

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique

UNIVERSITÉ ABDERRAHMANE MIRA - BÉJAÏA
FACULTÉ DE SCIENCES EXACTES
DÉPARTEMENT D'INFORMATIQUE



MÉMOIRE POUR L'OBTENTION DU DIPLÔME
DE MASTER EN INTELLIGENCE ARTIFICIELLE

Prédiction de la voie d'abord chirurgicale des tumeurs pétreuses : une approche basée sur le Deep Learning

Présenté par :
BENBERKANE Oussama
KHERBACHI Samy

Devant le jury composé de :

Mr AMROUN K	Pr	Encadreur
M BOUKERRAM S	MCA	Presidente
M TAHAKOURT Z	MCA	Examinatrice
Mr HIMEUR H	Pr	Examineur

2022/2023

Remerciements

Nous tenons tout d'abord à exprimer nos remerciements à notre directeur de recherche, Pr AMROUN, pour son soutien, ses conseils précieux qui ont été essentielles pour mener ce projet à bien.

Ensuite, nous tenons à remercier la Présidente de jury, M BOUKERRAM, et les deux examinateurs de notre soutenance, M TAHAKOURT et Mr HIMEUR, pour leur précieuse présence et leur évaluation minutieuse de notre travail.

Nous tenons également à remercier notre directeur de stage, Pr HIMEUR, pour sa contribution inestimable à notre projet. Sa présence et son expertise ont joué un rôle déterminant dans la réussite de ce projet. Nous lui sommes reconnaissants pour sa disponibilité et sa contribution essentielle tout au long de cette recherche. Egalement aux deux administrations celle de la section neurochirurgie du CHU Khelil Amrane, et celle de notre département.

Nous souhaitons également exprimer toute notre gratitude au Pr IGUEROUADA. Sa disponibilité et sa générosité dans le partage de ses connaissances ont grandement contribué à notre développement. Nous sommes profondément reconnaissants du temps qu'il a consacré à nous aider à progresser et à nous améliorer. Ainsi qu'à Mr Bouchebah pour nous avoir initié au concepts du Deep Learning et sa contribution tout au long de notre travail.

Nos remerciements vont également à nos collègues et amis qui nous ont apporté leur soutien et leurs encouragements pendant cette période. Leur soutien moral et leur camaraderie ont été d'une grande valeur pour nous.

Nous souhaitons également remercier notre famille pour leur soutien inconditionnel et leurs encouragements constants. Leur amour et leur encouragement ont été une source d'inspiration et de motivation tout au long de cette recherche.

BENBERKANE Oussama, KHERBACHI Samy

Liste des Abréviations

- IRM** : Imagerie par Résonance Magnétique
- CHU** : Centre Hospitalo Universitaire
- IA** : Intelligence Artificielle
- ML** : Machine Learning
- DL** : Deep Learning
- CNN** : Convolutional Neural Network
- RNN** : Recurrent Neural Network
- FC** : Fully Connected
- SVM** : Support Vector Machine
- KNN** : K-Nearest Neighbors
- VGG** : Visual Geometry Group
- DenseNet** : Densely Connected Convolutional Networks
- Val Loss** : Validation Loss
- Val Accuracy** : Validation Accuracy

Sommaire

Sommaire	I
Table des figures	1
Liste des tableaux	1
Introduction	1
I Neurochirurgie	3
I.1 Introduction	3
I.2 Définition	4
I.3 Rocher temporal	4
I.4 Tumeurs pétreuses	5
I.5 Les voies d'abord	6
I.6 IRM (neurochirurgie)	7
II Intelligence Artificielle	8
II.1 Introduction	8
II.2 Définition	8
II.3 Machine Learning	9
II.4 Les types de modèles ML	9
II.5 Réseaux de Neurons Artificiels	12
II.6 Deep Learning	13
II.7 Les types de Modèles de DL	13
II.8 Transfert learning	14
II.9 Ensembling	16
II.10 L'architecture et l'historique des réseaux de neurones à convolution (CNN)	17
III Etat De L'Art	20
III.1 Introduction	20

III.2 Travaux connexes	21
III.2.1 Les approches de classification d'IRM du cerveau	21
III.3 Tableau Comparatif	26
III.4 Conclusion	35
IV Proposition de notre approche	36
IV.1 Introduction	36
IV.2 Proposition de modeles pour le Transfer Learning	37
IV.2.1 Proposition 1 : VGG16	37
IV.2.2 Proposition 2 : Densenet121	38
IV.2.3 Proposition 3; InceptionV3	39
IV.3 Nos Modeles et Ensembling	41
V Test et résultats de notre proposition	43
V.1 Introduction	43
V.2 Préliminaires	43
V.2.1 Entraînement et validation	43
V.2.2 Techniques et concepts avancés du DL	47
V.3 Dataset	48
V.3.1 Données utilisées	48
V.3.2 Tri et labellisation de notre dataset	48
V.3.3 Prétraitement et Augmentation	49
V.4 Entraînement et test de nos modèles	51
V.4.1 Entraînement du modèle VGG16	52
V.4.2 Entraînement du modèle InceptionV3	53
V.4.3 Entraînement du modèle Densenet121	54
V.5 Résultats	56
V.5.1 Résultats du modèle VGG16	56
V.5.2 Résultats du modèle InceptionV3	58
V.5.3 Résultats du modèle Densenet121	61
V.6 Discussion des résultats	64
V.7 Application : Implémentation et Intégration des Modèles	66
Conclusion générale	70

Table des figures

I.1	Représentation de l'os temporal	4
I.2	Image IRM d'une tumeur schwannome	5
I.3	Image IRM d'une tumeur méninge	6
I.4	Exemple d'une IRM du cerveau	7
II.1	Représentation d'un apprentissage supervisé et non supervisé	10
II.2	Représentation d'un apprentissage semi-supervisé	11
II.3	Représentation d'un apprentissage par renforcement	12
II.4	Représentation d'un réseau de neurones	13
II.5	Représentation d'un Transfert learning	15
II.6	Représentation d'un Ensembling de trois classificateurs	16
II.7	Représentation de l'architecture d'un CNN	17
IV.1	Représentation de l'architecture du modèle VGG16	38
IV.2	Représentation de l'architecture du modèle Densenet121	39
IV.3	Représentation de l'architecture du modèle InceptionV3	40
IV.4	Représentation de l'architecture de notre modèle densenet121	41
IV.5	Représentation de l'architecture de notre modèle VGG16	41
IV.6	Représentation de l'architecture de notre modèle InceptionV3	42
IV.7	Représentation de l'architecture de notre modèle Ensembling de nos trois modèles	42
V.1	Représentation d'un underfitting et d'un Overfitting	46
V.2	Interface de classification des images IRM en voie d'abord sur notre application	49
V.3	Représentation du redimensionnement d'une image en 224 x 224	50
V.4	Représentation de la division de notre dataset pour l'entraînement et le test	50
V.5	Représentation d'un exemple d'augmentation de données sur une image	51
V.6	Visualisation des métriques d'évaluation de notre modèle basé sur VGG16 sur le dataset public	56
V.7	Visualisation des métriques d'évaluation de notre modèle basé sur VGG16 sur le dataset privé	57
V.8	Tableau représentant les résultats du modèle VGG16 sur les deux datasets	58

V.9	Visualisation des métriques d'évaluation de notre modèle basé sur InceptionV3 sur le dataset public	59
V.10	Visualisation des métriques d'évaluation de notre modèle basé sur InceptionV3 sur le dataset privé	60
V.11	Tableau représentant les résultats du modèle InceptionV3 sur les deux datasets	61
V.12	Visualisation des métriques d'évaluation de notre modèle basé sur Densenet121 sur le dataset public	62
V.13	Visualisation des métriques d'évaluation de notre modèle basé sur Densenet121 sur le dataset privé	63
V.14	Tableau représentant les résultats du modèle Densenet121 sur les deux datasets	64
V.15	Interface d'accueil de l'application	66
V.16	Interface de classification des images IRM en voie d'abord sur notre application	67
V.17	Interface d'administrateur pour la gestion des images IRM sur notre application	67
V.18	Interface de test sur notre application	68
V.19	Exemple d'une réponse correcte sur l'interface de test sur notre application .	68
V.20	Exemple d'une réponse incorrecte sur l'interface de test sur notre application	69
V.21	Interface pour charger sa propre image IRM sur notre application	69

Liste des tableaux

III.1 Tableau de comparaison des articles de l'état de l'art	34
V.1 Les paramètres utilisés pour notre approche	52

Résumé

Ce mémoire vise à développer un modèle d'intelligence artificielle basé sur des réseaux de neurones convolutifs (CNN) pour prédire l'approche chirurgicale optimale pour les tumeurs cérébrales situées à la base du crâne, en se basant sur des images d'IRM. Le Deep Learning, en particulier avec le Transfer Learning, a montré son efficacité dans la classification des tumeurs cérébrales à partir d'images d'IRM. Trois modèles populaires en Transfer Learning ont été utilisés et des techniques d'Ensembling ont été envisagées pour améliorer la classification. Les résultats préliminaires sont encourageants, mais des travaux futurs sont nécessaires pour augmenter la taille de l'ensemble de données et explorer d'autres architectures et techniques d'optimisation. Notre approche est ensuite implémentée dans une application web permettant aux résidents, médecins inexpérimentés et étudiants d'acquérir plus d'expérience en termes de choix de l'approche chirurgicale optimale. Ce mémoire ouvre la voie à des avancées cliniques significatives dans la classification des tumeurs cérébrales, en particulier pour les professionnels médicaux moins expérimentés.

Abstract

This thesis aims to develop an artificial intelligence model based on convolutional neural networks (CNN) to predict the optimal surgical approach for skull base brain tumors, based on MRI images. Deep learning, specifically with the addition of Transfer Learning, has demonstrated its effectiveness in classifying brain tumors from MRI images. Three popular Transfer Learning models were used, and Ensembling techniques were considered to improve the classification. The preliminary results are promising, but further work is needed to increase the dataset size and explore other architectures and optimization techniques. Our approach is then implemented in a web application that enables residents, inexperienced doctors, and students to gain more experience in terms of selecting the optimal surgical approach. This thesis paves the way for significant clinical advancements in brain tumor classification, particularly for less experienced medical professionals.

Introduction

La neurochirurgie est une spécialité médicale complexe qui se concentre sur le traitement des affections du système nerveux central, en particulier du cerveau et de la moelle épinière [37]. Les tumeurs cérébrales situées à la base du crâne, telles que les méningiomes et les schwannomes, posent des défis importants en raison de leur emplacement délicat. Pour aborder ces tumeurs chirurgicalement, une planification minutieuse et une connaissance approfondie de l'anatomie crânienne sont nécessaires [11]. Cependant, les médecins ou les résidents peuvent manquer d'expérience et de compétences nécessaires pour déterminer l'approche chirurgicale la plus appropriée pour chaque patient. C'est là que notre projet entre en jeu.

Notre travail vise à développer un modèle d'intelligence artificielle basé sur un réseau de neurones convolutifs (CNN) qui permettra de prédire l'approche chirurgicale optimale en se basant sur des données d'imagerie par résonance magnétique (IRM). Ce modèle sera accessible via une application web comprenant des tests d'évaluation.

Afin de mener à bien ce projet, nous avons effectué un stage au CHU Khelil Amrane de Bejaia en collaboration avec le Pr Himeur, un neurochirurgien spécialisé dans les interventions chirurgicales pour les tumeurs cérébrales situées à la base du crâne, ainsi que deux de ses étudiantes qu'il encadre pour une thèse de la même thématique. Il convient de mentionner que les deux principales voies d'abord chirurgicales utilisées pour les tumeurs cérébrales situées à la base du crâne sont la voie d'abord combinée et la voie d'abord trans-pétreuse. La voie d'abord combinée consiste en une approche qui combine les techniques d'abord trans-crânienne et trans-pétreuse. D'autre part, la voie d'abord trans-pétreuse implique le retrait d'une partie de l'os pétreux pour accéder aux tumeurs situées à la base du crâne, évitant ainsi une ouverture trans-crânienne plus étendue. Ces voies d'abord chirurgicales seront prises en considération dans notre modèle CNN afin d'aider à déterminer la meilleure approche chirurgicale pour chaque cas spécifique.

Le travail est structuré en plusieurs chapitres qui couvrent les domaines clés de la recherche. Le premier chapitre met l'accent sur la neurochirurgie et les tumeurs cérébrales situées à la base du crâne. Le deuxième chapitre explore l'IA en se concentrant sur les CNN. Le troisième chapitre parcourt l'état de l'art de l'IA appliquée à la neurochirurgie. Le quatrième chapitre aborde notre proposition et notre méthodologie utilisée pour développer un modèle CNN basé sur des données d'IRM. Le cinquième présente le test de notre modèle et les résultats obtenus avec les discussions associées ainsi que sur l'implémentation de notre modèle dans une application WEB. Le travail s'achève par une conclusion générale et souligne les perspectives futures.

Neurochirurgie

I.1 Introduction

La neurochirurgie est une branche spécialisée de la chirurgie qui se consacre au diagnostic, au traitement et à la gestion des affections touchant le système nerveux central et périphérique, comprenant le cerveau, la moelle épinière et les nerfs. C'est un domaine complexe et techniquement exigeant de la médecine, requérant une expertise approfondie et des compétences chirurgicales précises. [37]

Au fil des décennies, la neurochirurgie a connu une évolution significative grâce aux avancées technologiques et scientifiques [31]. Autrefois limitée aux interventions chirurgicales majeures, cette spécialité s'est élargie pour inclure des procédures moins invasives et des approches plus ciblées, ouvrant de nouvelles perspectives de traitement [45].

Les neurochirurgiens travaillent en étroite collaboration avec d'autres professionnels de la santé, tels que les neurologues, les radiologues, les oncologues et les spécialistes des soins intensifs, afin de fournir des soins complets et multidisciplinaires aux patients atteints de troubles neurologiques. Ils font face à une grande variété de conditions, allant des tumeurs cérébrales et des maladies vasculaires du cerveau aux traumatismes crâniens, aux malformations congénitales et aux troubles du mouvement. [37]

En raison de la complexité des affections neurologiques et des risques associés aux interventions chirurgicales, la neurochirurgie exige un haut niveau de compétence, d'expérience et de dévouement [35]. Ce chapitre explore les différentes facettes de la neurochirurgie, y compris les techniques chirurgicales de voies d'abord, et l'anatomie étudiée dans notre cas.

I.2 Définition

La neurochirurgie est une spécialité médicale qui se consacre au diagnostic, au traitement et à la gestion chirurgicale des affections affectant le système nerveux central (cerveau et moelle épinière) ainsi que le système nerveux périphérique. Elle vise à traiter des pathologies telles que les malformations vasculaires, les traumatismes crâniens, les hernies discales, les maladies dégénératives de la colonne vertébrale, les troubles du mouvement, les épilepsies, les infections du système nerveux, et, dans notre cas d'étude, les tumeurs cérébrales. [31]

I.3 Rocher temporal

Le rocher crânien, également connu sous le nom de rocher temporal, est une partie de l'os temporal située à la base du crâne, près de l'oreille. Il présente une forme en pyramide et abrite plusieurs structures importantes, notamment l'oreille interne, le nerf facial (VIIe paire de nerfs crâniens), le nerf vestibulocochléaire (VIIIe paire de nerfs crâniens), ainsi que des canaux et des cavités liés à ces structures. Le rocher crânien joue un rôle essentiel dans la fonction auditive et l'équilibre. [5]



FIGURE I.1 – Représentation de l'os temporal
[13]

I.4 Tumeurs pétreuses

Les tumeurs pétreuses sont des tumeurs qui se développent à proximité de l'os pétreux, une partie de l'os temporal située à la base du crâne. Ces tumeurs peuvent inclure des méningiomes, des schwannomes, des neurinomes, des chondromes et d'autres types de tumeurs qui se forment près de cette région [21]. En raison de leur localisation près de structures importantes telles que les nerfs crâniens et les vaisseaux sanguins, les tumeurs pétreuses peuvent présenter des défis particuliers lors de leur traitement chirurgical.

- **Schwannome** Un schwannome est une tumeur généralement bénigne qui se développe à partir des cellules de Schwann, qui sont responsables de la formation de la gaine de myéline entourant les nerfs périphériques. Ces tumeurs se forment principalement le long des nerfs crâniens ou des nerfs périphériques du corps. Les schwannomes peuvent causer des symptômes tels que des douleurs, des engourdissements, une faiblesse musculaire et des problèmes de coordination. [10]

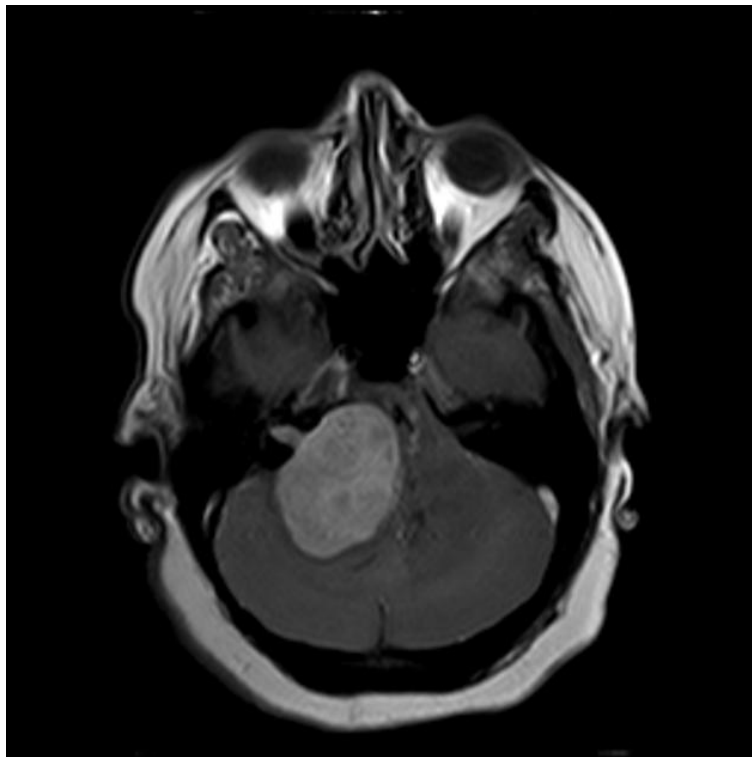


FIGURE I.2 – Image IRM d'une tumeur schwannome
[13]

- **Méningiome** Un méningiome est une tumeur qui se développe à partir des cellules des méninges, les membranes qui recouvrent le cerveau et la moelle épinière. La majorité des méningiomes sont bénins, bien qu'ils puissent parfois être malins. Ces tumeurs peuvent causer des symptômes tels que des maux de tête, des troubles visuels, des convulsions et des problèmes de coordination, en fonction de leur emplacement et de leur taille. [11]



FIGURE I.3 – Image IRM d'une tumeur méninge
[13]

I.5 Les voies d'abord

Les voies d'abord, en neurochirurgie, font référence aux différentes approches chirurgicales utilisées pour accéder à une lésion ou à une tumeur cérébrale. Elles déterminent le trajet emprunté par le chirurgien pour atteindre la zone cible [59]. Les voies d'abord peuvent varier en fonction de la localisation et de la taille de la lésion, ainsi que des structures anatomiques environnantes. Les principales voies d'abord comprennent l'abord trans-crânien, qui implique l'ouverture du crâne, et l'abord trans-pétreux, qui implique le retrait d'une partie de l'os pétreux près de l'oreille, et les voies combinées, qui impliquent l'utilisation de plusieurs trajets chirurgicaux pour atteindre la zone cible. [12]

- **Les voies trans-pétreuses** en neurochirurgie, sont des approches chirurgicales utilisées pour accéder à des lésions ou des tumeurs situées à proximité de l'os pétreux, une partie de l'os temporal à la base du crâne. Ces voies d'abord nécessitent le retrait d'une partie de l'os pétreux pour exposer la région cible [42]. Elles offrent un accès direct à des structures telles que le nerf facial (VIIe paire de nerfs crâniens) et le nerf vestibulocochléaire (VIIIe paire de nerfs crâniens). Les voies trans-pétreuses peuvent être utilisées dans le traitement de certaines tumeurs cérébrales et de maladies de l'oreille interne. [12]
- **Les voies combinées** en neurochirurgie, est une approche chirurgicale qui combine les techniques d'abord trans-crâniennes et trans-pétreuses pour accéder à une lésion ou à une tumeur située à la base du crâne. Cette approche permet de contourner certaines difficultés associées à chaque technique d'abord individuelle en offrant une exposition plus large et un accès plus direct à la zone cible [42]. Elle peut être utilisée lorsque la lésion se trouve dans une position complexe ou près de structures anatomiques sensibles. [12]

I.6 IRM (neurochirurgie)

Dans le domaine de la neurochirurgie, l'IRM (Imagerie par Résonance Magnétique) est une technique d'imagerie médicale avancée utilisée pour visualiser en détail les structures du cerveau et de la moelle épinière. L'IRM utilise des champs magnétiques puissants et des ondes radio pour générer des images en trois dimensions des tissus mous et des structures anatomiques du système nerveux central [24]. Cela permet aux neurochirurgiens d'obtenir des informations précieuses sur la localisation, la taille, la forme et les relations spatiales des lésions, telles que les tumeurs, les malformations vasculaires ou les anomalies structurelles. L'IRM est souvent utilisée dans la planification chirurgicale, l'évaluation préopératoire et le suivi postopératoire des patients neurochirurgicaux. [31]

Original Image

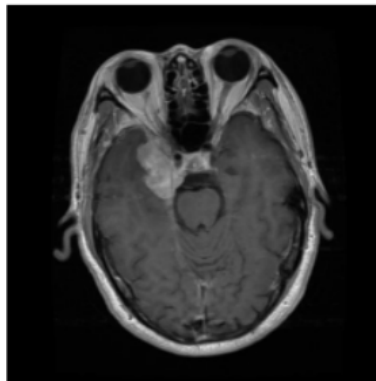


FIGURE I.4 – Exemple d'une IRM du cerveau [48]

Intelligence Artificielle

II.1 Introduction

L'intelligence artificielle (IA) est un domaine en constante évolution qui joue un rôle de plus en plus crucial dans de nombreux domaines, notamment celui de la médecine. L'IA se réfère à la capacité d'un système informatique ou d'une machine à imiter et à reproduire les fonctions cognitives humaines telles que l'apprentissage, la résolution de problèmes et la prise de décisions. [36]

Au cours des dernières décennies, l'IA a connu des avancées remarquables grâce aux progrès technologiques et aux algorithmes de plus en plus sophistiqués. Ses applications dans le domaine médical sont vastes et variées. Elle a permis le développement d'assistants virtuels intelligents, l'amélioration de la détection de maladies, la prédiction de résultats médicaux, la personnalisation des traitements, et bien d'autres applications encore.

L'IA continue de façonner notre monde et de transformer la manière dont nous interagissons avec la technologie et la société. Comprendre les concepts et les applications de l'IA dans le domaine médical est essentiel pour saisir les opportunités qu'elle offre et pour relever les défis qu'elle présente [36]. Dans ce chapitre, nous plongerons dans l'univers de l'IA et ses principes fondamentaux.

II.2 Definition

L'IA fait référence à la capacité des machines ou des systèmes informatiques à effectuer des tâches qui nécessitent normalement l'intelligence humaine. L'IA vise à créer des algorithmes et des modèles qui permettent aux machines de percevoir, de comprendre, de raisonner, d'apprendre et de prendre des décisions de manière autonome. [2]

II.3 Machine Learning

Le ML, apprentissage automatique, est une branche de l'IA qui se concentre sur le développement d'algorithmes et de modèles informatiques capables d'apprendre à partir de données et d'améliorer leurs performances sans être explicitement programmés. Plutôt que d'utiliser des règles strictes et prédéfinies, le ML permet aux machines d'analyser et d'interpréter les données pour détecter des schémas, des relations et des tendances, afin de prendre des décisions ou de prédire des résultats futurs. Les techniques couramment utilisées en ML incluent les réseaux de neurones, les arbres de décision, les méthodes de clustering et les algorithmes de régression. [3]

II.4 Les types de modèles ML

- **Supervised Learning**

L'apprentissage supervisé est une approche du ML dans laquelle un modèle est entraîné à partir d'exemples étiquetés. L'objectif de l'apprentissage supervisé est de prédire des valeurs ou des étiquettes correctes à partir de nouvelles données non étiquetées, en se basant sur les connaissances acquises à partir des exemples d'entraînement.

Dans le cadre de l'apprentissage supervisé, le modèle est présenté à un ensemble de données d'entraînement qui comprend à la fois les échantillons de données et les étiquettes correspondantes. Le modèle utilise ces exemples étiquetés pour apprendre les relations et les schémas entre leurs caractéristiques et les étiquettes, afin de généraliser et de faire des prédictions sur de nouvelles données non étiquetées.

L'apprentissage supervisé peut être utilisé dans une variété de tâches, telles que la classification, la régression et la prédiction. Dans la classification, le modèle apprend à assigner des classes ou des catégories à de nouvelles instances en fonction des exemples d'entraînement. [53]

- **Unsupervised Learning**

L'apprentissage non supervisé est une approche ML dans laquelle un modèle est entraîné sur un ensemble de données d'entrées non étiquetées. L'objectif est de découvrir des structures ou des modèles intrinsèques dans les données sans connaître les sorties attendues.

Les algorithmes d'apprentissage non supervisé sont couramment utilisés pour effectuer des tâches telles que la regroupement (clustering) des données en groupes similaires ou la réduction de dimension pour représenter les données de manière plus concise. Des exemples d'algorithmes d'apprentissage non supervisé comprennent les algorithmes de clustering k-means et l'analyse en composantes principales (PCA). [53]

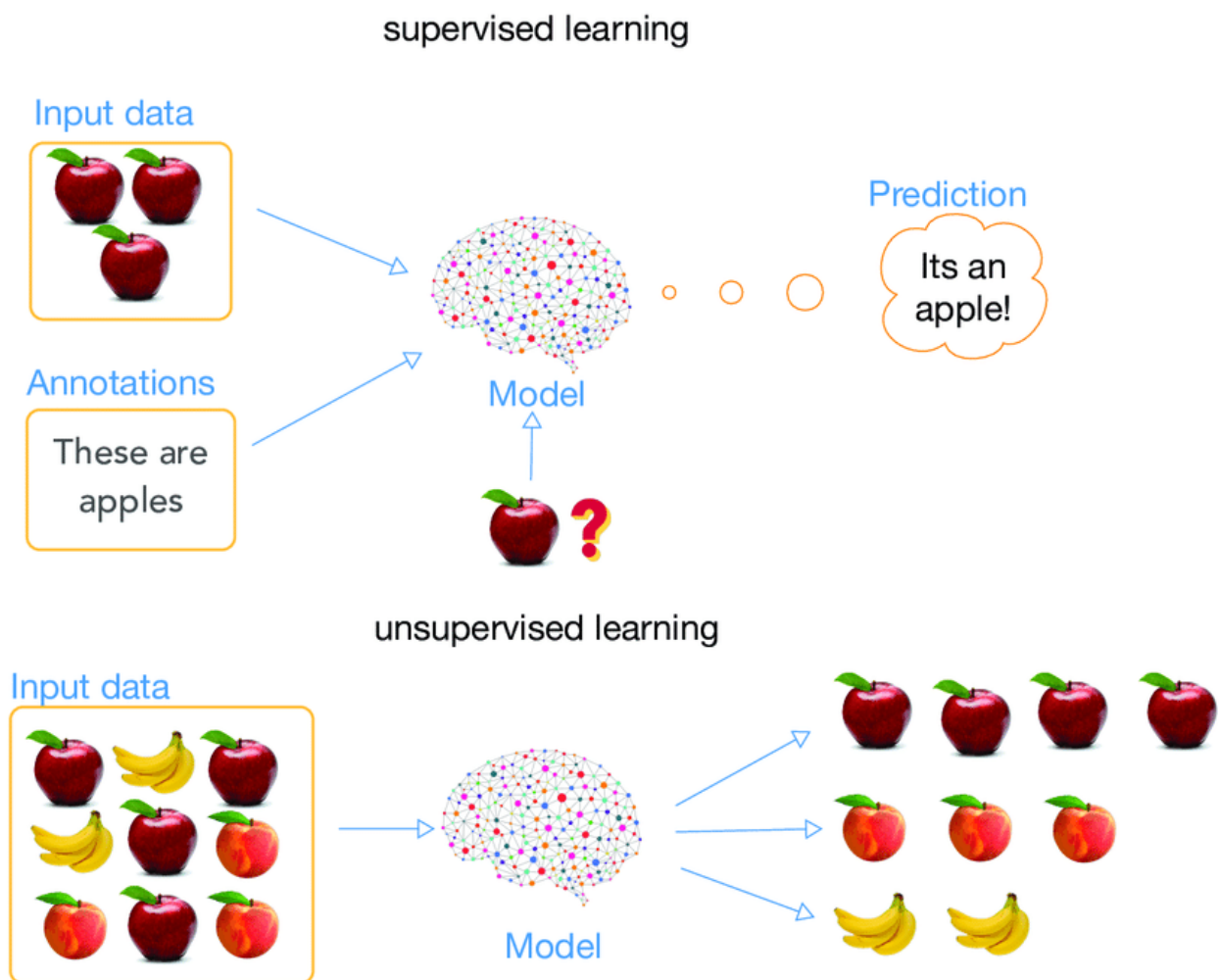


FIGURE II.1 – Représentation d'un apprentissage supervisé et non supervisé [60]

- **semi-supervised Learning**

L'apprentissage semi-supervisé est une approche ML qui utilise à la fois des données étiquetées et non étiquetées pour entraîner un modèle. Cela permet d'exploiter les informations des données non étiquetées pour améliorer les performances du modèle tout en réduisant la nécessité d'étiqueter manuellement de grandes quantités de données. Les algorithmes d'apprentissage semi-supervisé combinent des techniques d'apprentissage supervisé et non supervisé pour tirer parti de ces deux types de données. [53]

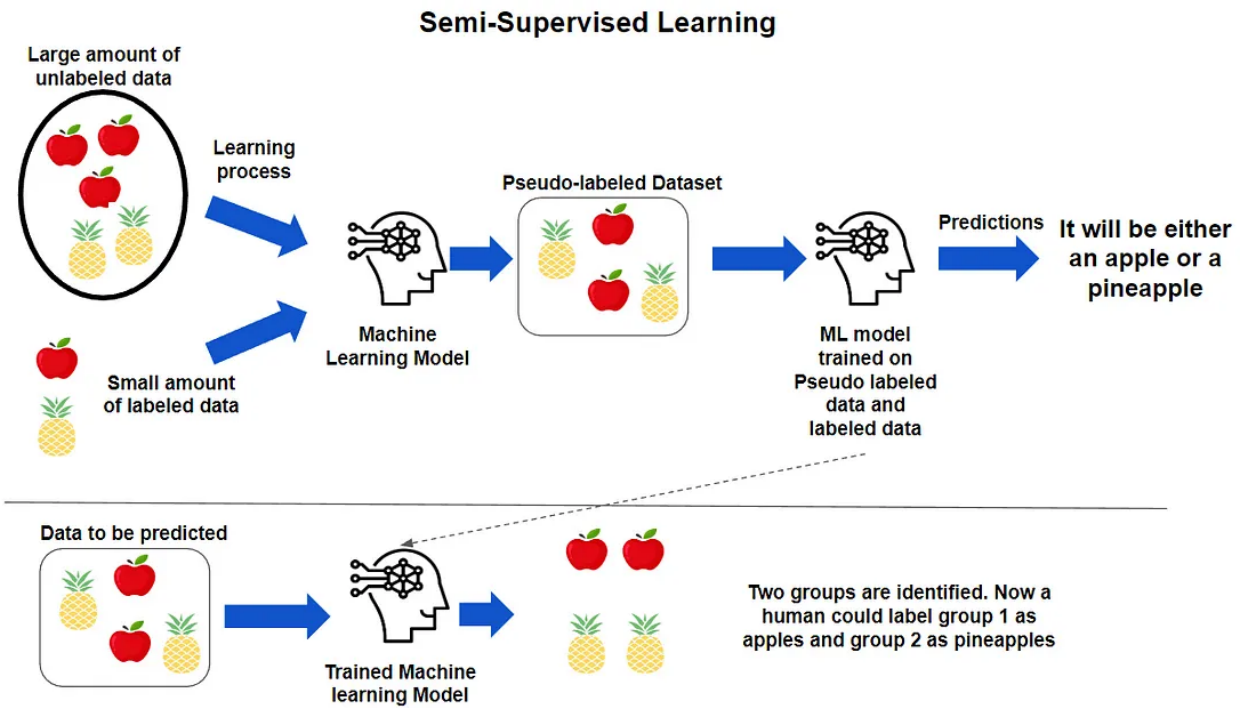


FIGURE II.2 – Représentation d'un apprentissage semi-supervisé [8]

- **Reinforcement Learning**

L'apprentissage par renforcement est une méthode ML où un agent apprend à prendre des décisions séquentielles en interagissant avec un environnement. L'agent reçoit des récompenses ou des pénalités en fonction de ses actions, ce qui lui permet d'ajuster son comportement pour maximiser les récompenses sur le long terme. Les algorithmes d'apprentissage par renforcement sont utilisés dans des domaines tels que les jeux, la robotique et la planification automatique. [53]

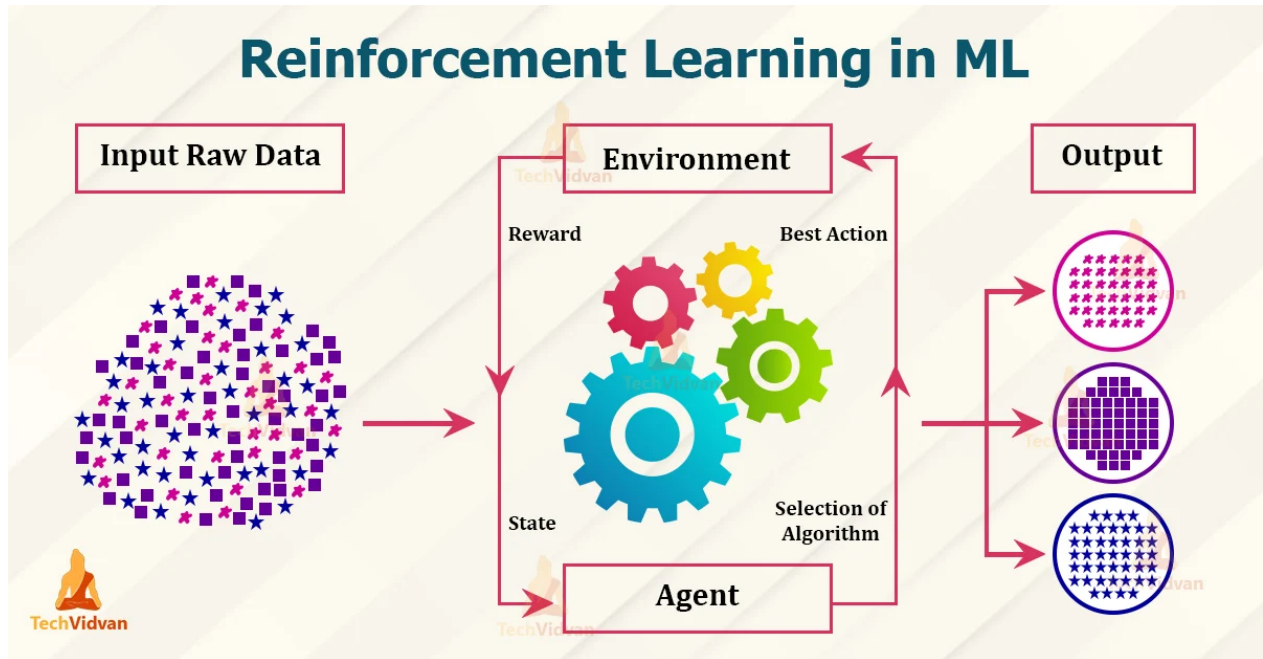


FIGURE II.3 – Représentation d'un apprentissage par renforcement [16]

II.5 Réseaux de Neurones Artificiels

Les réseaux de neurones artificiels, également appelés réseaux neuronaux ou simplement neurones, sont des modèles mathématiques inspirés du fonctionnement des réseaux de neurones biologiques présents dans le cerveau humain. Ces modèles sont utilisés dans le domaine de l'IA pour résoudre des problèmes complexes d'apprentissage et de reconnaissance de motifs. Les réseaux de neurones sont composés de couches de nœuds interconnectés, appelés neurones, qui traitent et transmettent des informations à travers des connexions pondérées. L'apprentissage dans les réseaux de neurones se produit par l'ajustement des poids des connexions en fonction des exemples d'entrée et de sortie fournis au réseau. [4]

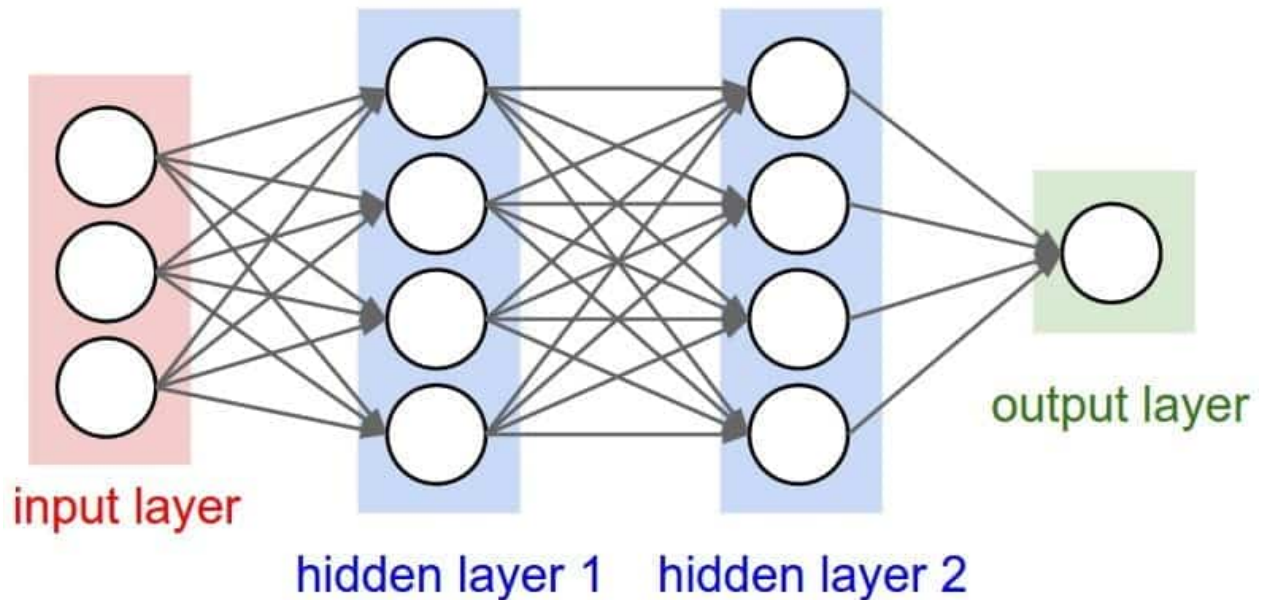


FIGURE II.4 – Représentation d’un réseau de neurones [7]

II.6 Deep Learning

L’apprentissage profond est une sous-branche du ML qui se concentre sur l’utilisation de réseaux de neurones artificiels profonds pour apprendre à partir de données. Le DL implique l’utilisation de réseaux de neurones à plusieurs couches, également appelés réseaux de neurones profonds, qui sont capables d’apprendre des représentations hiérarchiques complexes des données. Ces réseaux de neurones profonds sont conçus pour extraire des caractéristiques de plus en plus abstraites à mesure que l’information se propage à travers les différentes couches du réseau. Le DL a obtenu des résultats remarquables dans des domaines tels que la vision par ordinateur, la reconnaissance vocale, la traduction automatique et bien d’autres. [3]

II.7 Les types de Modèles de DL

Dans le domaine du DL, plusieurs types de modèles sont utilisés pour traiter différents types de données :

- **Réseaux de neurones à convolution (CNN)**

Les réseaux de neurones à convolution sont des modèles spécialement conçus pour traiter efficacement des données structurées en grille, telles que des images. Ils utilisent des couches de convolution pour extraire automatiquement des caractéristiques pertinentes des images, suivies de couches de regroupement (pooling)

et de couches entièrement connectées pour effectuer la classification ou la régression. Les CNN sont largement utilisés dans la vision par ordinateur [41].

- **Réseaux de neurones récurrents (RNN)**

Les réseaux de neurones récurrents sont conçus pour modéliser des données séquentielles, telles que des séquences de texte ou de données temporelles. Ils utilisent des boucles récurrentes pour propager les informations d'une étape à l'autre, leur permettant de capturer les dépendances à long terme dans la séquence. Les RNN sont couramment utilisés pour des tâches telles que la traduction automatique, la génération de texte et l'analyse de sentiments [41].

- **Réseaux de neurones générateurs adverses (GAN)**

Les réseaux de neurones générateurs adverses sont utilisés pour générer de nouvelles données réalistes en simulant un processus de génération de données. Ils se composent de deux réseaux : un générateur qui produit des échantillons synthétiques et un discriminateur qui essaie de distinguer les échantillons réels des échantillons synthétiques. Les GAN sont utilisés dans des domaines tels que la génération d'images, la synthèse de voix et la création de vidéos [41].

- **Réseaux de neurones autoencodeurs (AE)**

Les réseaux de neurones autoencodeurs sont des modèles utilisés pour apprendre des représentations comprimées des données en utilisant un processus d'encodage et de décodage. Ils sont composés d'un encodeur qui réduit la dimensionnalité des données en une représentation latente et d'un décodeur qui tente de reconstruire les données d'origine à partir de cette représentation. Les AE sont utilisés pour des tâches telles que la réduction de dimension, la détection d'anomalies et la génération de données [41].

- **Réseaux de neurones transformateurs**

Les réseaux de neurones transformateurs sont des modèles qui utilisent une architecture basée sur l'attention pour traiter des données séquentielles, telles que des phrases ou des séquences temporelles. Ils utilisent des mécanismes d'attention pour permettre aux différents éléments de la séquence d'interagir de manière dynamique lors du traitement de l'information. Les transformateurs sont couramment utilisés pour la traduction automatique et le traitement du langage naturel [41].

II.8 Transfert learning

Le transfert d'apprentissage est une technique ML qui consiste à appliquer les connaissances acquises lors de l'apprentissage d'une tâche à une nouvelle tâche similaire ou différente. Plutôt que de construire un modèle à partir de zéro, le transfert d'apprentissage

permet d'utiliser un modèle pré-entraîné sur une tâche antérieure, appelée tâche source, et de l'adapter à la nouvelle tâche, appelée tâche cible. [50]

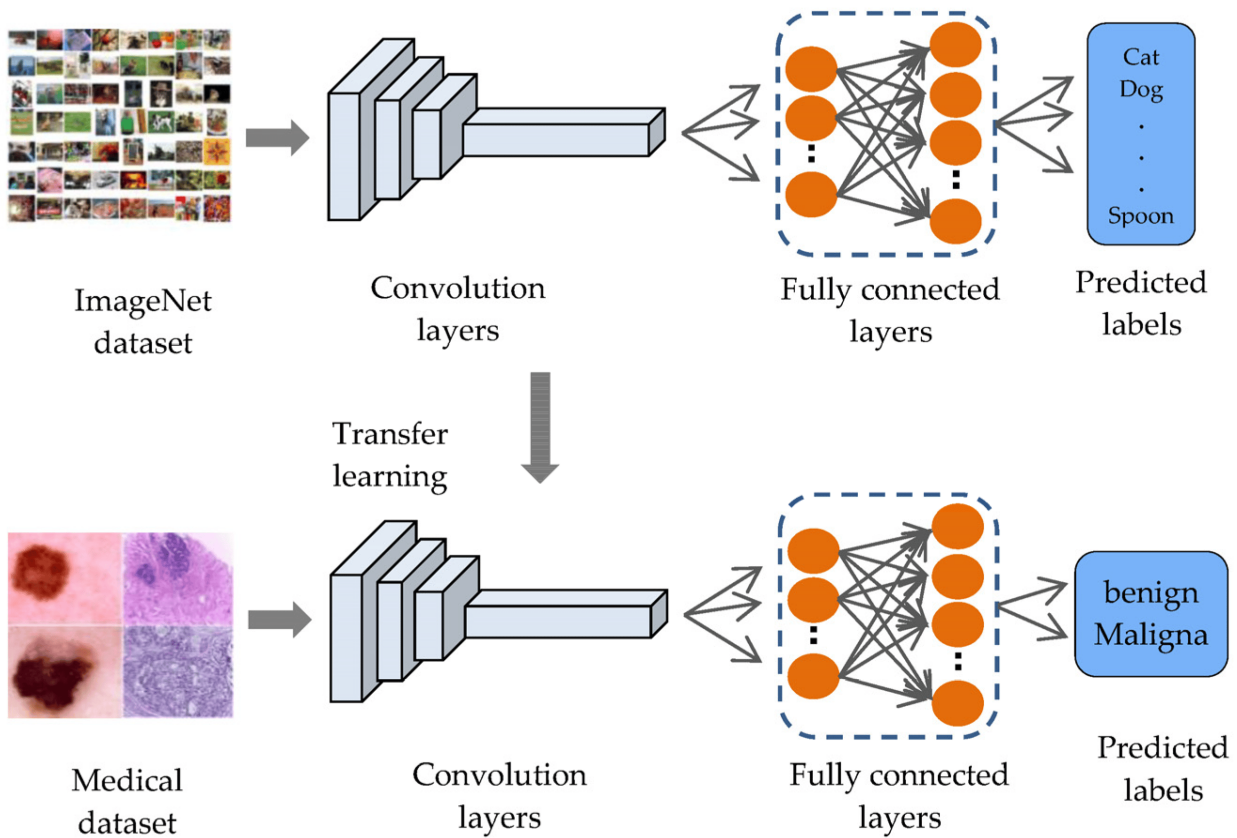


FIGURE II.5 – Représentation d'un Transfert learning [46]

- **La différence d'un Transfert learning avec poids gelés et non gelés**

Lorsque les poids sont gelés lors du transfert d'apprentissage, cela signifie que les poids des couches pré-entraînées ne sont pas mis à jour lors de l'apprentissage de la nouvelle tâche. Seules les couches supplémentaires ajoutées spécifiquement pour la nouvelle tâche sont entraînées. Cela est utile lorsque les caractéristiques apprises par les couches pré-entraînées sont suffisamment générales et adaptées à la nouvelle tâche. En revanche, lorsque les poids ne sont pas gelés, les poids des couches pré-entraînées peuvent être ajustés ou fine-tunés lors de l'apprentissage de la nouvelle tâche. Cela permet d'adapter davantage le modèle aux spécificités de la nouvelle tâche. Cependant, cela nécessite généralement plus de données d'apprentissage et peut présenter un risque de surapprentissage si les données d'apprentissage sont limitées [50].

- **La différence du Transfert learning avec et sans couches supérieures**

L'utilisation des couches supérieures se réfère à l'utilisation des couches les plus proches de la sortie du modèle pré-entraîné dans le transfert d'apprentissage. Ces couches capturent des caractéristiques plus abstraites et spécifiques à la tâche source. Si la nouvelle tâche est similaire à la tâche source, ces couches supérieures peuvent être utilisées directement sans modification [50].

II.9 Ensembling

L'Ensembling, également appelé apprentissage en comité ou méthode des ensembles, est une technique de ML qui vise à combiner les prédictions de plusieurs modèles individuels pour améliorer la performance globale et la stabilité des prédictions. En construisant un comité de modèles diversifiés et en agrégeant leurs prédictions, l'Ensembling permet de réduire le biais et la variance, d'augmenter la robustesse et d'obtenir de meilleures performances prédictives. Les techniques d'Ensembling les plus couramment utilisées sont le vote majoritaire, le vote pondéré, le bagging et le boosting. [28]

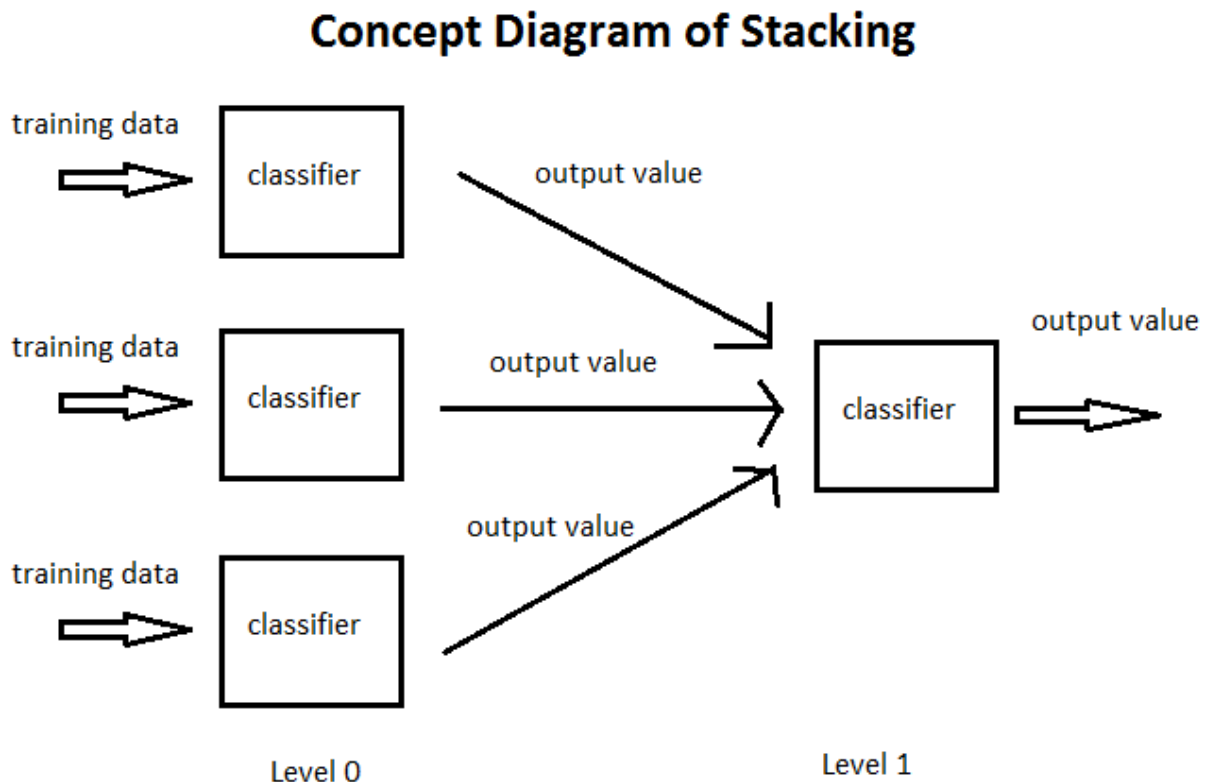


FIGURE II.6 – Représentation d'un Ensembling de trois classificateurs [1]

II.10 L'architecture et l'historique des réseaux de neurones à convolution (CNN)

Les réseaux de neurones à convolution (CNN) sont une classe de modèles d'apprentissage profond spécialement conçus pour traiter de manière efficace des données structurées en grille, telles que des images. Ils se basent sur l'utilisation de couches de convolution, qui permettent d'extraire automatiquement des caractéristiques significatives à partir des données d'entrée. Les CNN sont caractérisés par leur capacité à capturer des motifs locaux et des structures hiérarchiques dans les images, ce qui les rend particulièrement adaptés pour des tâches de vision par ordinateur telles que la classification d'images, la détection d'objets et la segmentation sémantique. [41]

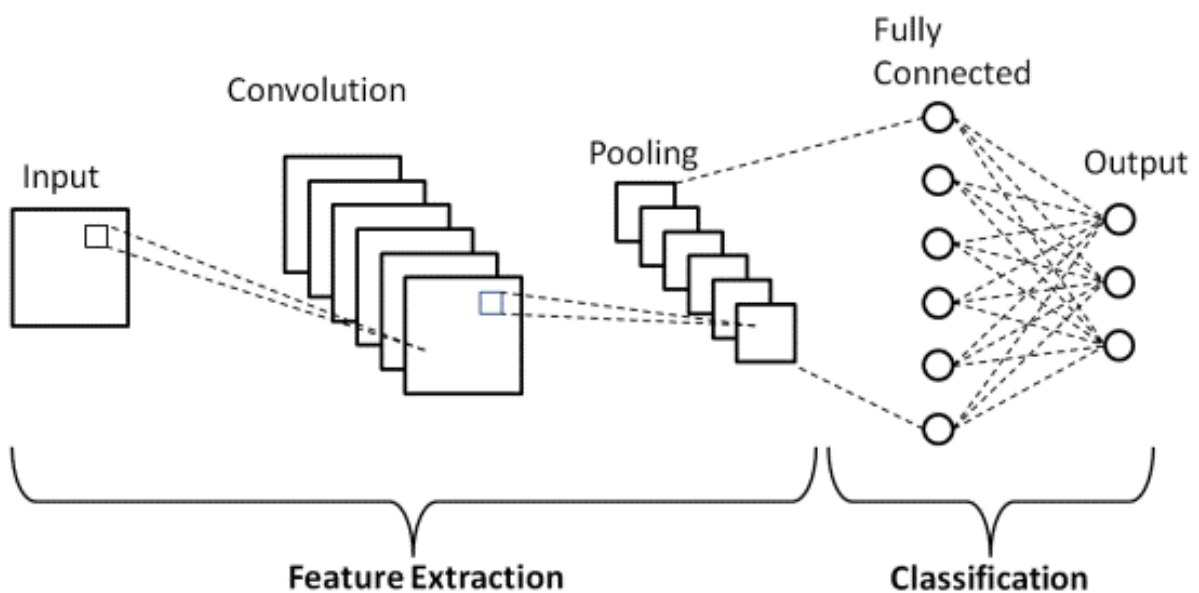


FIGURE II.7 – Représentation de l'architecture d'un CNN [20]

Convolution

La convolution est une opération fondamentale dans les CNN. Elle consiste à appliquer un filtre (auss appelé noyau) à une image ou à une autre représentation de données. Cette opération permet de détecter des motifs ou des caractéristiques spécifiques dans les données. L'utilisation de multiples filtres de convolution permet d'extraire des caractéristiques variées à différentes échelles. [41]

Pooling

Le pooling est une opération utilisée dans les CNN pour réduire la dimension des caractéristiques extraites. Les opérations de pooling, telles que le max pooling ou le average

pooling, réduisent la résolution spatiale des caractéristiques en conservant uniquement les valeurs maximales ou moyennes dans des régions spécifiques. Cela permet de réduire la complexité du modèle et de rendre l'apprentissage plus robuste face aux variations de position ou de taille des caractéristiques. [41]

Couches fully connected

Une couche fully connected, également appelée couche dense, est une couche où tous les neurones sont connectés à tous les neurones de la couche précédente. Dans ces couches, chaque neurone calcule une combinaison linéaire des entrées et les passe ensuite à une fonction d'activation non linéaire. Les couches fully connected sont souvent utilisées à la fin d'un réseau de neurones pour effectuer la classification ou la régression. [41]

Exemples de CNN

Quelques-uns des principaux CNN existants et un bref aperçu de leur historique :

- **LeNet-5** LeNet-5, développé par Yann LeCun et ses collègues dans les années 90, est considéré comme l'un des premiers CNN à succès. Il a été utilisé pour la reconnaissance de caractères manuscrits et a posé les bases du développement ultérieur des CNN. [40]
- **AlexNet** AlexNet, proposé par Alex Krizhevsky, Ilya Sutskever et Geoffrey Hinton en 2012, a marqué une avancée significative dans la performance des CNN. Il a remporté le concours ImageNet Large Scale Visual Recognition Challenge (ILSVRC) en 2012 et a montré la capacité des CNN à surpasser les méthodes traditionnelles de traitement d'images.[39]
- **VGGNet** VGGNet, développé par Karen Simonyan et Andrew Zisserman en 2014, est connu pour son architecture profonde et homogène avec des couches convolutives de petite taille (3x3). Il a établi de nouvelles références de performances sur diverses tâches de classification d'images. [55]
- **GoogLeNet/Inception** GoogLeNet, également connu sous le nom d'Inception, a été proposé par l'équipe de recherche de Google en 2014. Il a introduit le concept de blocs Inception, qui permettent de capturer des informations à différentes échelles spatiales en utilisant des convolutions de différentes tailles.[57]
- **ResNet** ResNet, développé par Kaiming He et ses collègues en 2015, a introduit le concept de blocs résiduels. Ces blocs permettent de former des réseaux beaucoup plus profonds sans souffrir de la dégradation de la performance, ce qui a ouvert la voie à la création de réseaux encore plus profonds. [33]

Nous avons également abordé, dans le chapitre 5, plusieurs définitions sur les concepts utilisés dans notre approche, concernant l'entraînement et le test d'un modèle et de ses métriques ainsi que sur les données et ses notions (voir la section Préliminaires du 5ème chapitre V.2).

Chapitre III

Etat De L'Art

III.1 Introduction

Dans ce chapitre, notre attention sera portée sur l'application de l'IA dans le domaine de la neurochirurgie, plus précisément sur l'utilisation des réseaux de neurones convolutifs (CNN) pour la classification des tumeurs cérébrales.

L'IA, en particulier les techniques de deep learning, a ouvert de nouvelles perspectives dans le domaine médical, offrant des solutions innovantes pour le diagnostic et le traitement des maladies neurologiques complexes.

Nous aborderons cette problématique en examinant en détail un certain nombre d'articles pertinents qui ont contribué significativement à ce domaine. Ces articles présenteront leurs méthodologies, leurs approches et les résultats obtenus, nous permettant d'explorer les avancées récentes et les défis associés à l'utilisation des CNN pour la classification des tumeurs cérébrales.

En nous basant sur ces travaux de recherche, nous pourrions appréhender les différentes techniques et architectures de CNN utilisées, les ensembles de données utilisés pour l'entraînement et la validation, ainsi que les performances obtenues par ces modèles. De plus, nous mettrons en évidence les éventuelles limites et les perspectives futures dans ce domaine, en nous appuyant sur les contributions et les recommandations des chercheurs.

L'objectif de ce chapitre d'état de l'art est de fournir une vision approfondie et complète de l'utilisation des CNN dans la classification d'IRM des tumeurs du cerveau, en mettant en évidence les progrès réalisés et en identifiant les opportunités de recherche future. En examinant les différentes approches et les résultats obtenus, nous espérons contribuer à l'enrichissement des connaissances et à l'avancement de l'IA dans le domaine de la neurochirurgie.

III.2 Travaux connexes

Plusieurs études se sont concentrées sur la classification des tumeurs cérébrales en utilisant des modèles d'apprentissage profond et des images par IRM. Cependant, il est important de noter que aucun des articles mentionnés n'aborde spécifiquement la prédiction des interventions chirurgicales pour les tumeurs pétreuses. Au lieu de cela, ils se concentrent principalement sur la classification et le diagnostic des tumeurs cérébrales. Néanmoins, nous avons inclus ces articles dans notre recherche car ils utilisent des réseaux neuronaux à convolution (CNN) pour analyser des images par IRM du cerveau, ce qui est en accord avec l'objectif de notre travail.

III.2.1 Les approches de classification d'IRM du cerveau

- **Article 1 : An Efficient Method to Classify Brain Tumor using CNN and SVM [54]**

Cet article présente une méthode efficace pour la classification des tumeurs cérébrales en utilisant un réseau neuronal convolutif (CNN) et une machine à vecteurs de support (SVM). La détection et la classification des tumeurs cérébrales à partir d'images d'IRM constituent un défi majeur dans le domaine médical. Pour résoudre ce problème, une méthode basée sur un modèle CNN-SVM est proposée pour classer les tumeurs cérébrales avec une précision élevée. Un CNN composé de 19 couches est construit, comprenant des couches de convolution 2D, de max-pooling, FC et de normalisation par lots. L'algorithme de classification softmax est utilisé sur un ensemble de données de 3064 images de trois types de tumeurs cérébrales (gliome, méningiome et pituitary). Ensuite, une SVM est utilisée pour améliorer la précision du modèle CNN en utilisant les caractéristiques extraites du modèle. La précision finale de cette méthode basée sur le CNN-SVM est de 97,1 %.

En conclusion, cet article présente une méthode de classification efficace pour les tumeurs cérébrales en utilisant un CNN et un SVM. Les résultats montrent une précision élevée dans la classification des trois classes de tumeurs cérébrales sur un ensemble de données d'images IRM. Cette approche offre des perspectives prometteuses pour améliorer la détection et la classification des tumeurs cérébrales, ce qui ouvre de nouvelles possibilités dans le domaine de l'imagerie médicale et de la recherche en neurochirurgie.

- **Article 2 : Brain Tumor Classification Using Convolutional Neural Network [25]**

Cet article porte sur la classification des tumeurs cérébrales en utilisant un CNN. L'étude se concentre sur la classification des tumeurs cérébrales à partir d'images d'IRM pondérées en T1 avec contraste amélioré. Le système proposé comprend deux étapes majeures : le prétraitement des images à l'aide de différentes techniques de traitement d'image, suivi de la classification des images prétraitées par le CNN. Les résultats obtenus sur un ensemble de données de 3064 images comprenant trois types de tumeurs cérébrales (gliome, méningiome, pituitary) montrent une précision de test de 94,39 %, une précision d'entraînement de 93,33 % et un recall de 93 %. Les CNN se révèlent ainsi être des outils puissants pour l'analyse des images médicales et ouvrent la voie à de nouvelles avancées dans le domaine de la neurochirurgie et de l'oncologie.

- **Article 3 : Multi-Grade Brain Tumor Classification using Deep CNN with Extensive Data Augmentation [52]**

Cet article propose un système de classification des tumeurs cérébrales multi-grades basé sur un CNN pour l'analyse des images IRM. Le système propose une technique de segmentation pour extraire les régions tumorales des images IRM, ainsi qu'une augmentation extensive des données pour améliorer l'entraînement du modèle de classification. Un modèle CNN pré-entraîné est ensuite ajusté en utilisant les données augmentées pour la classification des tumeurs cérébrales en fonction de leur grade. Les résultats expérimentaux montrent que le système proposé offre des performances convaincantes par rapport aux méthodes existantes. Les résultats expérimentaux montrent une précision élevée par rapport aux techniques existantes.

- **Article 4 : Brain Tumor Classification using Deep CNN Features via Transfer Learning [27]**

Cet article se concentre sur la classification en trois classes des tumeurs cérébrales, à savoir les gliomes, les méningiomes et les tumeurs pituitary. Les chercheurs ont utilisé le Transfer Learning profond en utilisant un modèle pré-entraîné, GoogLeNet, pour extraire des caractéristiques à partir d'images d'IRM cérébrale. Les caractéristiques extraites ont ensuite été classifiées à l'aide de modèles de classification éprouvés. Le système proposé a enregistré une précision de classification de 98 %, surpassant toutes les méthodes de pointe dans le domaine.

Une des principales contributions de cet article est d'évaluer le système de classification avec un nombre réduit d'échantillons d'entraînement, ce qui est souvent une contrainte dans le domaine médical. Les résultats montrent que le système proposé produit des résultats acceptables même avec un nombre limité d'échantillons d'entraînement, démontrant ainsi l'utilité du Transfer Learning dans les cas de données médicales limitées. Cela ouvre des perspectives pour l'application de techniques d'apprentissage profond dans des scénarios où les images médicales sont rares.

- **Article 5 : Brain Tumor Classification Based on Fine-Tuned Models and the Ensemble Method [49]**

Cet article se concentre sur la classification des tumeurs cérébrales à l'aide de modèles DL et ML et de la méthode d'Ensembling. Les auteurs soulignent l'importance d'une détection précise et rapide des tumeurs cérébrales pour sauver des vies, en particulier pour les gliomes, les méningiomes et les tumeurs pituitary. Alors que de nombreuses études ont exploré l'analyse et la classification des tumeurs cérébrales, peu se sont intéressées au problème d'extraction de caractéristiques. Les techniques d'extraction de caractéristiques conventionnelles présentent des limites, et il est nécessaire de disposer de systèmes capables de le faire et de classer les tumeurs avec précision.

L'étude propose un modèle qui combine le DL et le ML pour l'évaluation des tumeurs cérébrales. Le DL est utilisé pour l'extraction des caractéristiques à l'aide de modèles CNN tels que Inception-v3 et Xception. La classification des tumeurs cérébrales est effectuée à l'aide de différents algorithmes, notamment une classification softmax, Random Forest, SVM, KNN, et la technique d'Ensembling. Les résultats de ces approches sont comparés à des méthodes existantes.

Le modèle Inception-v3 obtient les meilleures performances avec une précision de test de 94,34% par rapport à d'autres méthodes récemment rapportées. L'utilisation de techniques DL et ML pour l'extraction des caractéristiques et la classification montre des perspectives prometteuses pour améliorer le diagnostic des tumeurs cérébrales et soutenir les professionnels de la santé dans leur processus de prise de décision.

- **Article 6 : Pre-trained deep learning models for brain MRI image classification [38]**

Cet article présente une étude sur l'utilisation de modèles DL pré-entraînés pour la classification d'images par résonance magnétique (IRM) du cerveau. L'article met l'accent sur l'utilisation de modèles de CNN pré-entraînés tels que VGG-19, VGG-16, ResNet50 et Inception V3, en utilisant des techniques d'augmentation de données et de Transfer Learning. Les modèles sont évalués en termes d'exactitude, de recall, de précision et de score F1 sur un ensemble de tests. Les résultats montrent que le modèle pré-entraîné VGG-19 avec Transfer Learning présente les meilleures performances.

Une des principales contributions de l'article est de démontrer que ces modèles pré-entraînés peuvent effectuer une classification précise des images brutes sans nécessiter d'extraction manuelle de caractéristiques. De plus, l'utilisation du Transfer Learning permet d'obtenir de bons résultats même avec un petit ensemble de données d'entraînement, ce qui aide à prévenir le surapprentissage.

- **Article 7 : Holistic Brain Tumor Screening and Classification Based on DenseNet and Recurrent Neural Network [61]**

Cet article présente une méthode de dépistage et de classification holistique des tumeurs cérébrales basée sur le CNN DenseNet et les RNN. Les algorithmes proposés extraient d'abord des caractéristiques à partir d'IRMs en utilisant un CNN dense, puis ces caractéristiques séquentielles sont utilisées pour la classification à l'aide d'un RNN. Contrairement aux méthodes traditionnelles, cette approche ne nécessite pas de segmentation manuelle ou automatique des régions d'intérêt. Les résultats obtenus sur un ensemble de données public et une population de 422 IRMs montrent l'efficacité de la méthode proposée.

- **Article 8 : MRI Brain Tumor Classification Using a Hybrid VGG16-NADE Model [56]**

Dans cette article les auteurs ont proposé une approche hybride pour la classification des tumeurs cérébrales à partir d'IRMs pondérées en T1 avec contraste. Cette méthode combine le CNN VGG16 et NADE (Neural Autoregressive Distribution Estimator). Le modèle hybride VGG16-NADE a été entraîné sur un ensemble de données comprenant 3064 images de trois catégories de tumeurs cérébrales. Les résultats ont montré que cette approche hybride surpassait d'autres méthodes, avec une précision de classification de

96,01 % et des performances globales supérieures en termes de détection précoce du cancer du cerveau.

Le modèle hybride VGG16-NADE a été comparé à différentes approches telles que le CNN hybride, le NADE, le CNN avec KELM, le CNN avec augmentation de données et le CNN avec GA. Les résultats ont montré que le modèle hybride VGG16-NADE présentait une précision élevée de 95,72 %, un recall de 95,64 % et un score F1 de 95,68 %. Il surpassait les autres approches évaluées, avec une aire sous la courbe ROC de 0,91 et un MCC de 0,3564.

- **Article 9 : Brain Tumor Classification Using VGG-16 and MobileNetV2 Deep Learning Techniques on Magnetic Resonance Images (MRI) [51]**

Cet article vise à classer différentes tumeurs cérébrales en mettant en œuvre des techniques DL appliquées aux images d'IRM. Deux modèles CNN, à savoir VGG-16 et MobileNetV2, sont utilisés pour classifier les images d'IRM dans les quatre catégories suivantes : Gliome, méningiome, Notumor et pituitary. Dans cette étude, les échantillons de données complets sont répartis en ensembles d'entraînement (70 %), de validation (15 %) et de test (15 %). Les modèles DL VGG-16 et MobileNetV2 sont tous deux entraînés sur 50 époques. Les résultats de cette étude indiquent que le modèle VGG-16 a obtenu une précision globale de 91,46 %. La valeur la plus élevée du score F1 est de 96,87 %, obtenue pour la classe notumor. En revanche, le modèle MobileNetV2 l'a surpassé pour la classification des tumeurs cérébrales, avec une précision de 97,46 % et un score F1 maximum de 98,87 % pour la classe notumor, ainsi qu'un score F1 de 98,72 % pour la tumeur pituitary. Au cours de la phase de mise en œuvre, il a été observé que des paramètres tels que le nombre d'époques affectent significativement la précision du modèle. Il est donc recommandé de former le modèle avec un nombre suffisant d'époques pour obtenir des résultats de classification précis.

III.3 Tableau Comparatif

Dans les tableaux ci-dessous nous effectuerons une étude comparative des approches proposées ci-dessus selon les 7 facteurs suivants :

- Titre : désigne le titre de l'article et les auteurs.
- Dataset : Indique les datasets utilisées pour l'implémentation de l'approche proposée.
- Approche : désigne l'approche de chaque article.
- Résultats : Les résultats obtenue par chaque approche.
- Avantages : Les avantages de l'approche abordée.
- Inconvénients : Les inconvénients de l'approche abordée.

Titre	Dataset	Approches	Résultats	Avantages	Inconvénients
An Efficient Method to Classify Brain Tumor using CNN and SVM (Zarin Anjuman Sejuti1, Md Saiful Islam) [54]	Un ensemble de données comprenant trois classes différentes d'images (gliome, pituitary, méningiome)	Une architecture CNN avec plusieurs couches de convolution, normalisation de lot, activation ReLU, max-pooling, couches FC et classifieur softmax. Les caractéristiques extraites du CNN sont utilisées pour la classification avec une classification SVM.	Précision de 97,1%	Haute précision, approche robuste, pas de prétraitement majeur requis.	La pertinence des données est discutable étant donné qu'il s'agit d'un ensemble de données publiques, accordance d'un niveau de priorité relativement faible à l'augmentation des données.
Brain Tumor Classification Using Convolutional Neural Network (Sunanda Das, O.F.M. Riaz Rahman Aranya, Nishat Nayla Labiba) [25]	3064 images pondérées en T1 avec contraste de patients atteints de méningiome, gliome et pituitary.	Utilisation d'un CNN pour la classification des tumeurs cérébrales.	Précision de 92% et recall de 89% obtenus sur l'ensemble de test.	Utilisation d'un CNN pour la reconnaissance de motifs et le traitement d'images.	Manque de détails sur l'architecture spécifique du CNN utilisé. Pas de mention des performances sur les différentes classes de tumeurs.

<p>Multi-Grade Brain Tumor Classification using Deep CNN with Extensive Data Augmentation (Muhammad Sajjad) [52]</p>	<p>Le système utilise deux ensembles de données : le jeu de données Radiopaedia et le jeu de données Brain Tumor. Le jeu de données Radiopaedia comprend 121 images d'IRM avec des tumeurs cérébrales réparties en quatre grades différents. Le jeu de données Brain Tumor comprend 3064 images d'IRM avec des tumeurs cérébrales collectées auprès de 233 patients.</p>	<p>La méthode utilise un modèle CNN finement ajusté basé sur l'architecture VGG-19. Le système se compose de trois étapes principales : segmentation des tumeurs, augmentation des données et extraction/classification des caractéristiques profondes.</p>	<p>La précision obtenue varie en fonction des grades, avec des valeurs allant jusqu'à 95,5% pour le grade I, 92,66% pour le grade II, 87,77% pour le grade III et 86,71% pour le grade IV.</p>	<p>Utilisation d'un CNN pour une meilleure précision dans la classification des tumeurs cérébrales de plusieurs grades. Utilisation de l'augmentation des données pour augmenter. Utilisation d'une architecture finement ajustée (VGG-19) par Transfer Learning.</p>	<p>Le système dépend fortement de la disponibilité de données d'entraînement de haute qualité et en quantité suffisante.</p>
---	--	---	--	---	--

<p>Brain tumor classification using deep CNN features via transfer learning (S. Deepak, P.M. Ameer) [27]</p>	<p>Le dataset utilisé est composé de 3064 images d'IRM du cerveau provenant de 233 patients, diagnostiqués avec l'un des trois types de tumeurs cérébrales : méningiome, gliome et tumeurs pituitary. Les images sont disponibles au format .mat</p>	<p>Une approche Transfer Learning a été utilisée en utilisant le CNN pré-entraîné GoogLeNet. Les couches du réseau ont été adaptées au domaine cible des tumeurs cérébrales en modifiant les couches FC et en remplaçant les couches softmax de classification. Les caractéristiques extraites ont ensuite été utilisées pour l'apprentissage avec des classifieurs supplémentaires tels que SVM et KNN.</p>	<p>La précision de classification obtenue avec le modèle Transfer Learning seul était de 92,3%. Avec SVM, la précision était de 97,8%, et avec le classifieur KNN, elle était de 98,0%.</p>	<p>Utilisation de Transfer Learning. Bonne performance du système de classification SVM et KNN.</p>	<p>L'ensemble de données utilisé était déséquilibré en termes du nombre d'images pour chaque type de tumeur, ce qui peut affecter les performances du système de classification, accordance d'un niveau de priorité relativement faible à l'augmentation des données.</p>
---	--	--	---	---	---

<p>Brain Tumor Classification Based on Fine-Tuned Models and the Ensemble Method (Neelum Noreen1, Sellapan Palaniappan1, Abdul Qayyum2, Iftikhar Ahmad3 and Madini O. Alassafi3) [49]</p>	<p>Contient 3 064 coupes d'IRM cérébrale provenant de 233 patients différents. Comprend trois classes principales de tumeurs cérébrales : Méningiome, gliome et pituitary. Réparti en 930 images de tumeurs de pituitary, 708 images de gliomes et 1 426 images de gliomes tumoraux.</p>	<p>Modèles utilisés : Inception-v3, Xception, SVM, RF et KNN. Classification ensembliste : combinaison des modèles Inception-v3 et Xception avec KNN, SVM et RF, utilisant une méthode d'Ensembling.</p>	<p>Accuarcy des différents modeles : Modele Ensemble1 : 94,34% Modele Ensemble2 : 93,79%</p>	<p>Utilisation de l'Ensembling pour bénéficier des résultats positifs des modèles. Utilisation du Transfer Learning.</p>	<p>Les deux modèles ensembles n'ont pas atteint une valeur élevée, La pertinence des données est discutable étant donné qu'il s'agit d'un ensemble de données publiques.</p>
--	--	--	--	--	--

<p>Pre-trained deep learning models for brain MRI image classification (Srigiri Krishnapriya, Yepuganti Karuna) [38]</p>	<p>Le dataset utilisé dans cette étude comprenait un total de 305 images d'IRM du cerveau après augmentation. Parmi ces images, 98 ne présentaient pas de tumeurs et 155 présentaient des tumeurs.</p>	<p>Les étapes fondamentales de la méthode proposée comprennent le prétraitement des images, l'utilisation de modèles pré-entraînés tels que VGG-19, VGG-16, ResNet50 et InceptionV3, ainsi que l'augmentation des données pour améliorer les performances d'apprentissage.</p>	<p>VGG19 Accuracy : 99.48%, Recall : 98.76%, precision : 100%, F1 score : 99.17%. VGG16 accuracy : 99%, recall : 98.18%, precision : approximately 100%, F1 score : 99.08%. Resnet50 Accuarcy : 97.92%, recall : 87.7%, precision : 77.77%, F1score : 82.24%. InceptionV3 Accuarcy : 81.25%, recall : 63.25%, precision : 53.84%, F1 score : 58.16%</p>	<p>L'utilisation du Transfer Learning a permis d'obtenir de bons résultats même avec un petit ensemble de données. L'augmentation des données a contribué à améliorer les performances d'apprentissage et à éviter le surajustement.</p>	<p>La taille relativement petite du dataset utilisé dans cette étude peut limiter la généralisation des résultats.</p>
---	--	--	--	--	--

<p>Holistic Brain Tumor Screening and Classification Based on DenseNet and Recurrent Neural Network (Yufan Zhou¹, Zheshuo Li¹, Hong Zhu², Changyou Chen¹, Mingchen Gao¹, Kai Xu³, and Jinhui Xu) [61]</p>	<p>Deux ensembles de données : un ensemble de données public et un ensemble de données propriétaire. L'ensemble de données public comprend 3064 images d'IRM de cerveau provenant de 233 patients, avec différentes types de tumeurs cérébrales. L'ensemble de données propriétaire comprend 422 scans images d'IRM diagnostiqués comme normaux ou avec différents types de tumeurs cérébrales (gliomes, méningiomes et métastases).</p>	<p>Deux modèles sont proposés : DenseNet-LSTM et DenseNet-DenseNet. DenseNet-LSTM utilise un auto-encodeur basé sur DenseNet pour extraire des caractéristiques des images originales, puis utilise un RNN pour la classification des séquences d'images. DenseNet-DenseNet empile les caractéristiques des séquences d'images en tant que tenseur 2D et utilise un autre réseau DenseNet pour la classification.</p>	<p>Dataset public Densenet-RNN : 84.61%, Densenet-LSTM : 92.13%, Densenet-Denset : 86.88%, Dataset privé Densenet-RNN : 60.00%, Densenet-LSTM : 71.10%, Densenet-Denset : 64.95%</p>	<p>L'approche proposée permet de traiter les séquences d'images 3D de manière holistique, sans avoir besoin d'étiqueter chaque image individuellement.</p>	<p>Les valeurs accuracy du dataset privé sont assez faibles.</p>
--	--	---	--	--	--

<p>MRI Brain Tumor Classification Using a Hybrid VGG16-NADE Model (Saran Raj Sowrirajan ,Surendiran Balasubramanian, Raja Soosaimarian Peter Ra) [56]</p>	<p>L'ensemble de données utilisé est un ensemble de données sur les tumeurs cérébrales en images d'IRM de tumeurs : méningiomes, gliomes et tumeurs pituitary.</p>	<p>L'approche utilisée consiste à appliquer l'augmentation des données et les processus d'apprentissage sur l'ensemble de données des tumeurs cérébrales en IRM. Le modèle hybride NADE (Neural Autoregressive Distribution Estimator) et VGG16 est entraîné sur les échantillons et comparé à d'autres approches existantes telles que le CNN standard, le CNN avec KELM, le CNN avec augmentation des données et le CNN avec GA.</p>	<p>Le modèle hybride NADE et VGG16 obtient un taux de précision de prédiction de 96,01%,</p>	<p>Extraction d'informations globalement distribuées grâce à un noyau plus large dans l'architecture Inception, utilisation du Transfer Learning, utilisation de l'augmentation de données.</p>	<p>La pertinence des données est discutable étant donné qu'il s'agit d'un ensemble de données publiques.</p>
--	--	--	--	---	--

Brain Tumor Classification Using VGG-16 and MobileNetV2 Deep Learning Techniques on Magnetic Resonance Images (MRI) (Rashmi Saini, Prabhakar Semwal, Tushar Hrishikesh Jaware) [51]	Le dataset est composé d'images d'IRM (Imagerie par Résonance Magnétique) du cerveau. Les images d'IRM sont classées en quatre catégories : Gliome, Méningiome, notumor et pituitary.	Deux modèles d'apprentissage profond, VGG-16 et MobileNetV2, sont utilisés pour classer les images d'IRM en quatre catégories de tumeurs cérébrales.	Le modèle VGG-16 a obtenu une précision globale de 91,46% et le modèle MobileNetV2 a atteint une précision de 97,46%.	La précision de 97,46% et l'utilisation du Transfer Learning.	La pertinence des données est discutable étant donné qu'il s'agit d'un ensemble de données publiques, accordance d'un niveau de priorité relativement faible à l'augmentation des données.
---	---	--	---	---	--

TABLE III.1 – Tableau de comparaison des articles de l'état de l'art

III.4 Conclusion

En conclusion, dans ce chapitre les articles examinés portent sur la classification des tumeurs cérébrales en utilisant des techniques DL, principalement des CNN. Les résultats obtenus dans ces études montrent une précision élevée dans la classification des tumeurs cérébrales, ce qui suggère que l'apprentissage profond peut être une approche efficace pour le diagnostic et la classification de ces affections médicales.

Les méthodes proposées dans les articles combinent souvent des techniques telles que le Transfer Learning, l'augmentation des données et l'Ensembling. L'utilisation de modèles pré-entraînés permet d'extraire automatiquement des caractéristiques pertinentes à partir des images sans nécessiter une extraction manuelle. De plus, l'augmentation des données permet d'augmenter la taille de l'ensemble d'entraînement, ce qui peut améliorer la capacité du modèle à généraliser et à classer avec précision les tumeurs cérébrales.

Après avoir pris tous ces éléments en considération, nous avons pris la décision d'utiliser trois modèles pour traiter notre problématique. Plus précisément, nous avons choisi d'utiliser les modèles VGG16, DenseNet121 et InceptionV3 pour extraire les caractéristiques des images d'IRM et les utiliser dans notre processus de classification. Ces modèles sont largement utilisés et ont démontré de bonnes performances dans le domaine de la vision par ordinateur, notamment pour la classification d'images IRM. Nous avons également pris en considération l'utilisation d'un Ensembling afin de bénéficier simultanément de la précision des trois modèles. En utilisant ces modèles, nous espérons obtenir des résultats précis et fiables dans notre tâche de classification d'images d'IRM.

En conclusion, les approches basées sur le DL, telles que l'utilisation de CNN et de techniques ML, offrent de nouvelles perspectives pour la détection et la classification des tumeurs cérébrales à partir d'images IRM.

Chapitre IV

Proposition de notre approche

IV.1 Introduction

Dans ce chapitre, nous proposons une approche basée sur le Transfer Learning pour résoudre notre problème de classification en deux classes : "voie combinée" et "voie transpétreuse". Cette approche permet de bénéficier des caractéristiques extraites par le modèle pré-entraîné et d'éviter l'entraînement complet d'un modèle à partir de zéro, ce qui peut être coûteux en termes de temps et de ressources.

Nous proposons d'utiliser trois modèles de référence pour le Transfer Learning : VGG16, Densenet121 et InceptionV3. Ces modèles ont été pré-entraînés sur des datasets massifs tels que ImageNet, ce qui leur confère une capacité à extraire des caractéristiques pertinentes à partir d'images. Nous présenterons brièvement chacun de ces modèles, y compris leur définition, et leur architecture.

Ensuite, nous décrirons notre approche pour l'adaptation de ces modèles à notre tâche spécifique. Nous ajouterons des couches supplémentaires à l'architecture de chaque modèle pour les adapter à la classification en deux classes. Nous expliquerons comment nous avons configuré ces couches supplémentaires, y compris l'ajout d'une couche de Global Average Pooling, d'une couche de Dropout et d'une couche dense avec une activation softmax. Ces modifications permettront aux modèles de générer des prédictions pour nos classes cibles : "voie combinée" et "voie transpétreuse".

Enfin, nous discuterons de notre approche d'Ensembling pour combiner les prédictions des trois modèles et obtenir une prédiction finale plus robuste et précise. Dans les sections suivantes, nous présenterons plus en détail chaque modèle de référence, décrirons notre approche pour leur adaptation et expliquerons comment nous avons utilisé l'Ensembling pour combiner leurs prédictions.

IV.2 Proposition de modeles pour le Transfer Learning

IV.2.1 Proposition 1 : VGG16

- **Definition**

Le modèle VGG16 est une architecture de CNN largement utilisée en vision par ordinateur et en ML. Il a été proposé par Simonyan et Zisserman de l'Université d'Oxford en 2014. "VGG" fait référence à l'abréviation de Visual Geometry Group, le groupe de recherche auquel appartiennent les auteurs.

VGG16 est caractérisé par sa profondeur, comprenant 16 couches de convolutions et de couches entièrement connectées. Il utilise des convolutions 3x3 avec un pas de 1 et des couches de pooling 2x2 pour extraire progressivement les caractéristiques des images. Les couches de convolution sont suivies de couches entièrement connectées qui servent de classificateurs.

L'une des particularités de VGG16 est que toutes les couches convolutives utilisent des filtres de taille 3x3 avec un padding de 1, ce qui permet de maintenir la taille des images et d'éviter une réduction trop rapide de la résolution.

VGG16 a été pré-entraîné sur un grand ensemble de données d'images appelé ImageNet, qui comprend des millions d'images appartenant à différentes classes. En raison de sa profondeur et de son architecture simple, ça a été largement utilisé comme base pour le Transfer Learning, où les poids pré-entraînés du modèle sont utilisés comme point de départ pour des tâches de classification. [55]

- **Architecture** L'architecture du modèle VGG16 :

1. Couche d'entrée : Cette couche reçoit les images en entrée avec une taille fixe (224x224 pixels).
2. Couches convolutives : Le modèle VGG16 comprend un total de 13 couches convolutives, chacune suivie d'une fonction d'activation ReLU pour introduire la non-linéarité. Les filtres utilisés sont des convolutions 3x3 avec un pas de 1 et un padding de 1 pour maintenir la taille des images.
3. Couches de pooling : Après chaque paire de couches convolutives, une couche de pooling de type max pooling est appliquée avec une fenêtre de 2x2 et un pas de 2, réduisant ainsi la taille spatiale des caractéristiques.
4. Couches entièrement connectées : Après les couches convolutives et les couches de pooling, le modèle VGG16 comprend trois couches entièrement connectées

approche favorise le partage et l'utilisation optimale des caractéristiques extraites à tous les niveaux du réseau.

Densenet121 a démontré de bonnes performances dans des tâches de classification d'images, de détection d'objets et de segmentation sémantique. Il a été pré-entraîné sur des datasets massifs, tels que ImageNet, qui contiennent des millions d'images provenant de différentes catégories. En utilisant des techniques de transfert de connaissances, ce modèle peut être adapté à des tâches spécifiques en fine-tuning ou en entraînant uniquement les dernières couches du réseau. [34]

• **Architecture** : L'architecture de base de Densenet121, est la suivante :

1. Convolution initiale : Convolution 7x7, stride 2, nombre de filtres : 64
2. MaxPooling : Fenêtre 3x3, stride 2
3. Bloc Dense 1 : Composé de plusieurs couches de bottleneck layers avec une croissance du nombre de filtres.
4. Transition 1 : Compression du nombre de filtres, Réduction de dimension (1x1 Convolution) et Pooling (2x2 avec un stride de 2)
5. Bloc Dense 2 : Composé de plusieurs couches de bottleneck layers avec une croissance du nombre de filtres.
6. Transition 2 : Compression du nombre de filtres, Réduction de dimension (1x1 Convolution) et Pooling (2x2 avec un stride de 2)
7. Bloc Dense 3 : Composé de plusieurs couches de bottleneck layers avec une croissance du nombre de filtres.[34]

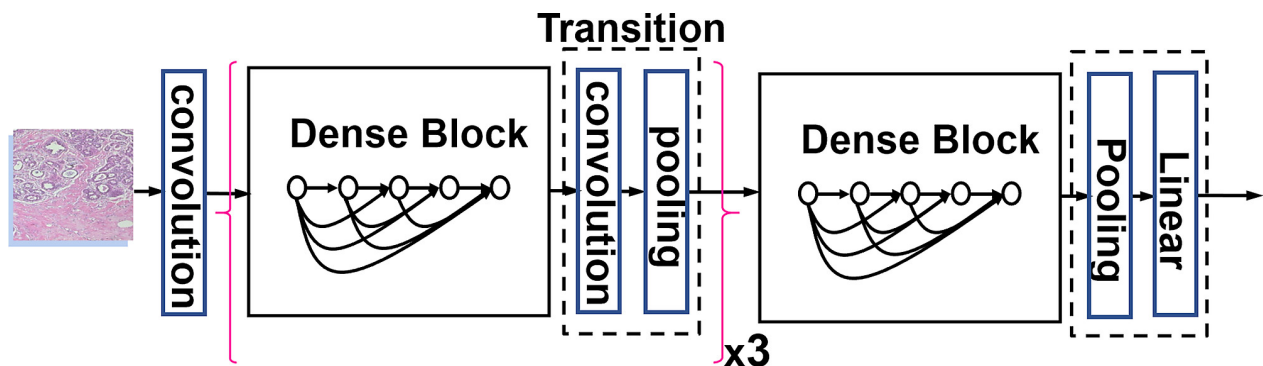


FIGURE IV.2 – Représentation de l'architecture du modèle Densenet121 [43]

IV.2.3 Proposition 3; InceptionV3

- **Definition** : Le modèle InceptionV3 est une architecture de réseau neuronal convolutif (CNN) qui a été développée pour la classification d'images. Il a été introduit par Christian Szegedy et al. en 2015. L'architecture InceptionV3 est basée

IV.3 Nos Modeles et Ensembling

Pour construire nos modèles, nous avons exploité la puissance du Transfer Learning en utilisant la bibliothèque TensorFlow pour ajouter des couches supplémentaires à l'architecture des trois modèles mentionné précédemment, Voici comment nous avons procédé :

Tout d'abord, nous avons ajouté une couche de Global Average Pooling pour réduire les dimensions des caractéristiques extraites. Ensuite, nous avons introduit une couche de Dropout avec un taux de 0,8 pour régulariser le modèle et éviter le surapprentissage. Enfin, nous avons ajouté une couche dense avec une fonction d'activation "softmax" pour effectuer la classification. Cette couche dense a été spécifiée avec 2 neurones pour la classification binaire des classes "voie combinée" et "voie transpétreuse" de notre dataset ou 4 neurones pour la classification des 4 classes du dataset publique.

Ainsi, en ajoutant ces couches supplémentaires et en configurant la dernière couche dense pour une classification à 4 puis à 2 classes, nous avons adapté le modèle à notre tâche spécifique.

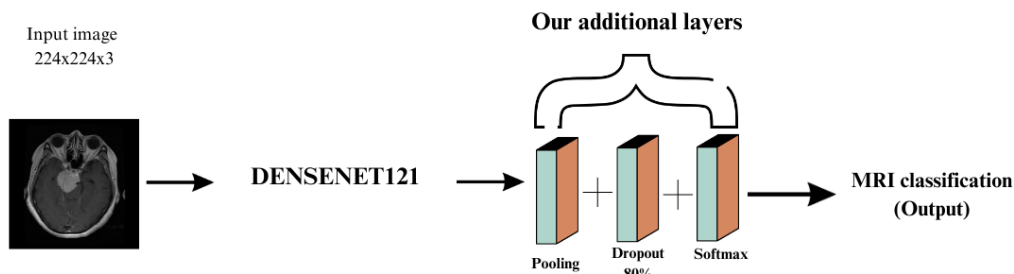


FIGURE IV.4 – Représentation de l'architecture de notre modèle densenet121

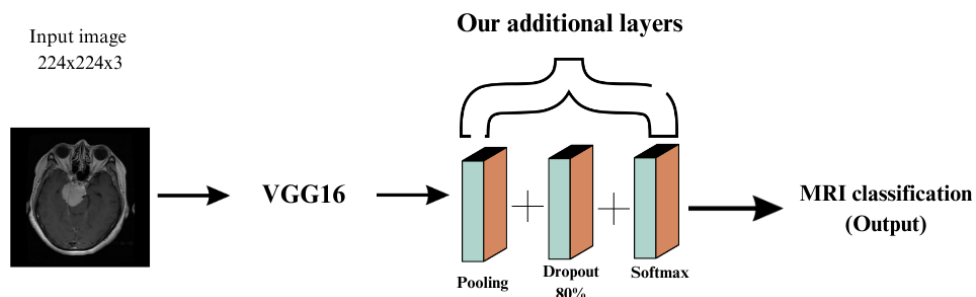


FIGURE IV.5 – Représentation de l'architecture de notre modèle VGG16

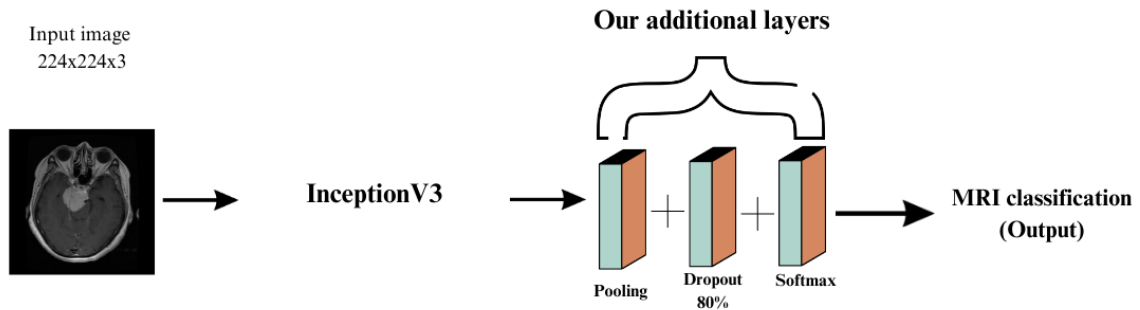


FIGURE IV.6 – Représentation de l’architecture de notre modèle InceptionV3

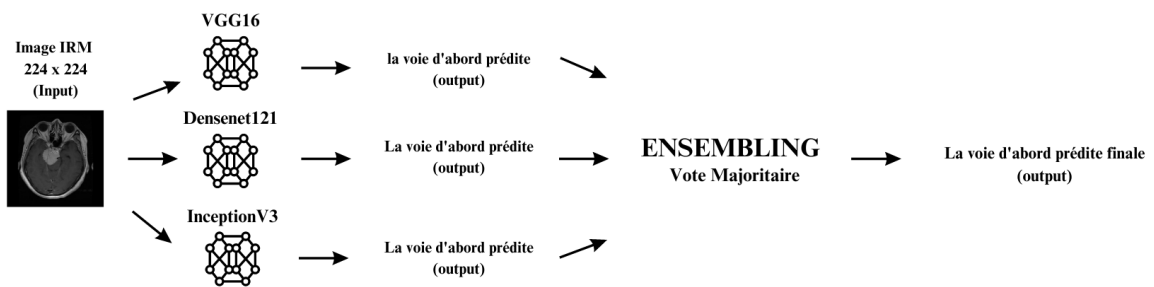


FIGURE IV.7 – Représentation de l’architecture de notre modèle Ensembling de nos trois modèles

Test et résultats de notre proposition

V.1 Introduction

Dans ce chapitre, nous présentons les résultats de notre proposition de modèle de classification pour les tumeurs cérébrales basé sur les CNN et le Transfer Learning. Nous commencerons par décrire le dataset utilisé, qui comprend à la fois un dataset public et notre propre dataset préparé spécifiquement pour cette étude. Ensuite, nous détaillerons les étapes de prétraitement et d'augmentation de données effectuées sur les deux datasets. Par la suite, nous présenterons les différentes métriques de performance utilisées pour évaluer les modèles, telles que l'accuracy, validation accuracy, Loss, validation Loss, ainsi que les epochs d'entraînement et les courbes d'apprentissage. Enfin, nous discuterons des résultats obtenus et analyserons les performances de notre approche.

V.2 Préliminaires

V.2.1 Entraînement et validation

les metriques d'evaluation

- **Accuracy** L'accuracy (exactitude) est une mesure de performance couramment utilisée en ML pour évaluer la performance d'un modèle. Elle représente le rapport entre le nombre de prédictions correctes et le nombre total d'échantillons.

L'accuracy est calculée en divisant le nombre de prédictions correctes par le nombre total d'échantillons et en multipliant le résultat par 100 pour obtenir un pourcentage. Une accuracy de 100% signifie que toutes les prédictions du modèle sont correctes, tandis qu'une accuracy de 0% indique que toutes les prédictions sont incorrectes. [23]

- **Validation Accuracy** La `val_accuracy` (exactitude de validation ou de test) mesure la capacité d'un modèle à donner des prédictions correctes sur un ensemble de données

de validation indépendant. Elle est calculée de la même manière que l'accuracy. [23]

- **Loss** La perte est une mesure de l'erreur entre les prédictions d'un modèle et les valeurs réelles. Elle quantifie la différence entre la sortie prédite et la sortie réelle et est utilisée pour guider l'optimisation du modèle. [23]
- **Validation Loss** La val_Loss est la perte calculée sur un ensemble de données de validation indépendant, qui est utilisé pour évaluer les performances du modèle sur des données non vues pendant l'entraînement. Elle permet de détecter l'éventuelle suradaptation du modèle aux données d'entraînement. [23]
- **Recall** Le recall (rappel) est une mesure de la capacité d'un modèle à identifier correctement les instances positives parmi toutes les instances réellement positives. Il est calculé en divisant le nombre de vrais positifs par la somme des vrais positifs et des faux négatifs. [26]
- **Precision** La précision est une mesure de la capacité d'un modèle à donner des prédictions positives correctes parmi toutes les prédictions positives effectuées. Elle est calculée en divisant le nombre de vrais positifs par la somme des vrais positifs et des faux positifs. [26]
- **F1 score** Le F1-score est une mesure de performance couramment utilisée en ML pour évaluer la précision et le rappel d'un modèle de classification. Il est calculé en prenant en compte à la fois la précision et le rappel afin de fournir une mesure équilibrée de la performance du modèle. [30]

Le F1-score est calculé à partir de la formule suivante :

$$\text{F1-score} = 2 * (\text{précision} * \text{rappel}) / (\text{précision} + \text{rappel})$$

Traitement de données

- **Données** Les données désignent des informations brutes, non traitées ou non structurées qui sont collectées, enregistrées ou générées dans le cadre d'une étude, d'une analyse ou d'un processus. Les données peuvent être de différents types, tels que des textes, des images, des vidéos, des enregistrements audio ou des mesures numériques. Dans le contexte du ML, les données sont essentielles pour entraîner des modèles et obtenir des résultats précis. [15]
- **Étiquetage et classification de données** L'étiquetage des données est le processus d'assignation d'étiquettes ou de balises à des données afin de les catégoriser, de les annoter ou de les classer. Cette étape est cruciale dans ML. Notamment l'apprentissage supervisé, car elle permet d'associer des informations significatives aux données, facilitant ainsi leur utilisation dans des tâches d'analyse et de modélisation.

L'étiquetage des données peut être effectué manuellement par des experts humains, ou automatiquement à l'aide de techniques ML [44]. Dans le processus manuel, des annotateurs formés examinent les données et appliquent des étiquettes en fonction des critères prédéfinis. Dans le processus automatique, des algorithmes ML sont utilisés pour prédire ou estimer les étiquettes en se basant sur des données d'entraînement préalablement étiquetées.

- **Prétraitement des données** Le prétraitement des données fait référence à l'ensemble des opérations appliquées aux données brutes afin de les préparer pour l'analyse ou l'ML. Cela peut inclure des étapes telles que le nettoyage des données en supprimant les valeurs manquantes ou aberrantes, la normalisation pour mettre les données à une échelle commune, la réduction de dimension pour extraire les caractéristiques les plus importantes, et d'autres techniques visant à améliorer la qualité et la cohérence des données. [6]
- **Augmentation de données** L'augmentation de données est une technique utilisée pour augmenter la quantité de données d'apprentissage en générant de nouvelles données à partir des données existantes. Cela peut inclure des opérations telles que la rotation, le recadrage, le redimensionnement, le renversement horizontal/vertical, le bruitage, etc. L'objectif est d'enrichir le jeu de données d'apprentissage et d'améliorer la généralisation du modèle en exposant celui-ci à une plus grande variété de cas. [17]

Entraînement du Modèle

L'entraînement fait référence au processus d'apprentissage d'un modèle à partir des données d'entraînement. Pendant l'entraînement, le modèle ajuste ses poids et ses paramètres en utilisant des algorithmes d'optimisation afin de minimiser l'erreur entre les prédictions du modèle et les valeurs réelles des données d'entraînement. L'objectif de l'entraînement est de permettre au modèle de généraliser à de nouvelles données et de produire des prédictions précises. [19]

Hyperparamètres

Les hyperparamètres sont des paramètres qui ne sont pas appris directement par le modèle pendant l'entraînement, mais qui déterminent la configuration et le comportement du modèle. Ils sont généralement définis avant le processus d'entraînement et peuvent inclure des paramètres tels que le taux d'apprentissage, la taille du lot (batch size), le nombre d'itérations, etc. Le choix approprié des hyperparamètres est crucial pour obtenir un modèle performant. [19]

Overfitting et Underfitting

L'Overfitting et l'Underfitting sont deux phénomènes courants dans le ML qui peuvent affecter les performances des modèles prédictifs.

L'Overfitting se produit lorsqu'un modèle ML est trop complexe et s'adapte excessivement aux données d'entraînement, capturant ainsi le bruit et les variations aléatoires présents dans ces données. Cela se traduit par une mauvaise généralisation du modèle, ce qui signifie qu'il ne parvient pas à bien se comporter sur de nouvelles données qui n'ont pas été utilisées pour son entraînement. En d'autres termes, le modèle "apprend par cœur" les données d'entraînement, mais il ne parvient pas à saisir les motifs et les structures sous-jacentes qui se généralisent à d'autres situations.

À l'inverse, l'Underfitting se produit lorsque le modèle est trop simple et ne parvient pas à capturer les motifs complexes des données d'entraînement. Cela se traduit par une sous-représentation des données et une faible performance prédictive. [19]

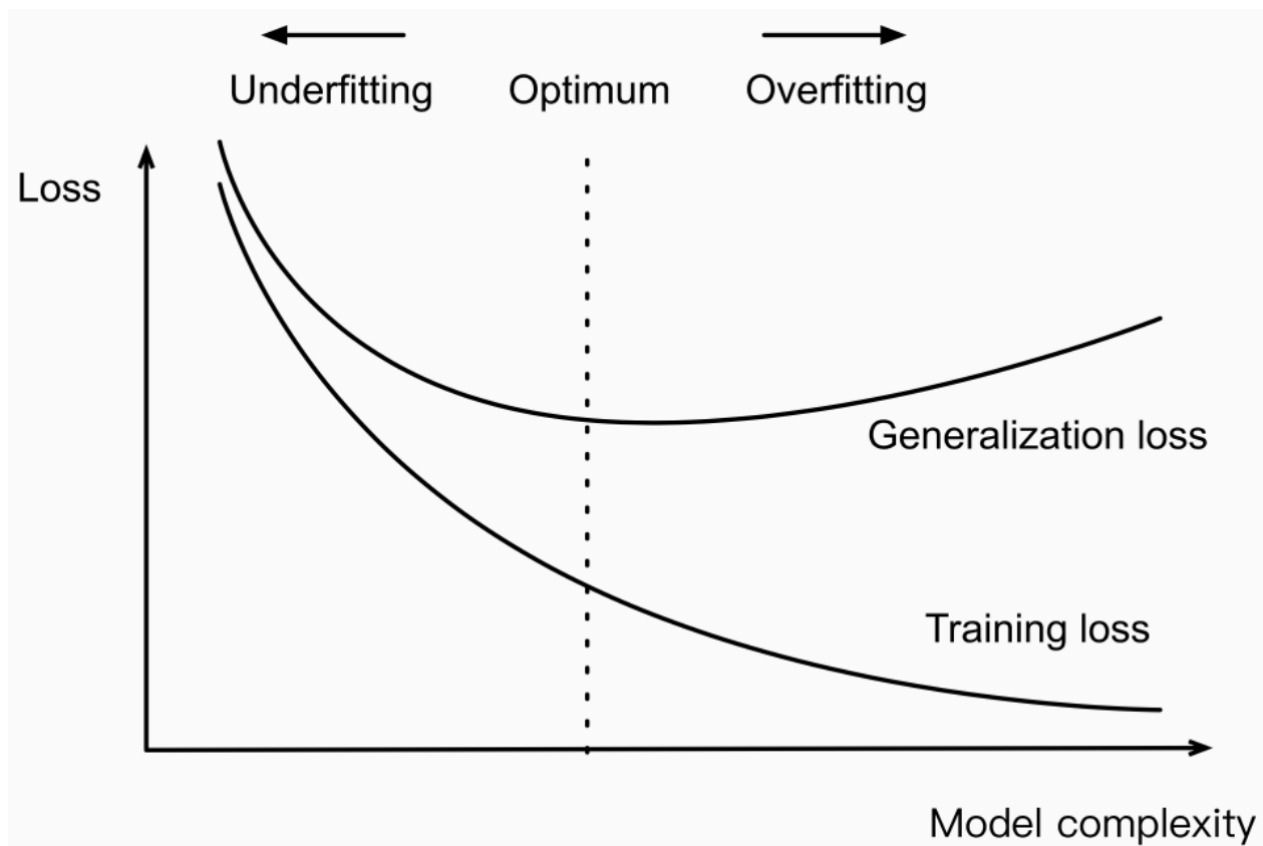


FIGURE V.1 – Représentation d'un underfitting et d'un Overfitting
[22]

V.2.2 Techniques et concepts avancés du DL

- **Callbacks** Un callback est une fonctionnalité qui permet d'exécuter des actions spécifiques à des étapes prédéfinies de l'entraînement du modèle, comme l'enregistrement des meilleurs poids, l'arrêt précoce en cas de stagnation de l'apprentissage, etc. [32]

- **Optimiseurs** Les optimiseurs sont des algorithmes utilisés dans le ML pour ajuster les paramètres d'un modèle afin de minimiser une fonction de perte ou d'erreur. Leur objectif est d'optimiser les performances du modèle en trouvant les valeurs optimales des paramètres.

Différents types d'optimiseurs existent, chacun ayant ses propres caractéristiques et avantages. Certains exemples courants d'optimiseurs sont la descente de gradient stochastique (SGD), l'optimiseur Adam, l'optimiseur RMSprop, et l'optimiseur Adagrad, parmi d'autres. [32]

- **Fonctions d'Activation**

Les fonctions d'activation sont des fonctions mathématiques appliquées à la sortie des neurones et introduisent une non-linéarité en transformant la somme pondérée des entrées en une valeur de sortie. Elles permettent aux réseaux de neurones de modéliser des relations non linéaires et d'effectuer des opérations complexes telles que la classification, la régression et la génération de séquences.

Il existe plusieurs fonctions d'activation couramment utilisées, telles que la fonction sigmoïde, la fonction tangente hyperbolique, la fonction ReLU (Rectified Linear Unit) et ses variantes, la fonction softmax, etc. [32]

- **Nombre de Neurones et de Couches (Architecture)** L'architecture du modèle fait référence à la structure et à la disposition des neurones et des couches dans le modèle. Le nombre de neurones et de couches peut varier en fonction de la complexité de la tâche. Une architecture bien choisie peut améliorer les performances du modèle en permettant une représentation plus riche et adaptée aux données. [32]

- **Dropout** Dropout est une technique de régularisation utilisée pour réduire le surapprentissage en désactivant aléatoirement certains neurones pendant l'entraînement. Cela encourage le modèle à développer une représentation robuste et évite la dépendance excessive à certains neurones spécifiques. [18]

V.3 Dataset

V.3.1 Données utilisées

Dataset public

Ce dataset [47] est composé d'une combinaison de trois autres datasets : figshare, SARTAJ et Br35H.

Il contient 7023 images d'IRM du cerveau humain, classifiées en quatre catégories : gliome, méningiome, notumor et pituitary.

Les images de la catégorie notumor proviennent spécifiquement du dataset Br35H.

Notre dataset

Nous avons utilisé des données open-source car nous n'avions pas suffisamment de cas réels disponibles. Ce sont des images IRM provenant d'une datasets de Kaggle[29], recueillies auprès de différents patients et représentent différentes perspectives et coupes du cerveau. Chaque image était associée à une étiquette de classe indiquant le type spécifique de tumeur présente. Parmi ces classes, nous avons spécifiquement sélectionné les méningiomes et schwannomes.

qu'il s'agisse d'un méningiome ou d'un schwannome. Ces images ont été soigneusement sélectionnées car elles représentent les types de tumeurs cérébrales de la région péri-pétreuses nécessitant une intervention chirurgicale par les voies d'abord étudiées. Ainsi, l'utilisation de ces images IRM constitue une étape cruciale dans le développement de notre dataset étiquetée par voie d'abord.

V.3.2 Tri et labellisation de notre dataset

Après l'obtention de deux ensembles de données distincts : l'ensemble de données Méningiomes et l'ensemble de données Schwannomes. Pour faciliter le tri et l'étiquetage de ces images, nous avons développé une application web personnalisée avec des fonctionnalités spécifiques. Cette application offrait la possibilité de supprimer les images non pertinentes, permettant ainsi de sélectionner uniquement les images significatives pour notre étude.

En utilisant l'application web, le neurochirurgien expert Pr. Himeur, a joué un rôle crucial dans le processus de tri et d'étiquetage. Il a examiné attentivement chaque image et a pris la décision de la supprimer si elle n'était pas pertinente pour notre étude. En outre, le Pr. Himeur a étiqueté chaque image en fonction de sa voie d'abord, en choisissant entre les options "voie combinée" et "voie trans-pétreuse". Cette étape d'étiquetage a été essentielle pour former notre ensemble de données final prêt à être utilisé dans notre tâche de classification.

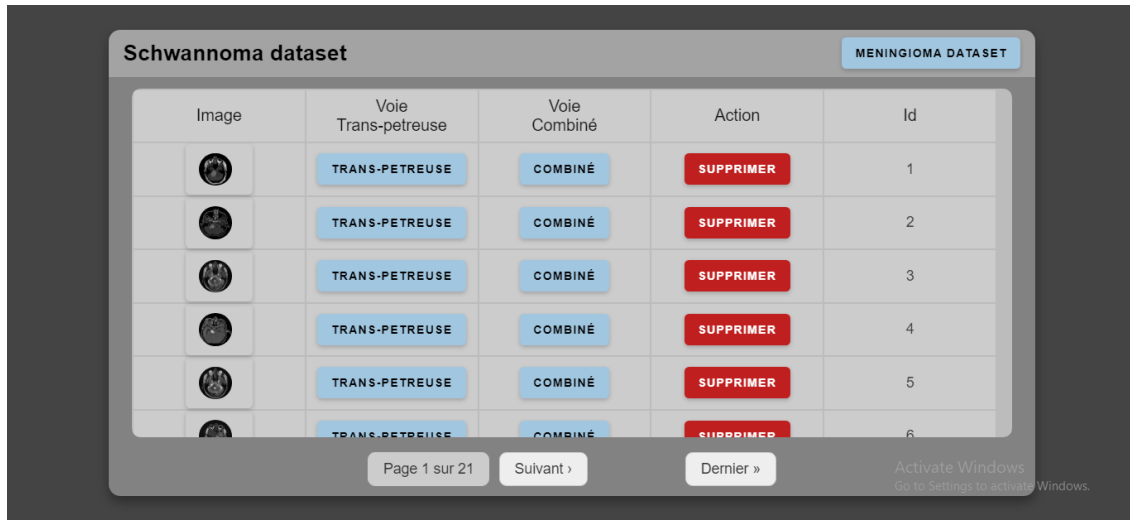


FIGURE V.2 – Interface de classification des images IRM en voie d’abord sur notre application

À la suite de ce processus de tri et d’étiquetage, nous avons obtenu un ensemble de données relativement restreint mais significatif. L’ensemble de données Meningioma comprenait seulement 273 images étiquetées dans la voie combinée, tandis que l’ensemble de données Schwannoma comprenait 436 images étiquetées dans la voie trans-pétreuse. Bien que le nombre d’images puisse sembler limité, il est important de souligner que ces données ont été soigneusement sélectionnées et étiquetées par le Pr. Himeur, un neurochirurgien expert. Malgré leur taille modeste, ces ensembles de données ont été préparés avec rigueur, garantissant la qualité et la pertinence des données utilisées pour nos expérimentations et nos modèles de classification.

V.3.3 Prétraitement et Augmentation

Prétraitement du dataset public

A part le redimensionnement de toutes les images en 224x224 pour être adéquat aux modèles, nous n’avons pas effectué de prétraitement car le dataset est déjà bien prétraité.

Prétraitement de notre dataset

Nous avons effectué plusieurs étapes pour préparer notre dataset en vue de l’entraînement et du test de nos modèles. Tout d’abord, nous avons développé un script personnalisé qui nous a permis de supprimer les images en double en se basant sur les valeurs RVB de chaque image. Ensuite, nous avons converti toutes les images en un format de codage commun, à savoir le format JPG.

Pour assurer la compatibilité avec le modèle que nous avons choisi, nous avons redimensionné toutes les images en une résolution de 224x224 pixels. Cette étape était nécessaire pour garantir que les images étaient adaptées à l’architecture spécifique du modèle.

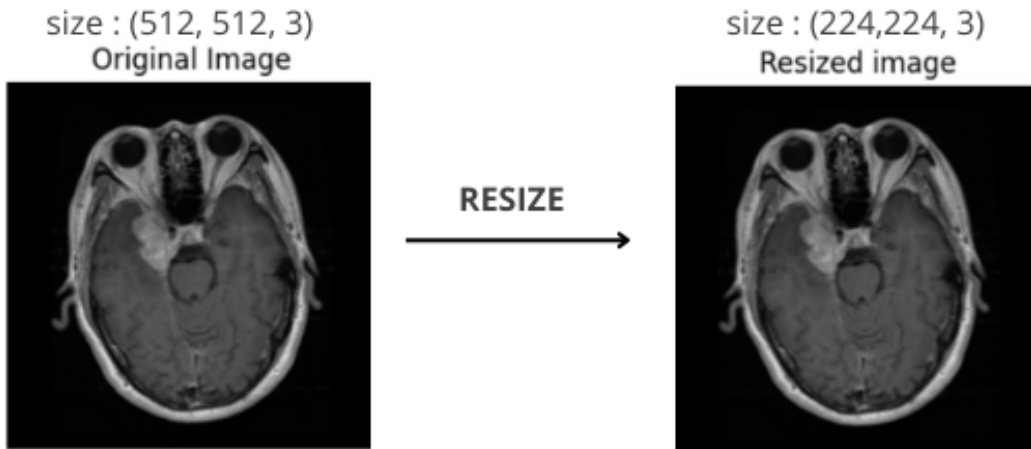


FIGURE V.3 – Représentation du redimensionnement d’une image en 224 x 224

Ensuite, nous avons divisé notre ensemble de données en deux parties distinctes : 80% pour l’entraînement et 20% pour les tests. Cette séparation nous a permis d’évaluer la performance de nos modèles sur des données non vues auparavant, assurant ainsi une évaluation impartiale de leur efficacité.

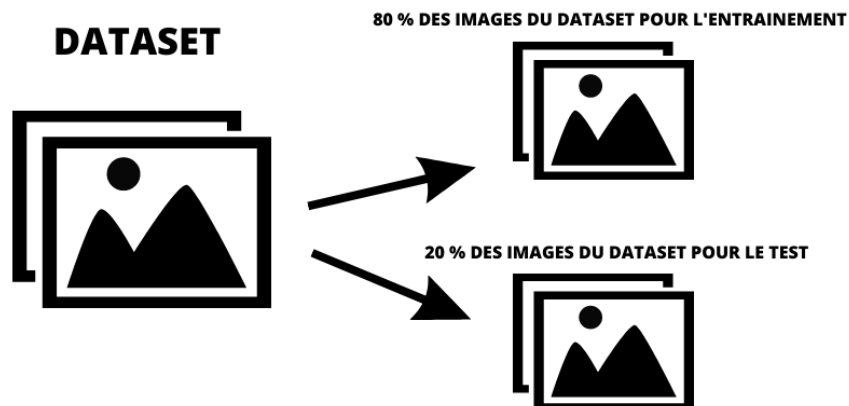


FIGURE V.4 – Représentation de la division de notre dataset pour l’entraînement et le test

Augmentation de données pour les deux datasets

Nous avons utilisé différentes techniques d'augmentation de données. Celles-ci comprennent des opérations telles que le changement de perspective, le zoom, la rotation, le retournement horizontal, la modification de la luminosité, ainsi que le décalage de largeur et de hauteur. Ces transformations ont permis d'augmenter la taille de notre ensemble d'entraînement tout en préservant les caractéristiques importantes des images. Ce qui est impératif pour notre dataset en vue du peu de données qu'on détient et les risques de surapprentissage.

Enfin, nous avons appliqué une étape de normalisation aux données, afin d'harmoniser les valeurs de pixels et d'optimiser l'apprentissage des modèles. Nous avons vérifié que cette normalisation était appliquée tant aux données d'entraînement qu'aux données de test, afin de maintenir la cohérence des échelles utilisées.

L'ensemble de ces étapes de prétraitement et d'augmentation des données garantit que notre ensemble de données est prêt à être utilisé dans l'entraînement et l'évaluation de nos modèles de classification.

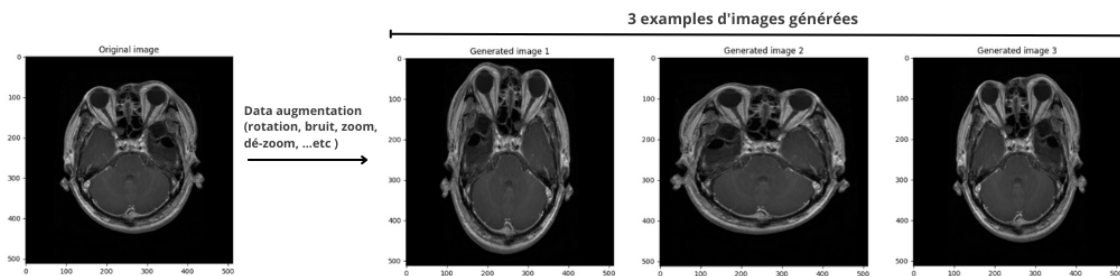


FIGURE V.5 – Représentation d'un exemple d'augmentation de données sur une image

V.4 Entraînement et test de nos modèles

Nous avons procédé à l'entraînement et au test de notre modèle en utilisant tout d'abord les données publiques classifiées en quatre catégories : glioma, méningioma, notumeur et Pituitary. Ensuite, sur les données étiquetées par le Pr Himeur pour les classes "voie transpétreuse" et "voie combinée" (notre dataset). Cela nous a permis d'élargir notre champ d'application et de tester l'efficacité de notre modèle sur des données provenant de différentes sources. L'entraînement a été effectué sur un total de 30 époques. Nous avons utilisé une méthode de validation croisée en utilisant à la fois un ensemble d'entraînement et un ensemble de validation.

Pour surveiller la progression de l'entraînement, nous avons utilisé des callbacks pour vérifier les critères d'arrêt de l'entraînement. Plus précisément, nous avons arrêté

l'entraînement lorsque nous avons observé une tendance de surapprentissage du modèle. Cela nous a permis d'éviter le surentraînement et de garantir que notre modèle généralise bien sur les nouvelles données.

Nous avons utilisé l'optimiseur Adam avec un taux d'apprentissage de $1e-4$ (0.0001) pour réguler le processus d'optimisation du modèle. La fonction de perte utilisée était la "categorical crossentropy", qui est couramment utilisée dans les tâches de classification multiclasse. Nous avons évalué les performances du modèle en utilisant la métrique d'exactitude ("accuracy" et "validation accuracy") et de perte ("Loss" et "validation Loss") pour mesurer la capacité du modèle à classer correctement les images.

Paramètres	Valeurs
Optimiseur	Adam
Fonction de perte	Categorical Crossentropy
Fonction d'activation (couche d'entrée)	ReLU
Nombre d'epochs	50
Fonction d'activation (couche de sortie)	Softmax

TABLE V.1 – Les paramètres utilisés pour notre approche

V.4.1 Entraînement du modèle VGG16

Sur le dataset public

Lors de l'entraînement de votre modèle VGG16 sur les données publiques, nous avons pu observer une amélioration progressive de ses performances au fil des epochs. Les résultats obtenus sont les suivants :

Epoch 1/20 : Loss : 1.1371, accuracy : 0.4884, val_Loss : 0.9035, val_accuracy : 0.6949

Epoch 2/20 : Loss : 0.7737, accuracy : 0.6857, val_Loss : 0.6252, val_accuracy : 0.7696

Epoch 3/20 : Loss : 0.4829, accuracy : 0.8291, val_Loss : 0.3239, val_accuracy : 0.8726

...

Epoch 18/20 : Loss : 0.0712, accuracy : 0.9804, val_Loss : 0.0513, val_accuracy : 0.9847

Epoch 19/20 : Loss : 0.0385, accuracy : 0.9877, val_Loss : 0.0283, val_accuracy : 0.9908

Ces résultats illustrent l'évolution de la perte (Loss) et de l'accuracy sur l'ensemble d'entraînement et l'ensemble de validation au fur et à mesure des différentes epochs. Le

modèle VGG16 a montré une progression constante de sa performance, avec une augmentation de l'accuracy et une diminution de la perte sur les deux ensembles de données.

Sur notre dataset

L'entraînement de notre modèle VGG16 sur notre dataset a été réalisé avec soin pour obtenir les meilleures performances possibles. Au cours de l'entraînement, nous avons observé que le modèle a atteint sa plus grande valeur d'accuracy et de validation accuracy à l'époch 20, bien avant la fin prévue des 50 epochs. Cependant, au-delà de l'époch 20, nous avons remarqué une tendance à l'Overfitting, où le modèle commençait à surapprendre les données d'entraînement spécifiques et à ne pas généraliser correctement sur de nouvelles données. Pour éviter cela, nous avons décidé d'arrêter l'entraînement à l'époch 20 et de considérer ce modèle comme notre version finale.

Voici les résultats de l'entraînement jusqu'à l'époch 20 :

```
Epoch 1/50 : Loss : 0.7111, accuracy : 0.6074, val_Loss : 0.5841, val_accuracy : 0.6170
Epoch 2/50 : Loss : 0.6631, accuracy : 0.5986, val_Loss : 0.6119, val_accuracy : 0.6170
Epoch 3/50 : Loss : 0.6531, accuracy : 0.6320, val_Loss : 0.5537, val_accuracy : 0.9220
...
Epoch 19/50 : Loss : 0.2472, accuracy : 0.9137, val_Loss : 0.2353, val_accuracy : 0.9078
Epoch 20/50 : Loss : 0.2094, accuracy : 0.9313, val_Loss : 0.3608, val_accuracy : 0.9007
```

V.4.2 Entraînement du modèle InceptionV3

Sur le dataset public

Lors de l'entraînement de votre modèle Inception sur les données publiques, nous avons obtenu des performances exceptionnelles. Le modèle a progressé rapidement dès les premières epochs, atteignant une précision et une validation accuracy élevée. Voici les résultats de l'entraînement jusqu'à l'époch 10 :

```
Epoch 1/10 : Loss : 0.4834, accuracy : 0.8193, val_Loss : 0.2030, val_accuracy : 0.9230
Epoch 2/10 : Loss : 0.1516, accuracy : 0.9473, val_Loss : 0.2340, val_accuracy : 0.9138
Epoch 3/10 : Loss : 0.0842, accuracy : 0.9701, val_Loss : 0.1351, val_accuracy : 0.9489
```

...

Epoch 8/10 : Loss : 0.0331, accuracy : 0.9898, val_Loss : 0.0486, val_accuracy : 0.9870

Epoch 9/10 : Loss : 0.0182, accuracy : 0.9946, val_Loss : 0.0549, val_accuracy : 0.9870

Epoch 10/10 : Loss : 0.0264, accuracy : 0.9911, val_Loss : 0.0505, val_accuracy : 0.9908

Sur notre dataset

Pendant l'entraînement de votre modèle InceptionV3 sur vos données, nous avons observé des progrès significatifs. L'accuracy et la validation accuracy ont atteint leur valeur maximale à l'époch 17, bien avant la fin prévue des 50 epochs. Au-delà de cette étape, nous avons remarqué que le modèle commençait à souffrir d'un Overfitting, ce qui signifie qu'il se spécialise trop dans les données d'entraînement et ne généralise pas de manière optimale sur de nouvelles données. Par conséquent, nous avons décidé d'arrêter l'entraînement à l'époch 17 et de considérer ce modèle comme notre version finale.

Voici les résultats de l'entraînement jusqu'à l'époch 17 :

Epoch 1/50 : Loss : 0.6499, accuracy : 0.7165, val_Loss : 3.1326, val_accuracy : 0.3830

Epoch 2/50 : Loss : 0.3449, accuracy : 0.8820, val_Loss : 1.4053, val_accuracy : 0.5816

Epoch 3/50 : Loss : 0.2286, accuracy : 0.9049, val_Loss : 0.6289, val_accuracy : 0.8085

...

Epoch 16/50 : Loss : 0.1129, accuracy : 0.9472, val_Loss : 0.3336, val_accuracy : 0.9220

Epoch 17/50 : Loss : 0.0988, accuracy : 0.9560, val_Loss : 0.2953, val_accuracy : 0.9220

V.4.3 Entraînement du modèle Densenet121

Sur le dataset public

Lors de l'entraînement de votre modèle Densenet121 sur les données publiques, nous avons observé des résultats prometteurs. L'accuracy et la validation accuracy ont atteint leur valeur maximale à l'époch 8, bien avant la fin prévue des 10 epochs. Au-delà de cette étape, nous avons constaté une tendance à l'Overfitting, indiquant que le modèle se spécialise excessivement dans les données d'entraînement et ne généralisent pas de manière optimale sur de nouvelles données. Par conséquent, nous avons décidé d'arrêter

l'entraînement à l'époch 8, considérant que le modèle avait atteint sa performance maximale.

Voici les résultats de l'entraînement jusqu'à l'époch 8 :

Epoch 1/10 : Loss : 0.7860, accuracy : 0.7530, val_Loss : 0.9225, val_accuracy : 0.7323

Epoch 2/10 : Loss : 0.2779, accuracy : 0.9088, val_Loss : 0.3006, val_accuracy : 0.9069

Epoch 3/10 : Loss : 0.1865, accuracy : 0.9387, val_Loss : 0.2603, val_accuracy : 0.9252

...

Epoch 7/10 : Loss : 0.0783, accuracy : 0.9750, val_Loss : 0.0623, val_accuracy : 0.9847

Epoch 8/10 : Loss : 0.0628, accuracy : 0.9802, val_Loss : 0.0226, val_accuracy : 0.9931

Sur notre dataset

Lors de l'entraînement de votre modèle DenseNet121 sur notre dataset, nous avons observé des performances prometteuses. Le modèle a atteint son plus haut niveau d'accuracy et de validation accuracy à l'époch 19, bien avant la fin prévue des 50 epochs. Au-delà de l'époch 19, nous avons constaté une tendance à l'Overfitting, où le modèle commençait à surapprendre les données d'entraînement et à ne pas généraliser de manière optimale sur de nouvelles données. Afin de prévenir cet effet d'Overfitting, nous avons décidé d'arrêter l'entraînement à l'époch 19 et de considérer ce modèle comme notre version finale.

Voici les résultats de l'entraînement jusqu'à l'époch 19 :

Epoch 1/50 : Loss : 0.8805, accuracy : 0.6989, val_Loss : 0.3135, val_accuracy : 0.8652

Epoch 2/50 : Loss : 0.5812, accuracy : 0.8222, val_Loss : 0.2206, val_accuracy : 0.9220

Epoch 3/50 : Loss : 0.4962, accuracy : 0.8627, val_Loss : 0.2123, val_accuracy : 0.9149

...

Epoch 18/50 : Loss : 0.1546, accuracy : 0.9296, val_Loss : 0.2592, val_accuracy : 0.9362

Epoch 19/50 : Loss : 0.1254, accuracy : 0.9472, val_Loss : 0.3496, val_accuracy : 0.9220

V.5 Résultats

V.5.1 Résultats du modèle VGG16

La figure V.6 présente les résultats de l'évaluation de notre modèle en utilisant VGG16 sur le dataset public. Les métriques d'évaluation sont affichées pour fournir un aperçu de la performance du modèle. Cette visualisation nous a permis d'analyser la relation entre l'entraînement et le test (validation), et ainsi de détecter l'Overfitting. L'exactitude (accuracy) et la perte (Loss) de validation ont atteint des valeurs remarquables de 0.9908 (99%) et 0.02 respectivement à l'epoch 19, indiquant une excellente capacité de généralisation du modèle sur de nouvelles données. On remarque aussi la similarité entre l'accuracy et la val_accuracy. Cela témoigne de l'efficacité du modèle VGG16 sur nos données publiques et de sa capacité à apprendre les motifs et les caractéristiques des images.

De plus, pour une évaluation plus complète, nous avons également testé notre modèle sur notre dataset privé. La figure V.7 présente les résultats de cette évaluation. À l'epoch 20, notre modèle a obtenu une accuracy de 93% et une val accuracy de 90% sur le dataset privé. Ces résultats confirment la performance prometteuse de notre modèle, non seulement sur les données publiques, mais aussi sur les données privées malgré la limitation en terme de quantité, . Cette cohérence entre les performances sur les deux ensembles de données souligne la capacité du modèle à généraliser efficacement.

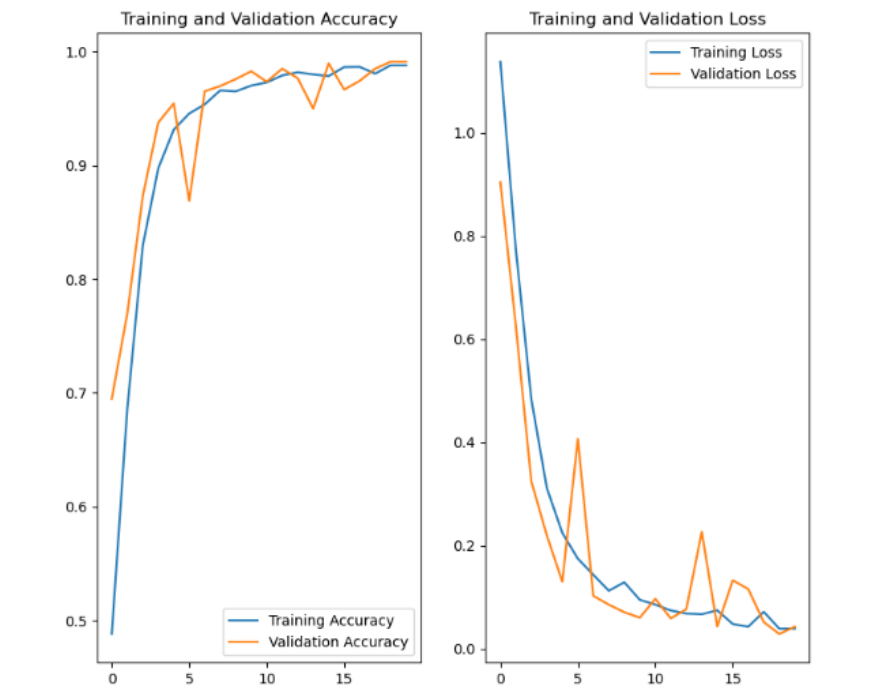


FIGURE V.6 – Visualisation des métriques d'évaluation de notre modèle basé sur VGG16 sur le dataset public

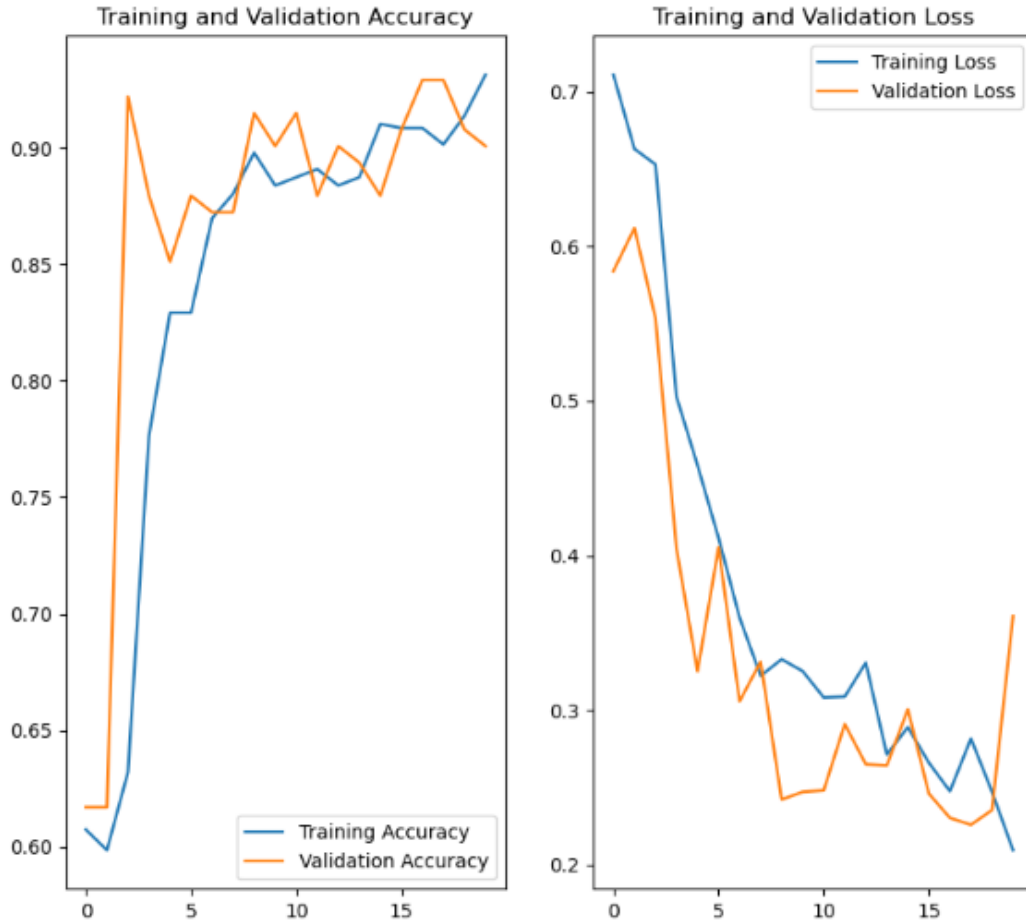


FIGURE V.7 – Visualisation des métriques d'évaluation de notre modèle basé sur VGG16 sur le dataset privé

Le tableau présenté dans la figure V.8 synthétise les différents résultats obtenus par notre modèle VGG16 sur les datasets public et privé. Ce tableau offre une vue d'ensemble des performances du modèle, en regroupant les métriques clés telles que l'accuracy, la val_accuracy, le Loss et le val_Loss pour chaque ensemble de données.

Data	Accuracy	Loss	Val_Accuracy	Val_Loss
Données utilisées	Exactitude	Perte	Exactitude de validation	Perte de validation
Dataset Public <ul style="list-style-type: none"> • Glioma • Minengioma • Notumor • Pituitary 	0.98	0.03	0.99	0.04
Notre Dataset <ul style="list-style-type: none"> • Combiné • Trans-petreuse 	0.93	0.02	0.90	0.36

FIGURE V.8 – Tableau représentant les résultats du modèle VGG16 sur les deux datasets

V.5.2 Résultats du modèle InceptionV3

La figure V.9 présente les résultats de l'évaluation de notre modèle utilisant l'architecture InceptionV3 sur le dataset public. Comme indiqué précédemment, les métriques d'évaluation affichées fournissent un aperçu de la performance du modèle. À l'époch 10, notre modèle a atteint une accuracy de 0.9911 et une val accuracy de 0.9908, démontrant ainsi une excellente capacité de généralisation sur de nouvelles données. Il est intéressant de noter la similarité entre la perte de l'entraînement et la val Loss, témoignant de l'efficacité de l'architecture InceptionV3 pour apprendre les motifs et les caractéristiques des images sur notre dataset public.

Nous avons également évalué notre modèle sur notre dataset privé. Les résultats de cette évaluation sont présentés dans la figure V.10. À l'époch 17, notre modèle a obtenu une accuracy de 0.9560 (95%) et une val accuracy de 0.9220 (90%) sur le dataset privé. Ces résultats confirment la performance prometteuse de notre modèle, tant sur les données publiques que sur les données privées. Cette cohérence entre les performances des deux ensembles de données met en évidence la capacité de généralisation de notre modèle, même lorsqu'il est confronté à des données peu nombreuses.

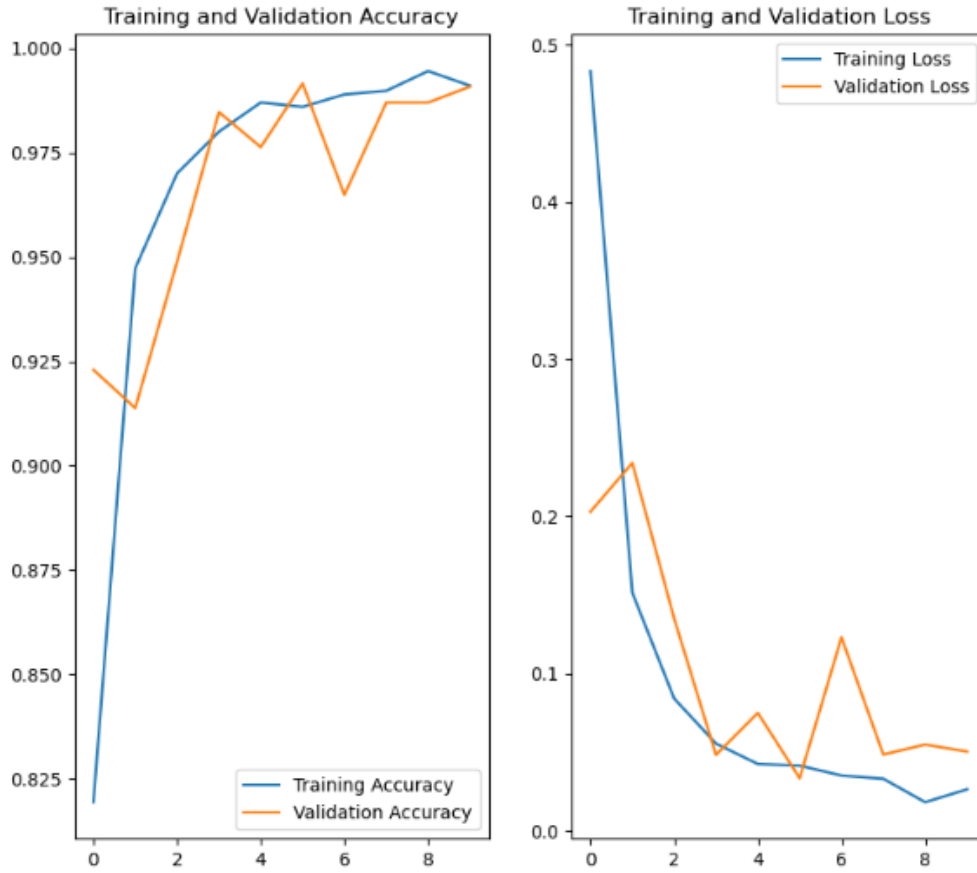


FIGURE V.9 – Visualisation des métriques d'évaluation de notre modèle basé sur InceptionV3 sur le dataset public

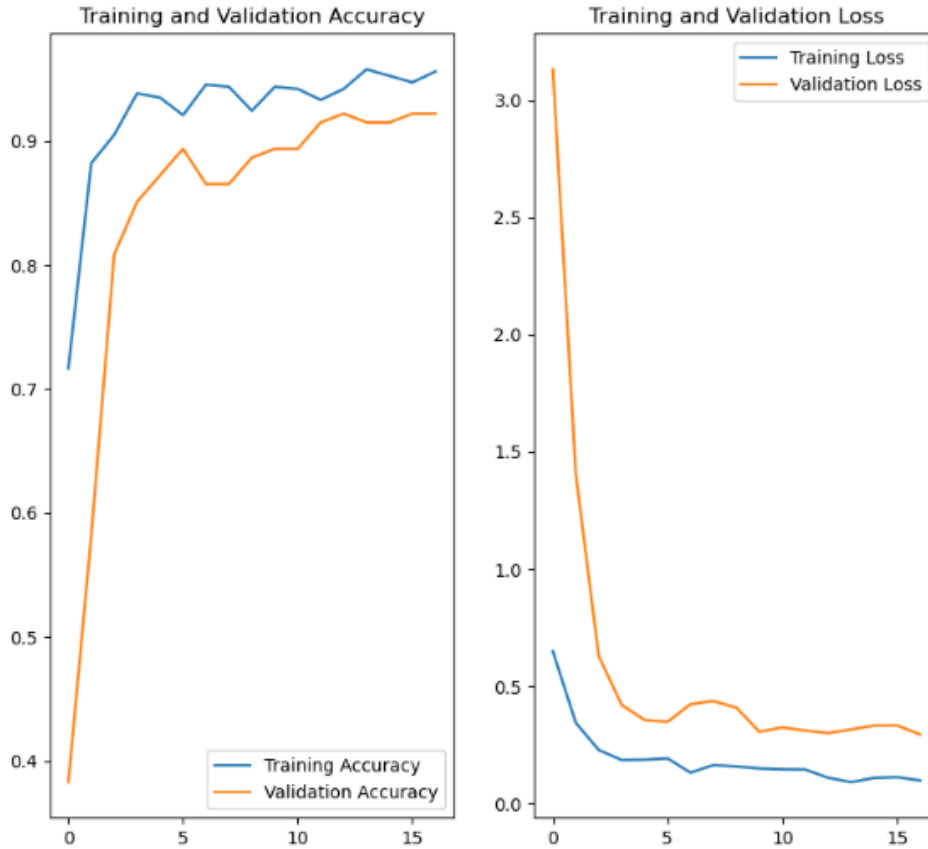


FIGURE V.10 – Visualisation des métriques d'évaluation de notre modèle basé sur InceptionV3 sur le dataset privé

Le tableau présenté dans la figure V.11 résume les différents résultats obtenus par notre modèle InceptionV3 sur les datasets public et privé. Ce tableau offre une vue concise des performances du modèle, en regroupant les métriques clés telles que l'accuracy, la val_accuracy, le Loss et le val_Loss pour chaque ensemble de données.

Data	Accuracy	Loss	Val_Accuracy	Val_Loss
Données utilisées	Exactitude	Perte	Exactitude de validation	Perte de validation
Dataset Public <ul style="list-style-type: none"> • Glioma • Meningioma • Notumor • Pituitary 	0.99	0.02	0.99	0.05
Notre Dataset <ul style="list-style-type: none"> • Combiné • Trans-petreuse 	0.95	0.09	0.92	0.29

FIGURE V.11 – Tableau représentant les résultats du modèle InceptionV3 sur les deux datasets

V.5.3 Résultats du modèle Densenet121

Les résultats de l'évaluation de notre modèle Densenet121 sur le dataset public sont présentés dans la figure V.12. Ces résultats fournissent un aperçu de la performance du modèle, en mettant en évidence son aptitude à apprendre les motifs et les caractéristiques des images. À l'époque 8, notre modèle a atteint une accuracy de 0.9802 (98%) et une val accuracy de 0.9931 (99%). Ces valeurs reflètent une excellente capacité de généralisation du modèle sur de nouvelles données, soulignant ainsi l'efficacité de l'architecture Densenet121.

Nous avons également évalué notre modèle sur notre dataset privé. Les résultats de cette évaluation sont présentés dans la figure V.13. À l'époque 19, notre modèle a obtenu une accuracy de 0.9472 (94%) et une val accuracy de 0.9220 (92%) sur le dataset privé. Ces résultats confirment la performance prometteuse de notre modèle, tant sur les données publiques que sur les données privées, et soulignent sa capacité à généraliser efficacement.

En résumé, ces résultats démontrent la fiabilité et la capacité de généralisation du modèle Densenet121 sur différents datasets d'images IRM.

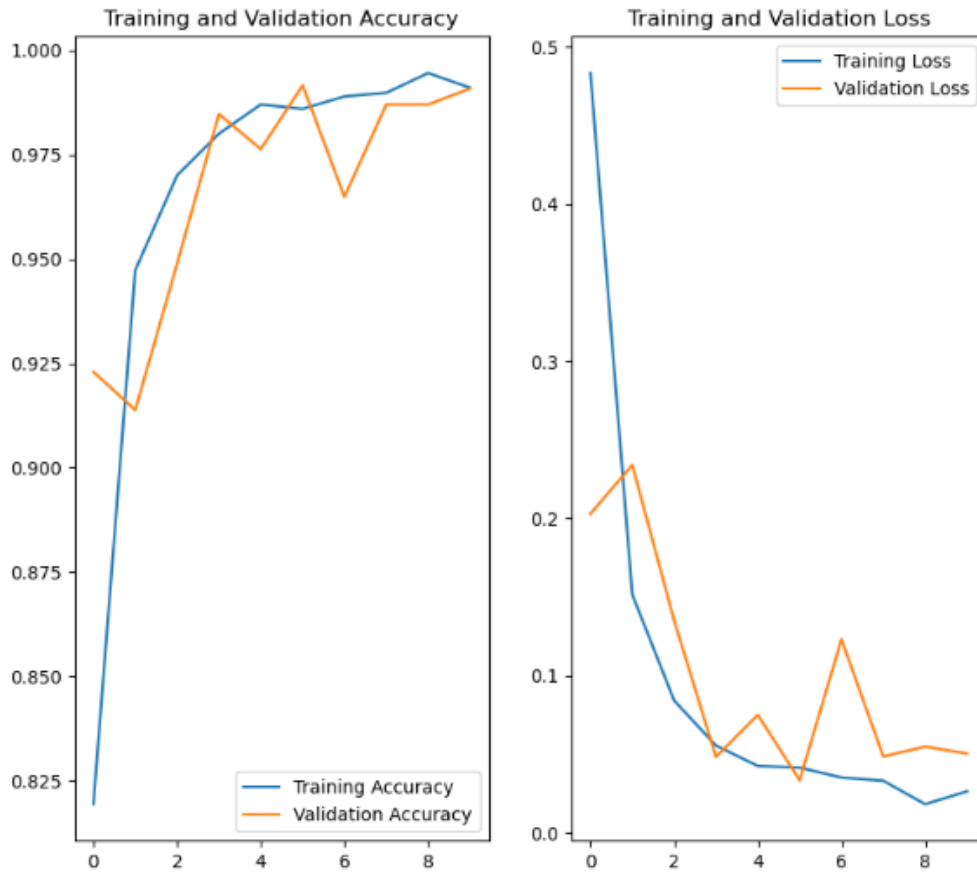


FIGURE V.12 – Visualisation des métriques d'évaluation de notre modèle basé sur Densenet121 sur le dataset public

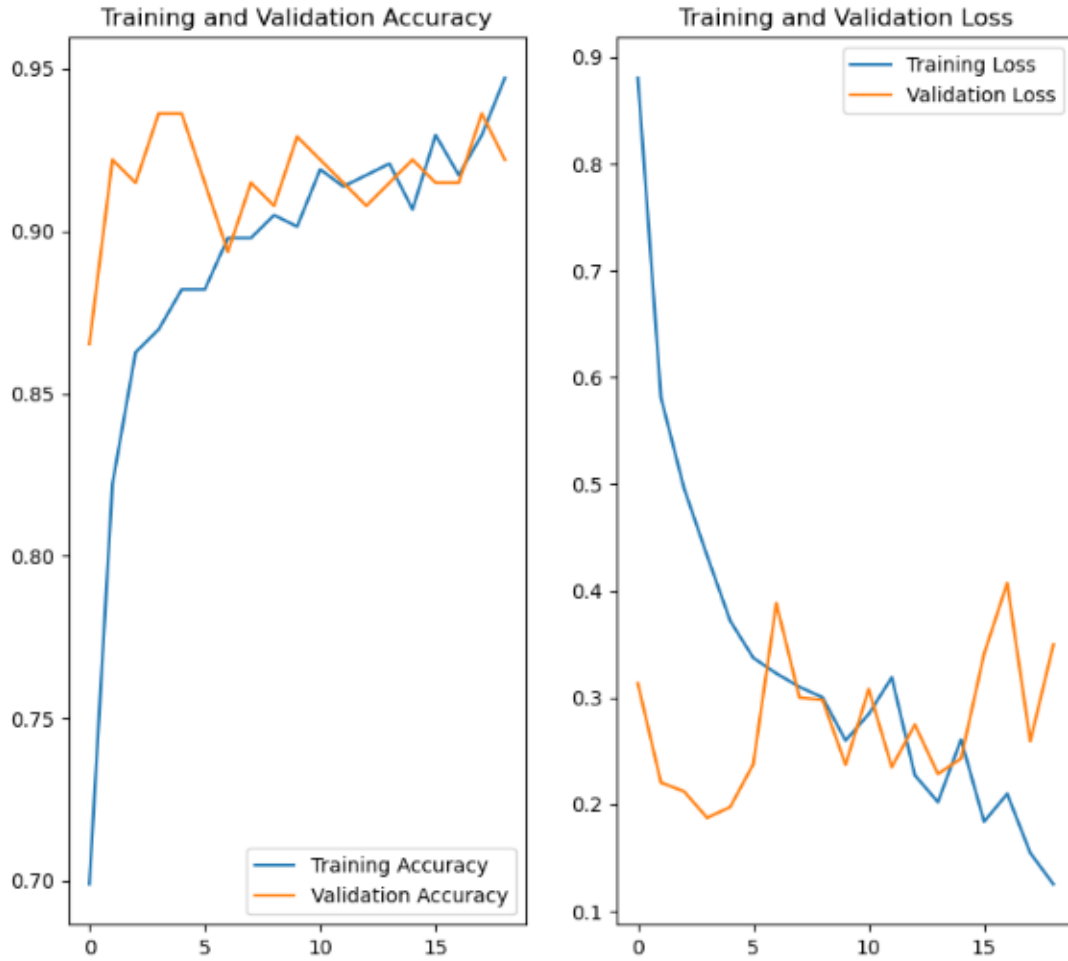


FIGURE V.13 – Visualisation des métriques d’évaluation de notre modèle basé sur DenseNet121 sur le dataset privé

Le tableau présenté dans la figure V.14 synthétise les différents résultats obtenus par notre modèle DenseNet121 sur les datasets public et privé. Ce tableau offre un aperçu global des performances du modèle, en regroupant les métriques clés telles que l’accuracy et la val accuracy pour chaque ensemble de données.

Data	Accuracy	Loss	Val_Accuracy	Val_Loss
Données utilisées	Exactitude	Perte	Exactitude de validation	Perte de validation
Dataset Publique <ul style="list-style-type: none"> • Glioma • Meningioma • Notumor • Pituitary 	0.98	0.06	0.99	0.02
Notre Dataset <ul style="list-style-type: none"> • Combiné • Trans-petreuse 	0.94	0.12	0.92	0.34

FIGURE V.14 – Tableau représentant les résultats du modèle Densenet121 sur les deux datasets

V.6 Discussion des résultats

Les résultats obtenus à partir de nos expérimentations et de nos modèles de classification sont encourageants. En plus de l'utilisation des trois modèles populaires en Transfer Learning : VGG16, Densenet121 et InceptionV3, nous envisageons également effectuer un Ensembling de ces modèles pour améliorer les performances de classification.

Lors de nos tests sur notre propre ensemble de données, composé d'images étiquetées dans les classes "voie combinée" et "voie transpétreuse", nous avons obtenu des résultats satisfaisants. Le modèle VGG16 a montré une Accuracy de 93,13% et une Loss de 0.21, le modèle Densenet121 a atteint une Accuracy de 94,72% et une Loss de 0.12, tandis que le modèle InceptionV3 a obtenu une Accuracy de 95,60% et une Loss de 0.09. Ces résultats indiquent une bonne convergence du modèle et démontrent leur efficacité dans notre tâche de classification.

Parallèlement, Le modèle VGG16 a montré une Val_Accuracy de 90,07% et une Val_Loss de 0.36, le modèle Densenet121 a atteint une Val_Accuracy de 92,20% et une Val_Loss de 0.34, tandis que le modèle InceptionV3 a obtenu une Accuracy de 95,60% et une Loss de 0.29. Ces résultats de validation (test) suggèrent que les modèles sont capables de reconnaître et de classer avec précision des éléments qui n'ont pas été inclus dans l'ensemble d'apprentissage initial. Cette capacité de généralisation est essentielle pour garantir la fiabilité et l'efficacité des modèles dans des scénarios réels où de nouvelles données sont constamment introduites.

Il est important de souligner que nos résultats sont d'une taille restreinte, car ils ont été obtenus à partir d'un dataset soigneusement étiqueté une image par une par un expert neurochirurien. Cependant, pour continuer à améliorer nos résultats, il serait bénéfique d'augmenter la taille de notre ensemble de données, comme le démontrent les résultats sur le dataset public. Les performances exceptionnelles de nos modèles sur ce dataset témoignent de leur capacité à reconnaître et à classer avec précision les images IRM. Ainsi, l'ajout de nouveaux cas à notre dataset privé permettrait une meilleure généralisation des modèles et la capture d'une plus grande variabilité des caractéristiques présentes dans ces images, et ainsi une amélioration de la précision des modèles.

Maintenant, examinons notre contribution sous la forme de couches supérieures ajoutées à la fin des modèles. La couche de pooling globale a permis de réduire la dimensionnalité des caractéristiques tout en conservant les informations clés. La couche de Dropout a contribué à la régularisation, réduisant le surapprentissage et améliorant la généralisation. Enfin, la couche dense avec activation softmax a facilité la classification des données en produisant des probabilités pour chaque classe. Ces ajouts ont renforcé la capacité de nos modèles à extraire des caractéristiques discriminantes à partir des images IRM en voie d'abord, améliorant ainsi leur performance et leur adaptabilité à de nouvelles données.

En outre, il serait intéressant d'explorer d'autres architectures de modèles de Transfer Learning pour cette tâche spécifique. Il existe de nombreuses variantes et architectures avancées qui pourraient potentiellement améliorer les performances de classification. L'utilisation de techniques d'optimisation plus avancées, telles que le fine-tuning des couches plus profondes des modèles, pourrait également être explorée pour améliorer les résultats.

En conclusion, nos résultats démontrent l'efficacité du Transfer Learning dans la classification d'images IRM en voie d'abord. Les modèles de Transfer Learning adaptés ont montré de bonnes performances de classification. L'Ensembling de ces modèles peut permettre d'améliorer les résultats. Cependant, des travaux futurs sont nécessaires pour explorer davantage les possibilités d'amélioration, notamment en augmentant la taille de l'ensemble de données et en explorant d'autres architectures et techniques d'optimisation.

V.7 Application : Implémentation et Intégration des Modèles

Introduction

Nous avons élaboré une application web spécifiquement dédiée à l'implémentation et à l'intégration de nos modèles de prédiction. Cette application a été développée en utilisant le framework Django de Python. Grâce à l'utilisation de Django, HTML5, CSS3 et JavaScript, nous avons pu créer une interface utilisateur conviviale et intuitive qui facilite l'interaction des utilisateurs avec notre modèle de prédiction. En tirant parti de Django en tant que framework Python, nous avons pu utiliser ses bibliothèques appropriées pour mettre en œuvre nos modèles, notamment Tensorflow et Numpy. L'utilisation de Django a également permis de garantir la robustesse, la flexibilité et la sécurité sur notre application web.

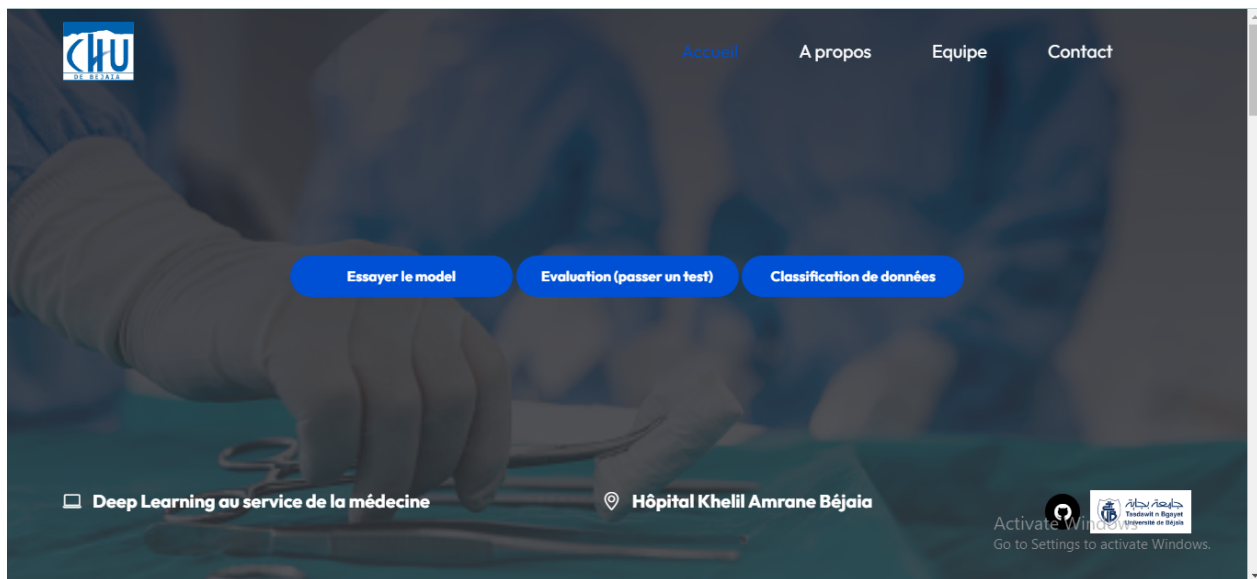


FIGURE V.15 – Interface d'accueil de l'application

En approfondissant, notre application vise à offrir une plateforme permettant aux utilisateurs de tester leurs compétences en matière de classification des voies d'abord, et a permettre aux experts de classer facilement les images IRM afin de les utiliser pour l'entraînement de nos modèles. En résumé, les deux objectifs sur notre application sont les suivants :

- faciliter la classification des images IRM.
- offrir un outil de test et d'évaluation pour les médecins et les résidents.

Partie de données

La première partie sur notre application est dédiée à la classification de données. L'application nous permet de stocker et d'afficher des images IRM dans notre système. Ces images sont ensuite soumises au Pr HIMEUR pour une classification rapide et précise de la voie d'abord appropriée. En un simple clic, le Pr peut fournir une expertise en identifiant la voie d'abord la plus adaptée pour chaque cas. La base de données resultante est ensuite utilisée pour entrainer nos modeles de prediction.

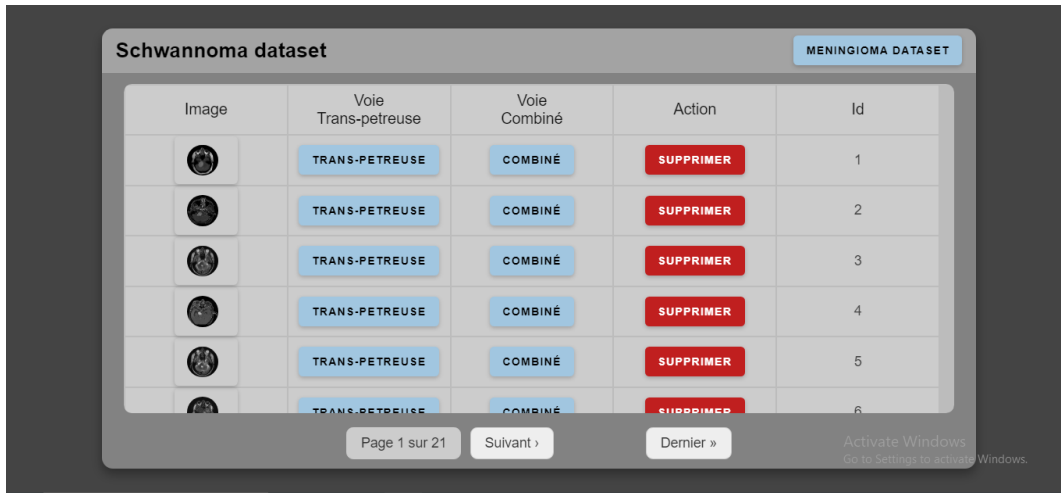


FIGURE V.16 – Interface de classification des images IRM en voie d'abord sur notre application

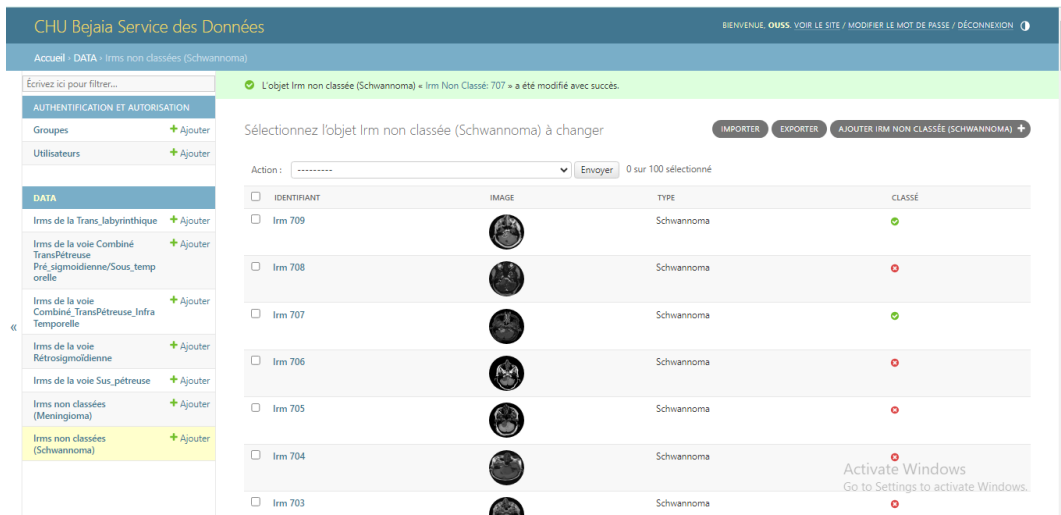


FIGURE V.17 – Interface d'administrateur pour la gestion des images IRM sur notre application

Partie des modèles

Cette partie offre aux médecins et aux résidents une opportunité de gagner de l'expérience en s'exerçant sur des cas réels. En utilisant les images IRM disponibles sur la plateforme, ou en chargeant leur propre image, les utilisateurs ont la possibilité de les classer en voie combinée ou voie transpétreuse. Le système compare en suite les réponses avec les prédictions générées par notre modèle CNN pour finalement produire une évaluation. Cela offre une opportunité précieuse pour les médecins et les résidents d'améliorer leur expertise grâce à une rétroaction immédiate basée sur les performances des modèles.

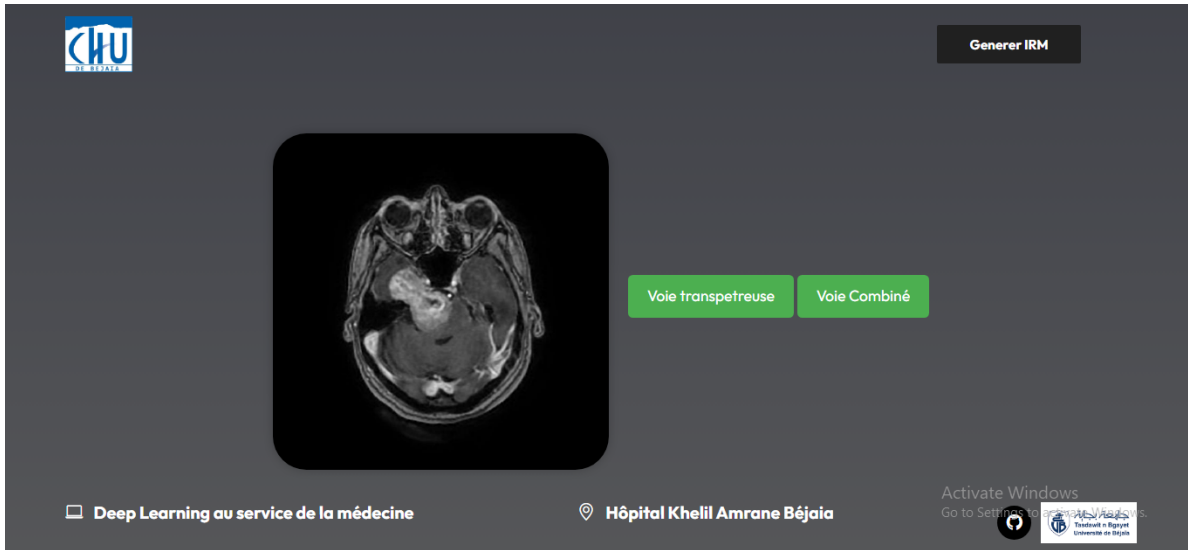


FIGURE V.18 – Interface de test sur notre application

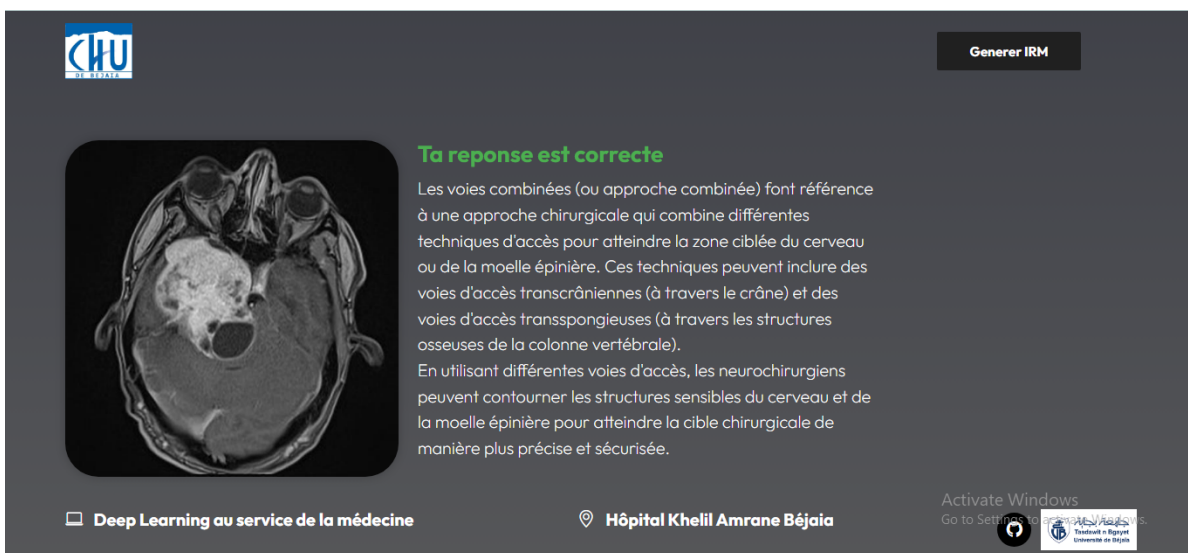


FIGURE V.19 – Exemple d'une reponse correcte sur l'interface de test sur notre application



FIGURE V.20 – Exemple d’une reponse incorrecte sur l’interface de test sur notre application

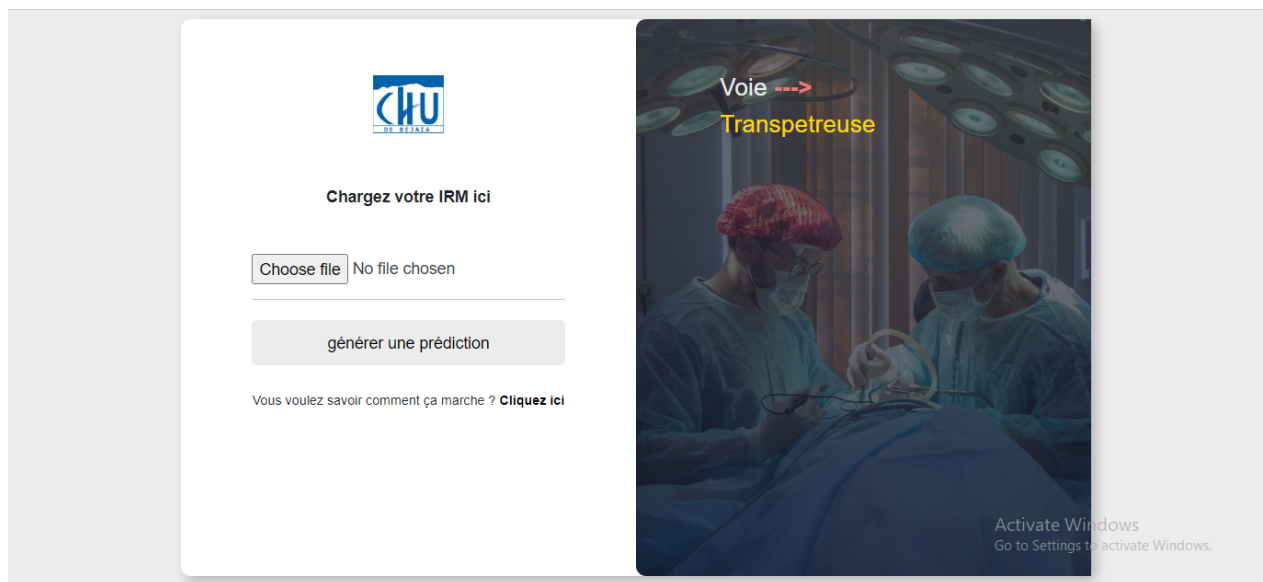


FIGURE V.21 – Interface pour charger sa propre image IRM sur notre application

Conclusion générale

En conclusion, notre étude s'est concentrée sur l'utilisation de techniques d'apprentissage profond, notamment des réseaux de neurones convolutifs (CNN), pour la classification des tumeurs cérébrales à partir d'images d'IRM. Les résultats obtenus à partir des articles examinés ont démontré une précision élevée dans la classification des tumeurs cérébrales, ce qui suggère que l'apprentissage profond est une approche efficace pour le diagnostic et la classification de ces affections médicales.

Les méthodes proposées dans ces articles ont souvent combiné des techniques telles que le Transfer Learning, l'augmentation des données et l'Ensembling. Le Transfer Learning a permis d'extraire automatiquement des caractéristiques pertinentes à partir des images, sans nécessiter une extraction manuelle. L'augmentation des données a contribué à améliorer la capacité du modèle à généraliser et à classer avec précision les tumeurs cérébrales. De plus, l'Ensembling, c'est-à-dire la combinaison de plusieurs modèles, a été considéré pour bénéficier de la précision des différents modèles.

Sur la base de ces éléments, nous avons décidé d'utiliser trois modèles populaires en Transfer Learning, à savoir VGG16, DenseNet121 et InceptionV3, pour extraire les caractéristiques des images d'IRM et les utiliser dans notre processus de classification. Nos expérimentations sur notre propre ensemble de données ont montré des résultats encourageants, avec des taux d'exactitude élevés et des pertes faibles pour chaque modèle à savoir 93% pour VGG16, 95% pour InceptionV3 et 94% pour Densenet121 d'accuracy.

Cependant, il est important de noter que nos résultats ont été obtenus à partir d'un ensemble de données de petite taille, il serait bénéfique d'augmenter la taille de notre ensemble de données pour capturer une plus grande variabilité des caractéristiques présentes dans les images d'IRM et de régulariser le nombre de données entre les classes.

En conclusion, nos résultats démontrent l'efficacité du Transfer Learning dans la classification des images d'IRM pour la détection et la classification des tumeurs cérébrales. L'utilisation de modèles de Transfer Learning adaptés, combinée à des techniques d'Ensembling, peut améliorer les performances de classification. Cependant, des travaux futurs sont nécessaires pour explorer davantage les possibilités d'amélioration, notamment

en augmentant la taille de l'ensemble de données, en explorant d'autres architectures de modèles et en utilisant des techniques d'optimisation avancées.

Bibliographie

- [1] A beginner's guide to stacking ensemble deep learning models. Article sur Analytics India Magazine. URL <https://analyticsindiamag.com/a-beginners-guide-to-stacking-ensemble-deep-learning-models/>. Consulté le 10/07/2023.
- [2] Artificial intelligence. Site web, . URL <https://www.ibm.com/topics/artificial-intelligence>. Consulté le 10/07/2023.
- [3] Machine learning. Site web, . URL <https://www.ibm.com/topics/machine-learning>. Consulté le 27/06/2023.
- [4] Neural networks. Site web, . URL <https://www.ibm.com/topics/neural-networks>. Consulté le 27/06/2023.
- [5] Partie pétreuse de l'os temporal. Site web. URL <https://www.imaios.com/fr/e-anatomy/structures-anatomiques/partie-petreuse-de-l-os-temporal-1536896096>. Consulté le 28/06/2023.
- [6] What is the purpose of image preprocessing in deep learning? Site web. URL <https://fr.isahit.com/blog/what-is-the-purpose-of-image-preprocessing-in-deep-learning>. Consulté le [date de consultation].
- [7] Réseau de neurones artificiels : définition et fonctionnement. Article sur LeBigData. URL <https://www.lebigdata.fr/reseau-de-neurones-artificiels-definition>. Consulté le [date de consultation].
- [8] The ultimate beginner guide of semi-supervised learning. Article sur Medium, . URL <https://medium.datadriveninvestor.com/the-ultimate-beginner-guide-of-semi-supervised-learning-3bd11cb19835>. Consulté le [date de consultation].
- [9] An overview of vgg16 and nin models. Article sur Medium, . URL <https://medium.com/mllearning-ai/an-overview-of-vgg16-and-nin-models-96e4bf398484>. Consulté le 22/07/2023].

- [10] Schwannoma. Site web, . URL <https://www.cancer.gov/pediatric-adult-rare-tumor/rare-tumors/rare-soft-tissue-tumors/schwannoma>. Consulté le 28/06/2023.
- [11] Meningioma. Site web, . URL <https://www.cancer.gov/rare-brain-spine-tumor/tumors/meningioma>. Consulté le 29/05/2023.
- [12] Base du crâne. Site web. URL <https://www.neurochirurgie-lariboisiere.com/base-du-crane/>. Consulté le 28/06/2023.
- [13] Fiche : Os temporal. Site web. URL <https://www.passeportsante.net/fr/parties-corps/Fiche.aspx?doc=os-temporal>. Consulté le [date de consultation].
- [14] A sustainable deep learning framework for object recognition using multi-layers deep features fusion and selection. URL https://www.researchgate.net/figure/Modified-Inceptionv3-architecture-for-features-extraction_fig3_342242847. Consulté le 22 Jun, 2023.
- [15] Data. Site web, . URL <https://www.techtarget.com/searchdatamanagement/definition/data>. Consulté le 28/06/2023.
- [16] Reinforcement learning tutorial. Tutoriel sur TechVidvan, . URL <https://techvidvan.com/tutorials/reinforcement-learning/>. Consulté le [date de consultation].
- [17] Data augmentation. Site web. URL https://www.tensorflow.org/tutorials/images/data_augmentation. Consulté le [date de consultation].
- [18] Dropout in neural networks. Site web, . URL <https://towardsdatascience.com/dropout-in-neural-networks-47a162d621d9>. Consulté le [date de consultation].
- [19] Learning process of a deep neural network. Site web, . URL <https://towardsdatascience.com/learning-process-of-a-deep-neural-network-5a9768d7a651>. Consulté le [date de consultation].
- [20] Basic cnn architecture : Explained. Article sur le blog d'UpGrad. URL <https://www.upgrad.com/blog/basic-cnn-architecture/>. Consulté le 22/07/2023.
- [21] Petrous apex lesions. Site web. URL <https://uvahealth.com/services/benign-brain-tumor/petrous-apex-lesions>. Consulté le 02/07/2023.
- [22] Overfitting and underfitting concepts : Interview questions. Article sur Vitalflux. URL <https://vitalflux.com/overfitting-underfitting-concepts-interview-questions/>. Consulté le 22/07/2023.

- [23] Y. Bengio, A. Courville, and P. Vincent. Representation learning : A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8) : 1798–1828, 2013.
- [24] R. W. Brown, Y. C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan. *Magnetic resonance imaging : physical principles and sequence design*. John Wiley & Sons, 2014.
- [25] S. Das, O. F. M. R. R. Aranya, and N. N. Labiba. Brain tumor classification using convolutional neural network. In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, pages 1–5, 2019. doi : 10.1109/ICASERT.2019.8934603.
- [26] J. Davis and M. Goadrich. The relationship between precision-recall and roc curves. In *Proceedings of the 23rd International Conference on Machine Learning (ICML)*, 2006.
- [27] S. Deepak and P. Ameer. Brain tumor classification using deep cnn features via transfer learning. *Computers in Biology and Medicine*, 111 :103345, 2019. ISSN 0010-4825. doi : <https://doi.org/10.1016/j.combiomed.2019.103345>. URL <https://www.sciencedirect.com/science/article/pii/S0010482519302148>.
- [28] T. G. Dietterich. Ensemble methods in ml. In *Multiple Classifier Systems, First International Workshop, MCS 2000, Cagliari, Italy, June 21-23, 2000. Proceedings*, pages 1–15. Non spécifié, 2000.
- [29] F. Feltrin. Brain tumor mri images (17 classes). <https://www.kaggle.com/datasets/fernando2rad/brain-tumor-mri-images-17-classes>, 2023. Kaggle.
- [30] G. H. Forman and M. Scholz. Apples-to-apples in cross-validation studies : Pitfalls in classifier performance measurement. *ACM SIGKDD Explorations Newsletter*, 12(1) : 49–57, 2010.
- [31] V. Gilard. *Apports de la Médecine de Précision à la Neurochirurgie : Applications aux Glioblastomes*. Thèse de doctorat, Normandie Université, 2021. fftel03517918f.
- [32] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
- [33] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [34] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
- [35] F. A. L. Ibañez, S. Hem, P. Ajler, E. Vecchi, C. Ciraolo, M. Baccanelli, ..., and A. Carrizo. A new classification of complications in neurosurgery. *World neurosurgery*, 75(5-6) :709–715, 2011.

- [36] W. H. W. Ishak and F. Siraj. Artificial intelligence in medical application : An exploration. *Health Informatics Europe Journal*, 16, 2002.
- [37] D. Kondziolka, S. M. Shin, A. Brunswick, I. Kim, and J. S. Silverman. The biology of radiosurgery and its clinical applications for brain tumors. *Neuro-oncology*, 17(1) : 29–44, 2015.
- [38] S. Krishnapriya and Y. Karuna. Pre-trained deep learning models for brain mri image classification. *Frontiers in Human Neuroscience*, 17 :1150120, 2023. doi : 10.3389/fnhum.2023.1150120. URL <https://doi.org/10.3389/fnhum.2023.1150120>. ORIGINAL RESEARCH article, Sec. Brain-Computer Interfaces.
- [39] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, 2012.
- [40] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324, November 1998.
- [41] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553) :436–444, 2015.
- [42] K. L. Li, V. Agarwal, H. S. Moskowitz, and W. M. Abuzeid. Surgical approaches to the petrous apex. *World Journal of Otorhinolaryngology-Head and Neck Surgery*, 6(02) : 106–114, 2020.
- [43] X. Li, X. Shen, Y. Zhou, X. Wang, and T.-Q. Li. The densenet-121 architecture. *PLOS ONE*, 2020. doi : 10.1371/journal.pone.0232127.g001. URL <https://doi.org/10.1371/journal.pone.0232127.g001>. Figure.
- [44] X.-y. Liu, J. Wu, and Z.-H. Zhou. Exploratory undersampling for class-imbalance learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(2) :539–550, 2008.
- [45] P. B. McBeth, D. F. Louw, P. R. Rizun, and G. R. Sutherland. Robotics in neurosurgery. *The American Journal of Surgery*, 188(4) :68–75, 2004.
- [46] A. A. Mukhlif, B. Al-Khateeb, and M. A. Mohammed. Incorporating a novel dual transfer learning approach for medical images. *Sensors*, 23 :570, 2023. doi : 10.3390/s23020570. URL <https://doi.org/10.3390/s23020570>. Received : 24 November 2022 / Revised : 27 December 2022 / Accepted : 1 January 2023 / Published : 4 January 2023.
- [47] M. Nickparvar. Brain tumor mri dataset. <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>, 2021. Kaggle.
- [48] M. Nickparvar. Brain tumor mri dataset. <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>, 2023.

- [49] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, and M. O. Alassafi. Brain tumor classification based on fine-tuned models and the ensemble method. *Computers, Materials Continua*, 2021. doi : 10.32604/cmc.2021.014158. Article.
- [50] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10) :1345–1359, 2010.
- [51] R. Saini, P. Semwal, and T. H. Jaware. Brain tumor classification using vgg-16 and mobilenetv2 deep learning techniques on magnetic resonance images (mri). *Conference Proceedings*, March 2023. URL <https://link.springer.com/conference/proceedings>. Conference paper.
- [52] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, and S. W. Baik. Multi-grade brain tumor classification using deep cnn with extensive data augmentation. *Journal of Computational Science*, 30 :174–182, 2019. ISSN 1877-7503. doi : <https://doi.org/10.1016/j.jocs.2018.12.003>. URL <https://www.sciencedirect.com/science/article/pii/S1877750318307385>.
- [53] I. H. Sarker. Machine learning : Algorithms, real-world applications and research directions. *SN Computer Science*, 2(160), 2021. doi : 10.1007/s42979-021-00592-x. URL <https://doi.org/10.1007/s42979-021-00592-x>.
- [54] Z. A. Sejuti and M. S. Islam. An efficient method to classify brain tumor using cnn and svm. In *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*, pages 644–648, 2021. doi : 10.1109/ICREST51555.2021.9331060.
- [55] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv :1409.1556*, 2014.
- [56] S. R. Sowrirajan, S. Balasubramanian, and R. S. P. Raj. Mri brain tumor classification using a hybrid vgg16-nade model. *Brazilian Archives of Biology and Technology*, 66, 2023. doi : 10.1590/1678-4324-2023220071. URL <https://doi.org/10.1590/1678-4324-2023220071>. Article - Engineering, Technology and Techniques.
- [57] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015.
- [58] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015.
- [59] D. Y. Wen and R. C. Heros. Surgical approaches to the brain stem. *Neurosurgery Clinics of North America*, 4(3) :457–468, 1993.
- [60] M. Yan, K. Liu, Z. Guan, X. Xinkai, X. Qian, and H. Bao. Background augmentation generative adversarial networks (bagans) : Effective data generation based on gan-augmented 3d synthesizing. *Symmetry*, 10 :734, 12 2018. doi : 10.3390/sym10120734.

- [61] Y. Zhou, Z. Li, H. Zhu, C. Chen, M. Gao, K. Xu, and J. Xu. Holistic brain tumor screening and classification based on densenet and recurrent neural network. *Lecture Notes in Computer Science*, pages 208–217, 2019. doi : 10.1007/978-3-030-11723-8_21. URL https://doi.org/10.1007/978-3-030-11723-8_21.