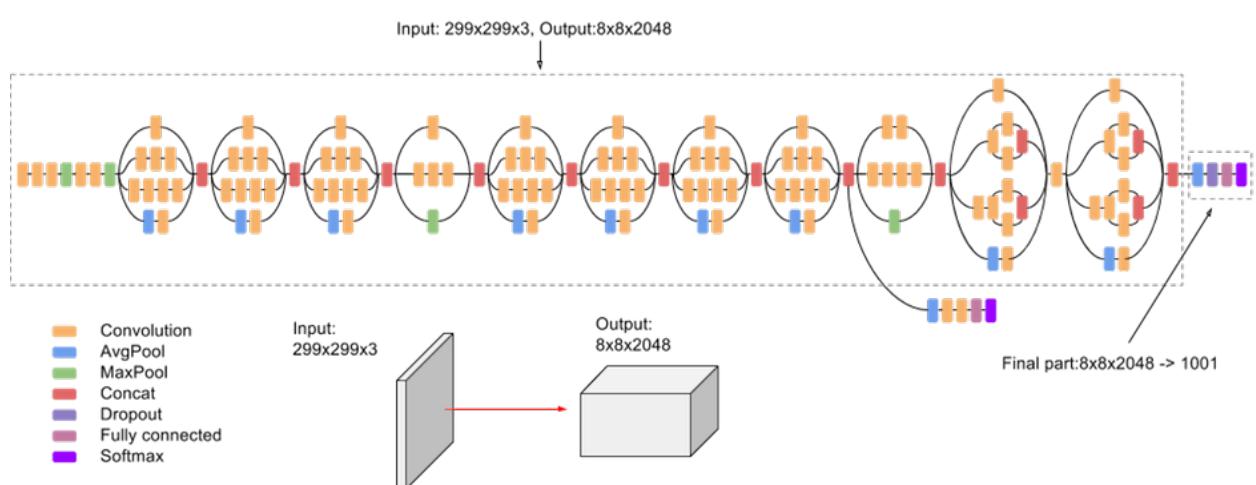


FIG. 4.15 : Organigramme du modèle Inception 3

Modèle	Accuracy	F1 score	Précision	Recall
VGG 16 (60%)	0.42	0.41	0.37	0.47
VGG 16 (80%)	0.94	0.94	0.95	0.94
VGG 19 (60%)	0.46	0.46	0.46	0.46
VGG 19 (80%)	0.96	0.96	0.96	0.96
Resnet 50 (80%)	0.84	0.84	0.87	0.82
Inception 3 (80%)	0.99	0.99	0.99	0.99
Inception 3 (90%)	0.99	0.99	0.99	0.99
Xception (80%)	0.99	0.99	0.99	0.99

TAB. 4.3 : Tableau comparatif des résultats des différents modèles en mode fine tuning pour la classification des modalités

Résultats de la troisième approche (combinaison de classifieurs)

Nous avons entraîné six modèles CNN (VGG 16, VGG 19, Resnet 50, Inception 3, Xception, NASNet) déjà pré-entraînés sur la base d'images ImageNet en mode extraction du transfert d'apprentissage sur une Base RSNA-MICCAI Brain, d'images IRM de différentes modalités d'images (quatre classes Flair , T1w, T1wCE , T2w : voir chapitre 3.4.1), avec augmentation de données. Après avoir entraîné nos différentes architectures et les avoir évalué, nous avons choisi les trois architectures les plus performantes (Inception 3, Xception et NASNet), pour les combiner en mode stacking afin d'avoir un modèle plus performant qui a nettement augmenté les résultats.

Modèle	Accuracy	F1 score	Précision	Recall
VGG 16	0.58	0.37	0.75	0.24
VGG 19	0.50	0.36	0.73	0.24
Resnet 50	0.63	0.51	0.70	0.41
Inception 3	0.84	0.85	0.89	0.81
Xception	0.88	0.88	0.89	0.87
NASNet	0.86	0.87	0.91	0.82
Combinaison Stacking	0.91	0.92	0.95	0.89

TAB. 4.4 : Tableau comparatif des résultats des différents modèles en mode fine tuning pour la classification des modalités et combinaison des modèles

4.5.2 Expérimentations et résultats de la classification des organes anatomique

Pour la classification des organes anatomiques, nous avons fait des tests sur six modèles CNN, tous déjà pré-entraînés sur la base d'images ImageNet. Pour avoir de meilleurs résultats nous avons utilisé la technique d'apprentissage par transfert en mode extraction, qui nous permet de profiter de la puissance de la base d'image ImageNet. Notre travail a été orienté sur une base de 8 classes d'images médicale de torse humain acquis par scanner, de rein d'humain acquis par scanner, de cerveau acquis par IRM, de torse acquis par IRM, de poitrine de femme acquis par ultrasound, de tête de foetus acquis par ultrasound et enfin de torse et de main acquis tout les deux par rayon x.

Modèle	Accuracy	F1 score	Précision	Recall
VGG 16	0.94	0.95	0.94	0.93
VGG 19	0.64	0.78	0.51	0.38
Resnet 50	0.66	0.89	0.54	0.39
Inception 3	0.98	0.97	0.99	0.96
XCEPTION	0.98	0.98	0.99	0.97
NASNet	0.98	0.98	0.99	0.96

TAB. 4.5 : Tableau comparatif des résultats des différents modèles en mode extraction pour la classification des organes anatomiques

4.6 Discussion des résultats

La classification des modalités d'imagerie médicale a été effectuée sur trois approches différentes, sur les deux méthodes nous avons utilisé l'apprentissage par transfert mais avec deux modes différent, et pour la dernière nous avons utilisé le mode combinaison stacking de l'apprentissage profond. Des études antérieures sur l'impact de la taille de la base d'images dans l'apprentissage profond à l'aide de CNN ont montré que pour bien entraîner un modèle et pour qu'il extrait efficacement les caractéristiques d'une image, doit être entraîné sur une base d'image grande. Par conséquent, nous avons utilisé pour les trois approches le transfert par apprentissage, qui nous permet de profiter de modèle déjà pré-entraîné sur la base d'images ImageNet contenant plus de 14 millions d'images.

La comparaison des performances de nos six architectures pour en sélectionner la meilleure dans notre cas est basée sur la valeur de F1-score, cela est dû aux coûts des faux positifs et des faux négatifs qui diffèrent en nombre, on obtient des faux positifs supplémentaires (fausses alertes) plutôt que de sauvegarder des faux négatifs.

Dans la première approche, nous pouvons voir que le modèle NASNet avec le classifieur softmax, a obtenu de meilleure résultats (avec un F1-score de 0.99) que tous les autres modèles, ainsi nous pouvons arriver à la conclusion suivante, NASNet en mode extraction de données n'a pas eu de challenge.

Dans la deuxième approche, nous pouvons remarquer que les résultats ont soit augmenté soit diminué en fonction du nombre de couches figées par la mode fine tuning. L'architecture Xception avec un taux de fine tuning de 80% a eu le meilleure taux et à améliorer significativement et améliorer son F1-score par rapport à la première approche en mode extraction, ceci nous montre que le bon choix de niveaux de fine tuning peut résoudre le problème de la petite taille de la base d'images.

La troisième approche, qui utilise la combinaison de trois architectures CNN (Inception 3, Xception et NASNet), qui ont eu les meilleurs résultats sur la base d'imagerie médicale des différent type d'IRM, a montré des résultats très optimiste, le modèle qui combine les trois modèles a améliorer le F-score du modèle le plus performant (Xception) de 0.4, ceci nous amène a la conclusion suivante, la combinaison stacking en DL permet une nette augmentation des résultats et score de nos modèles dans l'imagerie médicale.

Après avoir étudié les résultats des nos différentes approches, nous pouvons arriver à la conclusion suivante, la base d'image des différentes modalité d'imagerie médicale ne pose pas de challenge par rapport a la base d'image des différents types d'IRM,et que la classification des modalités d'imagerie médicale n'est pas un problème de l'apprentissage profond, la différence de caractéristiques entre les images est assez suffisante pour permettre aux architectures les plus performantes d'atteindre presque la perfection et le F1-score de 1, et que la problématique se trouve plus précisément en la classification des différents types du même modalité, car les différents images ont des caractéristiques très similaire et peux poser un challenge aux différentes architecture d'apprentissage profond

Pour la classification des organes anatomiques, nous pouvons remarquer que deux modèles (Xception et NASNet) avec le classifieur softmax ont eu un score de 0.98. Nous pouvons envisager d'essayer d'autres approches pour la classification d'organes, mais pour pouvoir explorer d'autres approches, il faut une base d'images d'organes anatomiques conséquentes pour pouvoir bien spécialiser nos architectures et améliorer les résultats.

Nous reconnaissons que, nous aurions pu comparer nos performances avec les performances de la littérature et de la communauté scientifiques. Mais ceci est principalement dû à l'unicité de la base d'images .

4.7 Conclusion

Nous avons présenté dans ce chapitre les résultats de notre expérimentation basée sur les différentes approches proposées pour la classification des modalités d'images médicales et des organes anatomiques en utilisant l'apprentissage profond par transfert. Les résultats obtenus démontrent l'efficacité de notre approche dans cette tâche complexe.

Nous avons exploré quatre approches distinctes et les résultats obtenus sont promet-

teurs. La première approche, basée sur le mode d'extraction du transfert d'apprentissage pour la classification des modalités, a donné des performances élevées en termes de précision. Cette approche permet de bénéficier des connaissances préalablement acquises par les modèles pré-entraînés, améliorant ainsi la capacité de classification des modalités d'images médicales.

La deuxième approche, basée sur le mode de TL par fine-tuning pour la classification des modalités, a montré une nette amélioration par rapport à la première approche. En adaptant spécifiquement les modèles pré-entraînés aux données médicales, nous avons pu obtenir des performances encore meilleures en termes de précision, de rappel et de mesure F1. Cette approche permet une meilleure adaptation des modèles aux caractéristiques spécifiques des modalités d'imagerie médicale.

Pour la troisième approche, on a utilisé le mode de combinaison (stacking) de l'apprentissage profond pour la classification des différents types d'imageries d'IRM, qui permet de générer un nouveau modèle à partir de différents modèle CNN déjà entraînés. Nous avons ainsi pu constater l'efficacité de cette technique qui a permis d'améliorer nettement les résultats des autres architectures sur la base d'images de différents types d'IRM.

En ce qui concerne, la quatrième approche, qui se concentre sur la classification des organes anatomiques en utilisant le mode de TL par extraction de caractéristiques, les résultats obtenus sont également encourageants. Les modèles pré-entraînés ont réussi à extraire des caractéristiques distinctives de différentes structures anatomiques, conduisant à des performances précises pour la classification des organes.

Ces résultats démontrent que notre approche basée sur l'apprentissage profond par transfert est efficace pour la classification des modalités d'images médicales et des organes anatomiques. Elle offre des performances élevées et une capacité à traiter la variabilité des données médicales, ce qui est essentiel pour des applications cliniques précises.

Dans l'ensemble, notre approche représente une avancée significative dans la classification des modalités d'images médicales et des organes anatomiques en utilisant l'apprentissage profond par transfert. Ces résultats ouvrent de nouvelles perspectives pour l'automatisation et l'amélioration des outils d'analyse d'images médicales, contribuant ainsi à l'amélioration des soins de santé et à la prise de décision médicale.

Conclusion générale

Conclusion générale

Dans le cadre de ce mémoire de Master, nous avons exploré l'utilisation de l'apprentissage profond par transfert pour l'identification automatique de la modalité d'images médicales et des organes anatomiques. Notre objectif était de comparer différents modèles de réseaux de neurones convolutifs (CNN) déjà existants dans le domaine du deep learning afin de déterminer celui offrant les meilleures performances pour cette tâche spécifique.

Nos résultats expérimentaux ont démontré que l'utilisation de l'apprentissage profond par transfert est une approche prometteuse pour résoudre le problème d'identification automatique des modalités d'images médicales et des organes anatomiques. En capitalisant sur les connaissances préalablement acquises par les modèles CNN pré-entraînés sur des tâches de classification d'images générales, nous avons pu obtenir des performances significatives.

Parmi les modèles CNN comparés, nous avons observé que certains ont montré une capacité supérieure à généraliser et à extraire des caractéristiques discriminantes pertinentes pour la tâche d'identification des modalités d'images médicales et des organes anatomiques. Cependant, il convient de souligner que le choix du modèle dépend également des caractéristiques spécifiques de l'ensemble de données et des contraintes de la tâche.

En termes de perspectives, plusieurs axes de recherche peuvent être envisagés, l'extension de notre étude à d'autres modalités d'imagerie médicale ou à des ensembles de données plus vastes et variés permettrait de mieux évaluer la robustesse et la généralisation des modèles.

Par ailleurs, l'intégration de techniques de renforcement et d'attention pourrait être explorée pour améliorer la localisation précise des organes anatomiques dans les images médicales. Ces approches pourraient aider à réduire les erreurs de classification et à fournir une interprétation plus fine des résultats.

Enfin, la collaboration avec des experts cliniciens et des radiologues pourrait contribuer à une meilleure intégration des modèles CNN dans la pratique médicale. Le développement de solutions conviviales et interprétables, permettant de visualiser les régions d'intérêt et d'expliquer les décisions prises par les modèles, est essentiel pour favoriser leur adoption et leur confiance par les professionnels de la santé.

En conclusion, notre étude comparative des différents modèles CNN existants pour l'identification automatique de la modalité d'images médicales et des organes anatomiques en utilisant l'apprentissage profond par transfert a ouvert de nouvelles perspectives pour l'amélioration du diagnostic médical assisté par ordinateur. Ces résultats encourageants ouvrent la voie à des développements futurs dans ce domaine, contribuant ainsi à une meilleure prise en charge des patients et à l'avancement de la médecine.

Bibliographie

Bibliographie

- [1] *Le traitement numérique des images.* (s. d.). *Images des mathématiques*. Accédé en Juillet 2023
- [2] Y.LeCun et al, « Gradient-based learning applied to document recognition, » *Proceedings of the IEEE*, vol. 86, pp., 1998 . Accédé en Juillet 2023
- [3] Y.LeCun et al, « Handwritten digit recognition with a backpropagation network, » in *Advances in neural information processing systems*, 1990 . Accédé en Juillet 2023
- [4] KRIZHEVSKY, Alex, SUTSKEVER, Ilya, HINTON et E.Geoffrey, « Imagenet classification with deep convolutional neural networks, » in *Advances in neural information processing systems*, 2012 . Accédé en Juillet 2023
- [5] K.Simonyan et A.Zisserman « Very deep convolutional networks for large-scale image recognition, » *arXiv preprint arXiv*, 2014 . Accédé en Juillet 2023
- [6] HE, Kaiming, ZHANG, Xiangyu, REN, Shaoqing et al. « Going deeper with convolutions, » in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015 . Accédé en Juillet 2023
- [7] HE, Kaiming, ZHANG, Xiangyu, REN, Shaoqing et al. « Deep residual learning for image recognition, » in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016 . Accédé en Juillet 2023
- [8] AI in Medical Imaging Informatics : current Challenges and Future Directions *Visual Interpretation of Convolutional Neural Network Predictions in Classifying Medical Image Modalities*, 2019 . Accédé en Juillet 2023
- [9] Incheol Kim ; Sivaramakrishnan Rajaraman ; Sameer Antan, *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS*, VOL. 24, NO. 7, 2020 . Accédé en Juillet 2023
- [10] Définitions : image - Dictionnaire de français Larousse. (s. d.). *Larousse.fr : encyclopédie et dictionnaires gratuits en ligne*. Accédé en Juillet 2023
- [11] 10 types de format d'image et leur utilisation. (2022, 11 avril). Blog HubSpot : marketing, vente, relation client et site web. Accédé en Juillet 2023
- [12] apprentissage automatique. (s. d.). La Vitrine linguistique de l'Office québécois de la langue française. Accédé en Juillet 2023
- [13] apprentissage profond. (s. d.). La Vitrine linguistique de l'Office québécois de la langue française. Accédé en Juillet 2023
- [14] réseau de neurones artificiels. (s. d.). La Vitrine linguistique de l'Office québécois de la langue française. Accédé en Juillet 2023
- [15] HAL open science, Gaël Latour, Laurianne Robinet , *La microscopie optique tridimensionnelle, de l'imagerie biomédicale à la caractérisation des objets du patrimoine*, 2020. Accédé en Juillet 2023

Bibliographie

- [16] *L'imagerie médicale par rayon x . (s. d.). HUG - Hôpitaux universitaires de Genève.* Accédé en Juillet 2023
- [17] *Médecine nucléaire / Elsan. (s. d.). Elsan.* Accédé en Juillet 2023
- [18] *Ilyes Talbi ; L'intelligence artificielle au service de la santé. 2020.* Accédé en Juillet 2023 Accédé en Juillet 2023
- [19] *Michal Drozdzal, Gabriel Chartrand, Eugene Vorontsov, Lisa Di Jorio, An Tang, Adriana Romero, Yoshua Bengio, Chris Pal, Samuel Kadoury ; Learning Normalized Inputs for Iterative Estimation in Medical Image Segmentation. 2017, PMC article, PMID : 29169029, DOI : 10.1016/j.media.2017.11.005* Accédé en Juillet 2023
- [20] *Gabriel Chartrand, PhD Phillip M. Cheng, MD, MS Eugene Vorontsov, BAscEng Sci Michal Drozdzal, PhD Simon Turcotte, MD, MSc Christopher J. Pal, PhD Samuel Kadoury, PhD An Tang, MD, MSc ; Deep Learning : A Primer for Radiologists1. 2017, Radio Graphics journal, pp. 37:2113–2131.* Accédé en Juillet 2023
- [21] *Roth et al, (2015). Anatomy-specific classification of medical images using deep convolutional nets. 2015. 10.1109/ISBI.2015.7163826.* Accédé en Juillet 2023
- [22] *Takiyama, H., Ozawa, T., Ishihara, S. et al. Automatic anatomical classification of esophagogastroduodenoscopy images using deep convolutional neural networks. Sci Rep 8, 7497 (2018).* Accédé en Juillet 2023
- [23] *S. Khan and S. -P. Yong, "A deep learning architecture for classifying medical images of anatomy object," 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Kuala Lumpur, Malaysia, 2017, pp. 1661-1668, doi : 10.1109/APSIPA.2017.8282299.* Accédé en Juillet 2023
- [24] *Yu, Y.; Lin, H.; Meng, J.; Wei, X.; Guo, H.; Zhao, Z. Deep Transfer Learning for Modality Classification of Medical Images. Information 2017, 8, 91.* Accédé en Juillet 2023
- [25] *Kim I, Rajaraman S, Antani S. Visual Interpretation of Convolutional Neural Network Predictions in Classifying Medical Image Modalities. Diagnostics (Basel). 2019 Apr 3;9(2) :38. doi : 10.3390/diagnostics9020038. PMID : 30987172; PMCID : PMC6627892.* Accédé en Juillet 2023
- [26] *Samuel Remedios, Dzung L. Pham , Jhon A. Butman et Snehashis Roy) (Classifying magnetic resonance image modalities with convolutional neural networks / Samuel Remedios, a, Dzung L. Phamb, John A. Butmanc, and Snehashis Royb / 17apr 2018.* Accédé en Juillet 2023
- [27] *F. Sica, G. Gobbi, P. Rizzoli and L. Bruzzone, "Φ-Net : Deep Residual Learning for InSAR Parameters Estimation," in IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 5, pp. 3917-3941, May 2021, doi : 10.1109/TGRS.2020.3020427.* Accédé en Juillet 2023

Bibliographie

- [28] *Classifying magnetic resonance image modalities with convolutional neural networks / Samuel Remedios,a,b, Dzung L. Phamb, John A. Butmanc, and Snehashis Royb / 17apr 2018.. Accédé en Juillet 2023*
- [29] *Chiang C-H, Weng C-L, Chiu H-W (2021) Automatic classification of medical image modality and anatomical location using convolutional neural network. PLoS ONE 16(6) : e0253205. Accédé en Juillet 2023*
- [30] *Msoud Nickparvar. (2021). Brain Tumor MRI Dataset [Data set]. Kaggle. Accédé en Juillet 2023*
- [31] *mahdi belhadi. (2023). mri filtreddataset [Data set]. Kaggle. Accédé en Juillet 2023*
- [32] *U.Baid, et al., “The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification”, arXiv :2107.02314, 2021 Accédé en Juillet 2023*
- [33] *Msoud Nickparvar. (2021). Brain Tumor MRI Dataset [Data set]. Kaggle. Accédé en Juillet 2023*
- [34] *Chest X-Ray Images [Data set]. kaggle. . Accédé en Juillet 2023*
- [35] *Figure-detection [Data set]. Kaggle . Accédé en Juillet 2023 Accédé en Juillet 2023*
- [36] *Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Dataset of breast ultrasound images. Data in Brief. 2020 Feb ;28:104863. DOI : 10.1016/j.dib.2019.104863.. Accédé en Juillet 2023*
- [37] *Ankit8467. (2023). Fetal Head UltraSound Dataset For Image Segment [Data set]. Kaggle.. Accédé en Juillet 2023*
- [38] *Mohamed Hany. (2020). Chest CT-Scan images Dataset [Data set]. Kaggle.. Accédé en Juillet 2023*
- [39] *MD NAZMUL ISLAM & Md Humaion Kabir Mehedi. (2021). CT KIDNEY DATASET : Normal-Cyst-Tumor and Stone [Data set]. Kaggle.. Accédé en Juillet 2023*
- [40] *U.Baid, et al., “The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification”, arXiv :2107.02314, 2021.. Accédé en Juillet 2023*
- [41] *chaouchi lynda, djani djouher, Application de l'apprentissage profond par transfert en imagerie médicale pour l'aide au diagnostic des lésions thyroïdiennes. Accédé en Juillet 2023*