



جامعة ابن طفيل
+⓪ΛΠΣ+ ΣΘΙ ΕΞΗΘΗ
Ibn Tofaïl University
Faculté des Sciences

Université Ibn Tofaïl
Faculté des Sciences, Kénitra

Mémoire de Projet de Fin d'Etudes

Master Intelligence Artificielle et Réalité Virtuelle

Deep Learning et la prédiction de la maladie d'Alzheimer par analyse des images TEP Scan

Établissement d'accueil : L a b o r a t o i r e S E T I M E F S K

Elaboré par : Mr. Soidad SOULE AHAMADA

Encadré par : Pr. TOUAHNI Raja
Pr. EDDAROUICH Souad

Soutenu le 19 SEPTEMBRE 2024, devant le jury composé de :

- Pr TOUAHNI Raja (FSK Université Ibn Tofaïl)
- Pr NOURI Anas (FSK Université Ibn Tofaïl)
- Pr EDDAROUICH Souad (CRMEF Rabat)
- Pr DABOUNOU Jaouad (FST Settat Université Hassan 2)

Dédicace

À

" Mon petit frère atteint de trisomie 21 et à tout ceux qui se battent pour lui offrir un meilleur avenir. "

Remerciement

Tout d'abord, je tiens à exprimer ma gratitude envers Allah le Tout-Puissant, pour m'avoir accordé la santé et la volonté nécessaires pour entreprendre et achever ce mémoire.

Je souhaite également adresser mes plus sincères remerciements au Professeur EDDAROUICH SOUAD. Sa patience, ses conseils avisés et sa disponibilité ont été d'une aide précieuse tout au long de ce travail. Je suis particulièrement reconnaissant pour la confiance qu'elle m'a accordée en acceptant de m'encadrer et de m'accompagner durant cette période.

Je tiens à remercier chaleureusement les membres du jury pour leur disponibilité et leur bienveillance en acceptant d'examiner ce mémoire. Je suis également reconnaissant envers l'ensemble de l'équipe du master Intelligence Artificielle & Réalité Virtuelle (IARV) de la faculté des Sciences de l'Université Ibn Tofaïl. Leur enseignement et leur soutien au cours de ces deux années de formation ont été essentiels pour la réalisation de ce travail.

Un grand merci à ma famille, à mes proches et à mes amis pour leur soutien inconditionnel et leurs encouragements constants. Enfin, je tiens à remercier toutes les personnes qui, de près ou de loin, ont contribué à la réalisation de ce mémoire.

Résumé

Ce mémoire présente une approche innovante pour la prédiction de la maladie d'Alzheimer à partir des données de tomographie par émission de positons (TEP scans). La maladie d'Alzheimer constitue un défi croissant pour les systèmes de santé à l'échelle mondiale, nécessitant des outils de diagnostic précoce et des interventions ciblées. Notre projet utilise des techniques avancées d'apprentissage automatique pour développer un modèle de prédiction précis et fiable, en se basant exclusivement sur l'analyse des TEP scans.

En exploitant les informations détaillées sur l'activité métabolique du cerveau fournies par les TEP scans, notre objectif est d'identifier les modèles et les caractéristiques spécifiques associés à la maladie, ouvrant ainsi la voie à une intervention précoce et à des traitements plus efficaces. Par ailleurs, nous intégrons l'intelligence artificielle explicable (XAI) dans notre approche pour garantir que les décisions du modèle soient interprétables et transparentes, ce qui est crucial pour la confiance et l'acceptation des professionnels de santé.

Ce mémoire fournit un aperçu de notre démarche, de la problématique, des objectifs et de la méthodologie utilisée pour la prédiction de la maladie d'Alzheimer à partir des données de TEP scans.

Mots clés : Traitement d'images, Machine Learning, Deep Learning, Classifications, IA Explicable

Introduction

La maladie d'Alzheimer représente un défi majeur pour les systèmes de santé à l'échelle mondiale, avec une augmentation alarmante des cas de démence liée à cette maladie neurodégénérative. Cette affection débilitante impacte non seulement la qualité de vie des patients mais aussi leurs proches et les structures de santé qui les prennent en charge.

Les symptômes progressifs de la maladie d'Alzheimer, tels que la perte de mémoire, la confusion et les changements de comportement, rendent le diagnostic précoce et la gestion efficace particulièrement complexes. Face à cette préoccupation croissante, les avancées technologiques dans les domaines de l'apprentissage automatique et de l'imagerie médicale offrent de nouvelles perspectives prometteuses pour le diagnostic précoce et la prise en charge de la maladie d'Alzheimer.

Ce mémoire se concentre sur une approche novatrice visant à prédire la maladie d'Alzheimer à partir de données de tomographie par émission de positons (TEP scans). Les techniques d'imagerie médicale, telles que les TEP scans, permettent d'observer les changements métaboliques dans le cerveau, fournissant ainsi des informations précieuses pour la détection précoce des anomalies associées à la maladie.

Notre projet vise à exploiter des techniques avancées d'apprentissage automatique pour développer un modèle de prédiction précis et fiable de la maladie d'Alzheimer, en se basant exclusivement sur l'analyse des données de TEP scans. En utilisant des algorithmes sophistiqués de DeepLearnig, nous cherchons à identifier les modèles et les caractéristiques spécifiques qui permettent de diagnostiquer la maladie à un stade précoce, ouvrant ainsi la voie à une intervention plus efficace et à des traitements ciblés.

De plus, nous intégrons l'intelligence artificielle explicable (XAI) dans notre approche pour garantir que les décisions du modèle soient interprétables et transparentes, ce qui est essentiel pour la confiance et l'adoption de ces technologies par les professionnels de santé.

Cette introduction fournit un aperçu du cadre général de notre projet, mettant en lumière l'importance et la pertinence de notre approche dans le contexte actuel de la santé publique. Par la suite, nous explorerons en détail la problématique, les objectifs et la méthodologie utilisée, offrant ainsi une compréhension approfondie de notre démarche pour la prédiction de la maladie d'Alzheimer à partir des données de TEP scans.

Table des matières

Dédicace	2
Remerciement	3
Résumé	4
Introduction	5
Table des matières	6
Table des figures	9
1 Traitement des images Médicales par l'IA	11
1.1 Introduction	11
1.2 Les images TEP scan dans l'étude de la maladie d'Alzheimer	12
1.2.1 Principe de TEP scan	13
1.2.2 TEP Scan et Radiotraceurs	13
1.3 Importance de la détection précoce et de la classification précise	14
1.4 Traitement des images médicales	15
1.4.1 Méthodes classiques de traitement d'images médicales	15
1.4.2 Limitations des approches classiques	15
1.5 Apprentissage Automatique appliqué à la classification de la maladie d'Alzheimer	16
1.5.1 Vue d'ensemble des techniques de Machine Learning	16
1.5.2 Exemples des méthodes classiques de Machine Learning	16
1.5.3 Limitations des méthodes classiques de Machine Learning	17
1.6 Deep Learning et ses progrès récents	18

1.6.1	Deep Learning et ses applications en vision par ordinateur	18
1.6.2	Succès récents dans l'utilisation du Deep Learning pour la classification de la maladie d'Alzheimer	18
1.6.3	Le Deep Learning et les changements métaboliques difficiles à évaluer au stade précoce	18
1.6.4	Exemples d'architectures des réseaux neuronaux profonds	19
1.6.4.1	Les réseaux de neurones convolutionnels (CNN)	19
1.6.4.2	Les autoencodeurs	24
1.7	IA Explicable	27
1.7.1	Qu'est-ce que l'IA Explicable ?	27
1.7.2	Pourquoi l'IA Explicable ?	27
1.7.3	Les avantages de l'IA Explicable	28
1.7.4	Cas d'utilisation pour l'IA Explicable	29
1.7.5	Quelques méthodes de l'IA Explicable	32
1.8	Conclusion	34
2	Approche de la classification des images TEP de la maladie d'Alzheimer	36
2.1	Introduction	36
2.2	Base de Données	36
2.2.1	Description de la Base de Données	37
2.2.2	Format des Données	38
2.2.3	Les Classes de la Base de Données	38
2.2.4	Source et pertinence de la Base de Données	38
2.3	Collection des Données	39
2.3.1	Méthodologie	39
2.3.2	Vue d'Ensemble des Données Collectées	41
2.4	Prétraitement	42
2.4.1	Conversion en niveau de gris des images TEP scans	42
2.4.2	Normalisation des images TEP scans	43
2.4.3	Débruitage des images TEP scans avec un Autoencodeur	43
2.5	Construction du modèle CNN pour la classification	43
2.5.1	Architecture du Modèle	43

2.5.2	Compilation et Entraînement du modèle CNN pour la classification	46
2.6	Explicabilité du Modèle avec LIME : IA Explicable	48
2.7	Conclusion	49
3	Résultats et Discussion	50
3.1	Introduction	50
3.2	Résultats	50
3.2.1	Approche 1 : Débruitage des images suivi de la classification	50
3.2.1.1	Description de l'Approche et Résultats	50
3.2.1.2	Test	62
3.2.1.3	Discussion	63
3.2.2	Approche 2 : Classification Directe des images sans Débruitage	64
3.2.2.1	Description de l'Approche et Résultats	64
3.2.2.2	Test	69
3.2.2.3	Discussion	70
3.2.3	Explicabilité du Modèle	71
3.2.3.1	Utilisation de LIME pour l'Explicabilité	71
3.2.3.2	Discussion	74
3.3	Comparaison avec d'autres résultats	74
3.3.1	Études comparatives	74
3.3.2	Discussion	75
3.4	Conclusion	76
	Conclusion Générale et Perspectives	77
	Bibliographie	79
	Annexe	83

Table des figures

1.1	Appareil TEP Scans	12
1.2	Résultats TEP Scans	12
1.3	Images TEP scans	13
1.4	Architecture de CNN	20
1.5	Autoencodeur	24
1.6	IA Explicable	28
2.1	Base de Données	37
2.2	Images collectées	41
3.1	Repartition des Images Sélectionnées	51
3.2	Images Non Bruitées	52
3.3	Images Bruitées	52
3.4	Architecture de l'autoencodeur	53
3.5	Progression de l'erreur quadratique moyenne	54
3.6	L'ensemble des images TEP	55
3.7	Images TEP avant débruitage	55
3.8	Images Débruitées	56
3.9	Architecture du CNN	57
3.10	Courbe de Perte de l'approche 1	58
3.11	Courbe de Précision de l'approche 1	59
3.12	Matrice de confusion du test de l'approche 1	60
3.13	Test de l'approche 1	62
3.14	Courbe de Perte de l'approche 2	65
3.15	Courbe de Précision de l'approche 2	66

3.16 Matrice de confusion du test de l'approche 2	67
3.17 Test de l'approche 2	69
3.18 Perturbations	72
3.19 Superpixels	73
3.20 Interface Graphique	83
3.21 Test avec l'interface graphique	83

Chapitre 1

Traitement des images Médicales par l'IA

1.1 Introduction

La maladie d'Alzheimer est une condition médicale complexe et débilitante, résultant d'une lente dégénérescence des neurones, qui commence souvent au niveau de l'hippocampe, une structure cérébrale fondamentale pour la mémoire, pour ensuite s'étendre progressivement au reste du cerveau. Cette dégénérescence entraîne des déficits cognitifs et fonctionnels significatifs, touchant non seulement la mémoire, mais aussi les fonctions exécutives, l'orientation dans le temps et l'espace, ainsi que l'autonomie des individus.[1]

La maladie d'Alzheimer se caractérise par la formation de plaques amyloïdes et de dégénérescences neurofibrillaires dans le cerveau, conduisant à une perte progressive de neurones et à une détérioration des fonctions cognitives [2]. Malheureusement, malgré les avancées spectaculaires dans la compréhension des facteurs de risque et des mécanismes de cette maladie, aucun traitement curatif n'est encore disponible, laissant les patients et leurs familles confrontés à un défi immense et souvent dévastateur.

Ce fléau insidieux affecte principalement les personnes âgées, mais des formes précoces peuvent également se manifester, ajoutant ainsi une couche de complexité à sa compréhension et à sa prise en charge. Bien que la recherche progresse, la maladie d'Alzheimer ne connaît pas de répit, et la quête de solutions efficaces pour ralentir sa progression ou améliorer la qualité de vie des patients demeure une priorité cruciale dans le domaine médical et scientifique.

1.2 Les images TEP scan dans l'étude de la maladie d'Alzheimer

Les images de tomographie par émission de positons (TEP scan) constituent une ressource inestimable dans l'étude des maladies neurodégénératives, en particulier la maladie d'Alzheimer. Cette technique d'imagerie offre une visualisation détaillée du métabolisme cérébral, permettant ainsi de détecter des changements subtils dans l'activité métabolique des différentes régions du cerveau. (Voir figure 1.1 et 1.2)



FIGURE 1.1 – Appareil TEP Scans

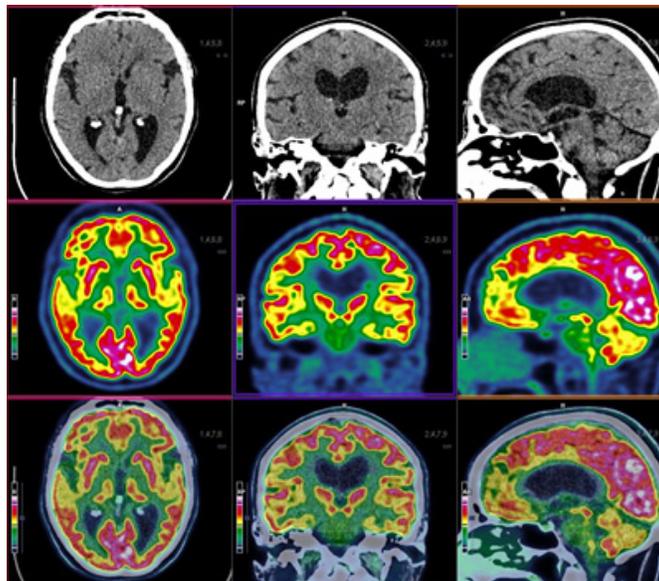


FIGURE 1.2 – Résultats TEP Scans

1.2.1 Principe de TEP scan

Un TEP scan (tomographie par émission de positons) est un test d'imagerie qui permet de montrer le fonctionnement des organes et des tissus. Lorsqu'il est associé à un traceur radioactif comme Amyvid™, le TEP scan peut détecter la maladie d'Alzheimer avant qu'elle n'apparaisse sur d'autres tests d'imagerie. Le traceur se lie aux plaques présentes dans le cerveau des patients atteints de la maladie d'Alzheimer. [3]

En se fixant sur les cellules cibles, le traceur émet des positons qui, en percutant les électrons environnants, émet des photons (particules de lumière) qui partent dans deux directions opposées. Une couronne de détecteurs disposée autour du patient capte ces couples de photons et un traitement informatique permet de reconstituer, à partir de ces émissions de photons, une image de l'organe étudié. [4] (Voir figure 1.3)

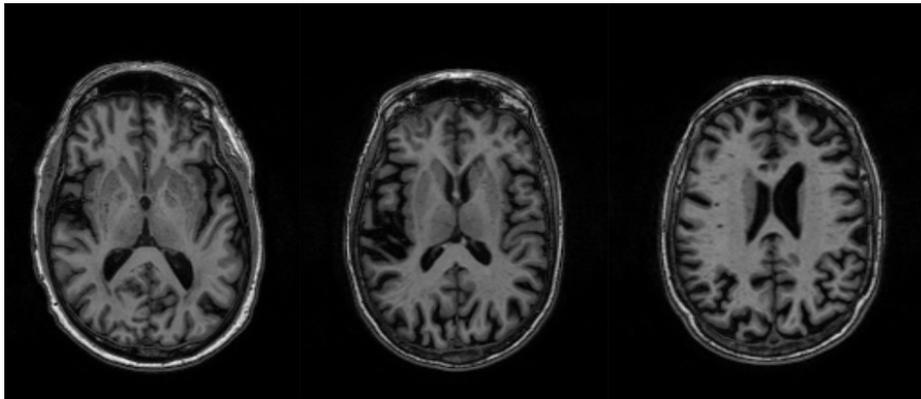


FIGURE 1.3 – Images TEP scans

1.2.2 TEP Scan et Radiotraceurs

La tomographie par émission de positons (TEP) est une technique d'imagerie médicale développée grâce aux avancées en physique médicale, biochimie, et neuropharmacologie au cours des trente dernières années [5]. Les radiotraceurs utilisés en TEP, qui étaient initialement des précurseurs ou des analogues métaboliques pour des études en laboratoire, ont été adaptés pour l'imagerie médicale. Ces radiotraceurs se fixent à des molécules spécifiques dans le corps, permettant ainsi de suivre leur distribution et leur accumulation dans différents tissus, ce qui est crucial pour la détection et le suivi de maladies comme l'Alzheimer.

Une nouvelle étude révèle que des TEP scan effectués avec un traceur de la protéine tau seraient en effet capables de prédire l'emplacement de futures atrophies cérébrales. Après des tests sur des patients atteints de la maladie d'Alzheimer à un stade précoce, puis lors d'une visite de suivi entre un et deux ans plus tard, l'étude semble indiquer que la protéine tau joue un rôle plus direct dans la dégénérescence et l'atrophie du cerveau que le peptide β -amyloïde. [6]

Les TEP scans permettent donc de visualiser les plaques présentes dans le cerveau, qui sont les principales responsables de l'endommagement et de la destruction des cellules nerveuses dans la maladie d'Alzheimer.

1.3 Importance de la détection précoce et de la classification précise

La détection précoce de la maladie d'Alzheimer revêt une importance capitale pour plusieurs raisons. Tout d'abord, elle permet une intervention thérapeutique précoce, ce qui peut ralentir la progression de la maladie et améliorer la qualité de vie des patients. De plus, une détection précoce offre aux patients et à leur famille la possibilité de planifier l'avenir et de prendre des décisions concernant les soins et le soutien nécessaires.

La classification précise des patients atteints de la maladie d'Alzheimer est tout aussi cruciale. Une identification précise des patients permet une gestion appropriée de la maladie, en permettant aux médecins de proposer des traitements adaptés et de surveiller l'évolution de la maladie de manière plus efficace.

En outre, l'utilisation de techniques avancées d'apprentissage automatique permet la classification des cellules cérébrales à partir d'images de TEP scans, fournissant ainsi des informations précieuses sur les caractéristiques anatomiques et fonctionnelles du cerveau associées à la maladie. De plus, une classification précise facilite la recherche clinique et permet le développement de nouveaux traitements et de stratégies de prise en charge personnalisées. [7]

En somme, la maladie d'Alzheimer représente un défi majeur pour la santé publique mondiale, nécessitant une détection précoce et une classification précise pour améliorer les résultats cliniques des patients et répondre aux besoins croissants de la population vieillissante.

1.4 Traitement des images médicales

1.4.1 Méthodes classiques de traitement d'images médicales

Les méthodes traditionnelles de traitement d'images médicales pour la détection et la classification de la maladie d'Alzheimer ont souvent utilisé des techniques de segmentation d'images et d'extraction de caractéristiques (Features). Ces approches reposent sur des algorithmes classiques de traitement d'images, tels que la transformation de Fourier, les filtres de Gabor et les descripteurs de texture, pour extraire des caractéristiques pertinentes à partir des images cérébrales.

Ensuite, des techniques de classification classiques telles que les machines à vecteurs de support (SVM), les réseaux de neurones artificiels (ANN) ou les méthodes de voisinage le plus proche (KNN) sont souvent utilisées pour classer les images médicales en fonction de leurs caractéristiques extraites.

1.4.2 Limitations des approches classiques

Bien que les méthodes classiques de segmentation des images et d'extraction des caractéristiques (en Anglais Features) aient été largement utilisées dans la détection et la classification de la maladie d'Alzheimer, elles présentent plusieurs limitations importantes.

Tout d'abord, ces approches reposent souvent sur des caractéristiques manuellement conçues, ce qui peut limiter leur capacité à capturer la complexité et la variabilité des images cérébrales.

En outre, ces méthodes nécessitent souvent une expertise approfondie pour sélectionner et concevoir des caractéristiques appropriées, ce qui rend le processus laborieux et sujet aux biais de l'utilisateur.

De plus, elles sont souvent sensibles au bruit et à la variabilité des données, ce qui peut entraîner une performance suboptimale, en particulier lorsque les images présentent des artefacts ou des variations de qualité. Enfin, ces approches peuvent manquer de capacité à généraliser à de nouveaux ensembles de données ou à capturer des modèles complexes et non linéaires présents dans les images cérébrales.

Cependant, bien que les méthodes classiques aient été largement utilisées dans le domaine du traitement d'images médicales pour la détection et la segmentation des images cérébrales de la

maladie d'Alzheimer, elles présentent des limitations significatives en termes de capacité à capturer la complexité et la variabilité des images cérébrales, ainsi que de généralisation à de nouveaux ensembles de données.

1.5 Apprentissage Automatique appliqué à la classification de la maladie d'Alzheimer

1.5.1 Vue d'ensemble des techniques de Machine Learning

L'Apprentissage Automatique (Machine Learning) a démontré son importance dans le domaine médical, notamment pour le diagnostic des maladies à partir des données collectées sur les IRM [8]. Ces techniques visent à permettre aux ordinateurs d'apprendre à partir des données, de prendre des décisions, ou de faire des prédictions basées sur ces apprentissages. Dans le contexte de la classification de la maladie d'Alzheimer, les techniques de Machine Learning sont utilisées pour extraire des modèles et des relations à partir des caractéristiques des images cérébrales, afin de discriminer entre les différents stades de la maladie ou entre les patients atteints de la maladie et les sujets sains.

1.5.2 Exemples des méthodes classiques de Machine Learning

Plusieurs algorithmes classiques de Machine Learning ont été appliqués à la classification de la maladie d'Alzheimer. Parmi eux, on trouve les machines à vecteurs de support (SVM), les arbres de décision et les méthodes de classification ensembliste comme les forêts aléatoires. Ces algorithmes sont souvent utilisés pour entraîner des modèles à partir de données d'images cérébrales préalablement extraites de patients atteints de la maladie d'Alzheimer et de sujets sains, afin de prédire le diagnostic ou le stade de la maladie à partir de nouvelles images. [9]

- Les SVM sont des algorithmes de classification supervisée qui cherchent à trouver l'hyperplan optimal séparant les différentes classes dans un espace de caractéristiques. Ils sont efficaces dans des espaces de grande dimension et sont robustes aux problèmes de surapprentissage. [10]

- Les arbres de décision utilisent une structure arborescente pour modéliser les décisions et leurs conséquences. Chaque nœud interne représente un test sur une caractéristique, chaque branche représente le résultat du test, et chaque feuille représente une classe. [11]
- Les forêts aléatoires (Random Forest) sont des ensembles de plusieurs arbres de décision, entraînés sur différentes parties des données avec des caractéristiques aléatoires. Elles améliorent la robustesse et la précision par rapport à un seul arbre de décision. [12]

1.5.3 Limitations des méthodes classiques de Machine Learning

Bien que les techniques de l'apprentissage automatique (Machine Learning) aient montré des promesses dans la classification de la maladie d'Alzheimer, elles présentent également des limitations importantes. Premièrement, ces approches reposent souvent sur des caractéristiques préalablement extraites des images cérébrales, ce qui peut limiter leur capacité à capturer des informations pertinentes et discriminantes. De plus, ces techniques peuvent être sensibles aux limitations des modèles de Machine Learning, en particulier lorsque les ensembles de données sont petits ou déséquilibrés.

L'apprentissage automatique est donc un outil puissant, capable de créer ses propres modèles à partir des données, même celles non qualifiées. Cependant, cet avantage peut aussi être une faiblesse.

En outre, les algorithmes classiques de Machine Learning peuvent rencontrer des difficultés à modéliser des relations complexes et non linéaires présentes dans les données d'images cérébrales. En conséquence, ils peuvent avoir des performances limitées lorsqu'ils sont confrontés à des données présentant une grande variabilité, des modèles subtils ou le fléau de dimensionalité. Enfin, ces approches peuvent manquer de capacité à généraliser à de nouveaux ensembles de données ou à s'adapter à des changements dans la distribution des données au fil du temps.

1.6 Deep Learning et ses progrès récents

1.6.1 Deep Learning et ses applications en vision par ordinateur

L'apprentissage profond ou Deep Learning est une branche du Machine Learning qui se concentre sur l'apprentissage de représentations hiérarchiques des données. Il utilise des architectures de réseaux neuronaux profonds pour apprendre automatiquement à partir des données et extraire des caractéristiques complexes et abstraites. En vision par ordinateur, le Deep Learning a révolutionné de nombreux domaines, y compris la détection d'objets, la reconnaissance faciale, et bien sûr, la classification d'images médicales telles que celles de la maladie d'Alzheimer.

1.6.2 Succès récents dans l'utilisation du Deep Learning pour la classification de la maladie d'Alzheimer

L'apprentissage profond (Deep Learning) a montré des succès remarquables dans la classification de la maladie d'Alzheimer en utilisant des données d'imagerie cérébrale. Les réseaux neuronaux profonds peuvent apprendre à partir de grandes quantités de données d'images médicales, capturant ainsi des motifs complexes et des relations non linéaires présentes dans les données.

Ces modèles sont capables de généraliser à de nouveaux ensembles de données et de surpasser les performances des approches traditionnelles de Machine Learning dans de nombreux cas. Ceci est appuyé par un article basé sur la détection précoce de la maladie d'Alzheimer grâce au Deep Learning des données de PET-FDG [13]. Cet article montre que l'apprentissage profond basé sur les données du PET-FDG améliorerait la capacité de l'imagerie cérébrale à prédire précocement la maladie d'Alzheimer, d'après une étude publiée dans la revue "Radiology" qui identifie les changements métaboliques avec une sensibilité significative.

1.6.3 Le Deep Learning et les changements métaboliques difficiles à évaluer au stade précoce

Le diagnostic précoce de la maladie d'Alzheimer revêt une importance capitale, car les traitements sont généralement plus efficaces lorsqu'ils sont administrés au début de la progression de la maladie. Les recherches ont démontré une corrélation entre l'évolution de la maladie et les altérations du métabolisme cérébral, notamment l'absorption de glucose dans certaines régions du

cerveau. Toutefois, ces changements métaboliques peuvent être difficiles à détecter.

Dans ce contexte, une étude publiée dans la revue “Radiology” a exploré le potentiel de l’Intelligence Artificielle (IA) pour améliorer la précision de l’imagerie cérébrale dans la prédiction de la maladie d’Alzheimer [14]. L’étude visait à déterminer si l’IA pouvait aider à identifier ces modifications subtiles du métabolisme cérébral plus tôt et avec une plus grande précision que les méthodes traditionnelles. Les résultats de cette recherche pourraient ouvrir la voie à des diagnostics plus précoces et des interventions thérapeutiques plus efficaces, offrant ainsi de meilleurs résultats pour les patients.

1.6.4 Exemples d’architectures des réseaux neuronaux profonds

Plusieurs architectures de réseaux neuronaux profonds ont été explorées pour la classification d’images médicales de la maladie d’Alzheimer, offrant des approches diverses pour la capture et l’exploitation des informations présentes dans les images cérébrales.

Parmi celles-ci, les réseaux de neurones convolutionnels (CNN) sont particulièrement populaires en raison de leur capacité à extraire des caractéristiques spatiales et locales pertinentes à partir des images.

Dans ce même contexte, les autoencodeurs peuvent être utilisés pour extraire des caractéristiques significatives des images cérébrales tout en réduisant les dimensions, ce qui peut faciliter la classification en fournissant des représentations plus discriminantes.

1.6.4.1 Les réseaux de neurones convolutionnels (CNN)

a. Définition et Principes de Base des CNN

Les réseaux de neurones convolutionnels (CNN) sont des réseaux de neurones multicouches spécialement conçus pour traiter des données structurées en grille, telles que les images ou les vidéos.

Les CNN ont été inventés à la fin des années 1990 par Yann LeCun, un chercheur français. Avec Yoshua Bengio et Geoffrey Hinton, ces trois pionniers sont largement reconnus comme les pères du Deep Learning. [15]

Ces réseaux sont capables d’apprendre par étapes à reconnaître les caractéristiques (Features)

d'une image. Une fois qu'une caractéristique est apprise à un endroit de l'image, le réseau sera capable de la reconnaître n'importe où dans l'image, ce qui les rend particulièrement efficaces pour le traitement et l'analyse des données visuelles.

Les réseaux de neurones convolutifs (CNN) se distinguent des autres réseaux neuronaux par leurs performances supérieures avec des entrées de signaux d'image, de parole ou audio. [16] Ils présentent trois principaux types de couches et une fonction d'activation comme principes de base. (Voir figure 1.4)

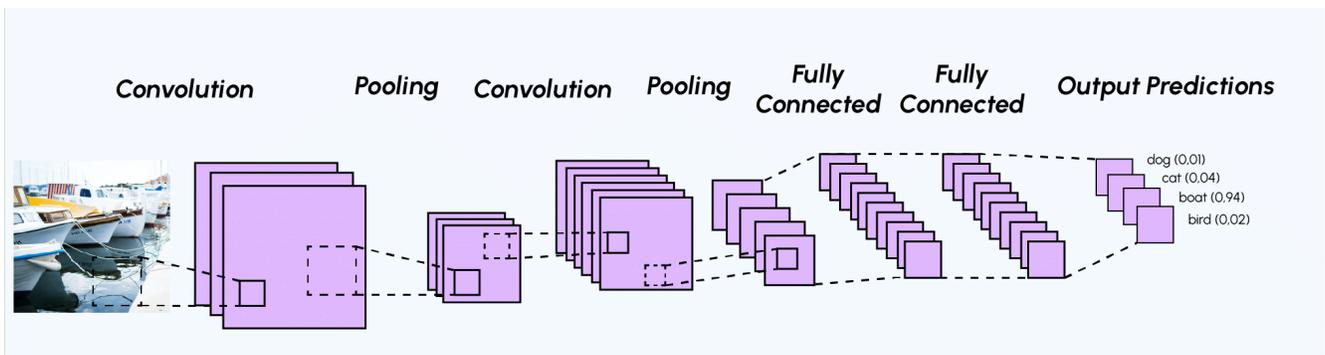


FIGURE 1.4 – Architecture de CNN

Principe de base :

- **Couches Convolutives (Convolutional Layers)** : La couche de convolution est la composante clé des réseaux de neurones convolutionnels et constitue toujours au moins leur première couche. Son but est de repérer la présence d'un ensemble de caractéristiques (**Features**) dans les images reçues en entrée. Pour cela, elle applique des filtres (**noyaux ou Kernel**) sur l'image d'entrée en réalisant un filtrage par convolution. Le principe est de faire "glisser" une fenêtre représentant une caractéristique sur l'image et de calculer le produit de convolution entre cette caractéristique et chaque portion de l'image balayée. Ces filtres permettent de détecter des motifs locaux tels que les bords, les textures et les formes. Chaque filtre génère une carte de caractéristiques qui met en évidence la présence de ces motifs dans l'image.
- **Couches de Pooling (Pooling Layers)** : La couche de pooling est souvent placée entre deux couches de convolution dans un réseau de neurones convolutionnels. Elle reçoit en entrée plusieurs cartes de caractéristiques (**feature maps**) et applique à cha-

cune d'elles l'opération de pooling, qui consiste à réduire la taille des images tout en préservant leurs caractéristiques importantes. Pour cela, l'image est découpée en cellules régulières, et au sein de chaque cellule, on conserve la valeur maximale (**max-pooling**) ou la valeur moyenne (**average-pooling**). Cette réduction de la dimensionnalité des cartes de caractéristiques permet de diminuer le nombre de paramètres et de calculs dans le réseau, améliorant ainsi son efficacité et aidant à prévenir le surapprentissage.

— **Couches Complètement Connectées (Fully Connected Layers)** : La couche entièrement connectée porte bien son nom : chaque nœud de la couche de sortie se connecte directement à un nœud de la couche précédente. Contrairement aux couches partiellement connectées, où les valeurs de pixel de l'image d'entrée ne sont pas directement connectées à la couche de sortie, chaque nœud de la couche entièrement connectée est lié à tous les nœuds de la couche précédente. Cette couche effectue la classification en se basant sur les caractéristiques extraites par les couches précédentes et leurs différents filtres. Tandis que les couches de convolution et de pooling utilisent souvent des fonctions d'activation comme **ReLU** (Rectified Linear Unit), les couches entièrement connectées exploitent généralement une fonction d'activation softmax pour classer les entrées, produisant un vecteur de distribution de probabilité de 0 à 1. En fin de réseau, ces couches intègrent les caractéristiques extraites pour effectuer la classification finale.

— **Fonctions d'Activation** : Les fonctions d'activation sont cruciales dans les réseaux de neurones, car elles introduisent de la non-linéarité, permettant au réseau d'apprendre des modèles complexes.

1. **ReLU (Rectified Linear Unit)** : La fonction ReLU est l'une des fonctions d'activation les plus populaires en raison de sa simplicité et de ses bonnes performances dans les réseaux de neurones profonds.

$$\text{ReLU}(x) = \max(0, x) \tag{1.1}$$

2. **Sigmoid** : La fonction sigmoid est souvent utilisée dans les réseaux de neurones pour transformer les sorties en valeurs comprises entre 0 et 1, ce qui est particulièrement utile pour les problèmes de classification binaire.

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (1.2)$$

3. **Tanh (Tangente Hyperbolique)** : La fonction tanh est similaire à la sigmoid mais s'étend de -1 à 1, ce qui peut être utile pour centrer les données et rendre l'apprentissage plus rapide.

$$\text{Tanh}(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (1.3)$$

4. **Leaky ReLU** : Leaky ReLU est une variante de ReLU qui permet un petit gradient lorsque l'unité est inactive, ce qui peut aider à résoudre le problème des unités "mortes". Une unité morte est un neurone qui ne contribue plus à la sortie du réseau et dont les poids ne sont plus mis à jour. Cela peut réduire l'efficacité de l'apprentissage et conduire à un réseau sous-optimal.

$$\text{Leaky ReLU}(x) = \begin{cases} x & \text{si } x > 0 \\ \alpha x & \text{sinon} \end{cases} \quad (1.4)$$

5. **Softmax** : La fonction softmax est utilisée principalement à la dernière couche d'un réseau de neurones pour des tâches de classification multiclasse. Elle convertit les scores en probabilités.

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (1.5)$$

b. Avantages des CNN pour la Classification des Images Cérébrales

Les réseaux de neurones convolutionnels (CNN) sont des outils d'intelligence artificielle puissants pour la reconnaissance d'images, utilisant l'apprentissage profond [17] pour accomplir des tâches à la fois génératives et descriptives. Leur capacité à traiter des informations visuelles complexes les rend particulièrement utiles pour la classification des images cérébrales, notamment dans le contexte du diagnostic de la maladie d'Alzheimer.

- **Extraction de Caractéristiques Complexes** : Les CNN sont composés de couches convolutives, où chaque neurone ne traite que les informations provenant d'une petite partie du champ visuel, générant ainsi une carte de caractéristiques. Cette structure permet aux CNN d'extraire des caractéristiques locales importantes, telles que les bords, les textures et les formes, qui sont essentielles pour détecter des altérations métaboliques dans les images cérébrales.
- **Adaptabilité et Généralisabilité** : Les CNN sont robustes aux variations interindividuelles et aux artefacts dans les images cérébrales, ce qui leur permet de généraliser efficacement à différents patients et conditions cliniques. Cette adaptabilité améliore la précision et la fiabilité des diagnostics en capturant des variations subtiles et des motifs non linéaires dans les données.

c. Exemples d'Applications Réussies dans le Diagnostic Précoce de la Maladie d'Alzheimer

Plusieurs études ont démontré l'efficacité des CNN dans le diagnostic précoce de la maladie d'Alzheimer en utilisant des images cérébrales. Parmi celles-ci, on trouve :

- **Classification Précise des Stades de la Maladie** : Les CNN ont été utilisés pour classer avec précision les différents stades de la maladie d'Alzheimer en identifiant les marqueurs spécifiques associés à chaque stade. Par exemple, Suk et Shen (2013) ont montré que les CNN pouvaient différencier les stades précoces de la maladie avec une grande précision en utilisant des images IRM. [18]

- **Diagnostic Différentiel** : Les CNN ont également été utilisés pour différencier la maladie d’Alzheimer d’autres pathologies cérébrales présentant des symptômes similaires. Cette capacité de diagnostic différentiel est essentielle pour recommander un traitement approprié et améliorer la qualité de vie des patients. L’étude de Payan et Montana (2015) a illustré cette application en utilisant des CNN pour différencier la maladie d’Alzheimer de troubles cognitifs légers et d’autres démences. [19]

Ces exemples démontrent le potentiel des CNN dans le diagnostic précoce de la maladie d’Alzheimer et soulignent leur capacité à améliorer les soins aux patients grâce à une détection plus précise et plus rapide des altérations métaboliques cérébrales.

1.6.4.2 Les autoencodeurs

a. Définition et Principes de Base des Autoencodeurs

Les autoencodeurs sont un type de réseau de neurones utilisés pour l’apprentissage non supervisé des représentations de données. Leur principal objectif est de construire une nouvelle représentation d’un jeu de données, généralement plus compacte et avec moins de descripteurs, permettant ainsi de réduire la dimensionnalité du jeu de données. L’architecture d’un autoencodeur comprend deux parties essentielles : l’encodeur et le décodeur. (Voir figure 1.5)

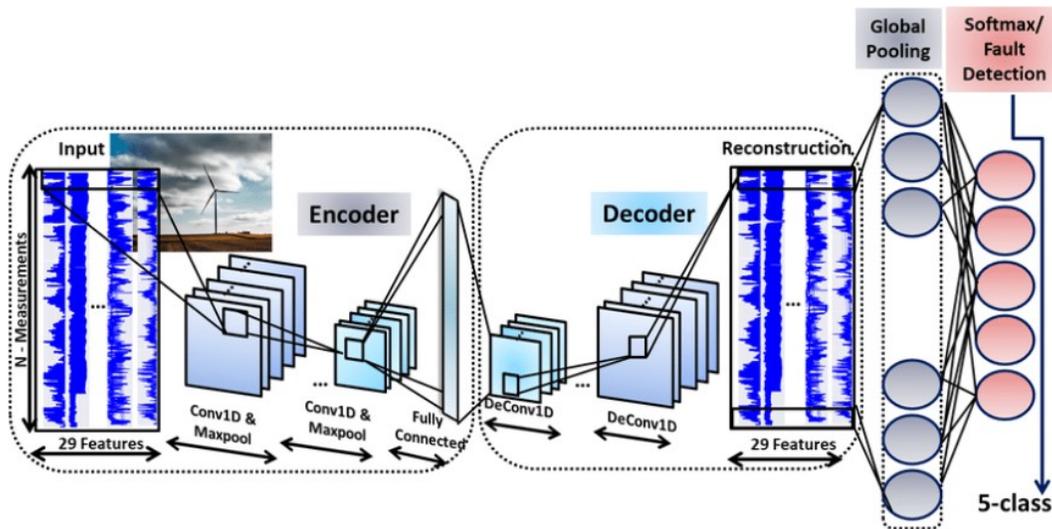


FIGURE 1.5 – Autoencodeur

Principe de base :

- **Encodeur** : Constitué d'un ensemble de couches de neurones, l'encodeur traite les données d'entrée pour construire de nouvelles représentations dites "encodées". Cette étape réduit les dimensions des données tout en conservant les informations les plus pertinentes.
- **Decodeur** : Les couches de neurones du décodeur reçoivent les représentations encodées et tentent de reconstruire les données d'origine. La qualité de cette reconstruction est mesurée en comparant les données reconstruites aux données initiales, permettant ainsi de calculer l'erreur de reconstruction.

L'entraînement de l'autoencodeur consiste à ajuster ses paramètres pour minimiser cette erreur de reconstruction. On s'intéresse principalement à la nouvelle représentation créée par l'encodeur plutôt qu'à la reconstruction finale des données.

Les autoencodeurs peuvent varier dans leur complexité et architecture, pouvant inclure des couches convolutives pour l'analyse d'images ou des couches récurrentes pour le traitement des séries temporelles.

b. Utilisation des autoencodeurs dans la Classification d'Images Médicales

Les autoencodeurs sont largement utilisés en imagerie médicale en raison de leur capacité à apprendre des représentations compactes et significatives des données. Ils ont un impact sur :

- **La Réduction de la Dimensionnalité** : En compressant les images cérébrales en codes latents plus petits, les autoencodeurs facilitent l'analyse de données complexes tout en conservant les caractéristiques essentielles.
- **La Détection d'Anomalies** : Les autoencodeurs peuvent identifier des anomalies en apprenant la distribution normale des données et en repérant les écarts significatifs de cette distribution.

- **La Préformation de Réseaux Supervisés** : En apprenant initialement une bonne représentation des données, les autoencodeurs peuvent préformer des réseaux supervisés, accélérant ainsi leur entraînement et améliorant leur performance pour la classification des images médicales.

c. Applications Réussies des Autoencodeurs dans le Diagnostic Médical

Les autoencodeurs ont démontré leur efficacité dans plusieurs applications de diagnostic médical, en particulier pour la détection et la classification des maladies neurodégénératives.

- **Diagnostic Précoce de la Maladie d’Alzheimer** : Les autoencodeurs sont utilisés pour extraire des caractéristiques significatives des images IRM cérébrales, facilitant la classification des stades précoces de la maladie d’Alzheimer. Par exemple, Suk et Shen (2013) ont utilisé des autoencodeurs pour apprendre des représentations compactes des images IRM, améliorant ainsi la précision du diagnostic précoce. [18]
- **Analyse des Images de TEP** : Les autoencodeurs analysent les images de tomographie par émission de positons (TEP), identifiant les altérations métaboliques caractéristiques de la maladie d’Alzheimer. Ces modèles détectent des anomalies subtiles dans les motifs d’absorption de glucose dans le cerveau, souvent précoces dans le développement de la maladie. Par exemple, Suk et al. (2014) ont démontré l’efficacité des autoencodeurs pour extraire des caractéristiques importantes des images de TEP, améliorant ainsi la détection précoce des altérations métaboliques liées à la maladie d’Alzheimer. [20]
- **Réduction des Artefacts et du Bruit** : En apprenant les représentations latentes des données d’imagerie, les autoencodeurs peuvent réduire les artefacts et le bruit, améliorant ainsi la qualité des images utilisées pour le diagnostic. Par exemple, Gondara (2016) a démontré l’efficacité des autoencodeurs pour la réduction du bruit dans les images médicales, ce qui permet d’obtenir des images plus claires et plus précises pour le diagnostic clinique. [21]

1.7 IA Explicable

1.7.1 Qu'est-ce que l'IA Explicable ?

L'Intelligence Artificielle Explicable (XAI) est un ensemble de processus et de méthodes qui permettent aux utilisateurs humains de comprendre, d'interpréter et de faire confiance aux résultats produits par les algorithmes de l'Intelligence Artificielle.

Comprendre le raisonnement derrière les résultats produits par les systèmes basés sur l'IA présente de nombreux avantages [16]. Cela peut permettre aux développeurs de garantir le bon fonctionnement du système, de répondre aux exigences réglementaires et de permettre aux personnes affectées par une décision de contester ou de modifier ces résultats. Cependant, rendre l'IA explicable peut également poser des défis, notamment en ce qui concerne la protection de la vie privée [22], l'accès à la propriété intellectuelle et la fiabilité des explications fournies.

Toutes les méthodes d'IA ne sont pas également affectées par les préoccupations relatives à l'explicabilité, et diverses approches peuvent être utilisées pour fournir des explications. Il existe des exemples de systèmes d'IA qui sont peu explicables mais peuvent être déployés sans problème, tandis que d'autres cas nécessitent des méthodes explicables et un cadre plus large pour garantir la responsabilité à chaque étape, de la collecte des données à la prise de décision. [23]

Les ingénieurs ou les data scientists de l'IA doivent prendre en compte les besoins des différents utilisateurs du système et réfléchir aux types d'explications qui pourraient être nécessaires et dans quel contexte elles seront utilisées.

1.7.2 Pourquoi l'IA Explicable ?

L'Intelligence Artificielle Explicable (XAI) est une branche de l'IA dédiée à la compréhension et à l'explication des décisions prises par les algorithmes. Contrairement à l'IA classique, souvent difficile à interpréter, l'XAI vise à fournir des explications claires et compréhensibles sur la logique derrière les décisions des machines. [24] (Voir figure 1.6)

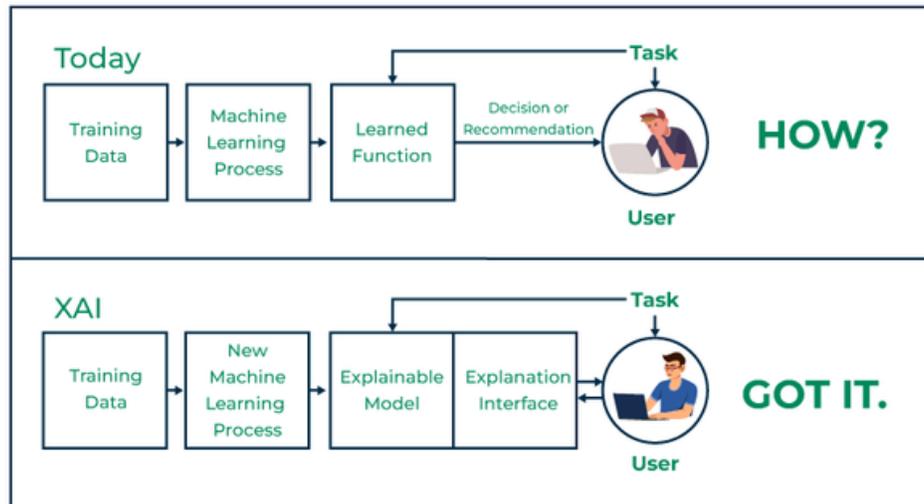


FIGURE 1.6 – IA Explicable

L’explicabilité consiste à rendre transparents les éléments et les variables pris en compte par le système d’IA pour produire un résultat. Cela inclut la capacité à relier les variables d’entrée à leurs conséquences sur les prévisions et les décisions, facilitant ainsi la compréhension du processus décisionnel.

Les explications fournies par l’XAI doivent être adaptées au niveau de compréhension des utilisateurs finaux, qu’il s’agisse de développeurs, de régulateurs ou de toute autre partie prenante. L’objectif est de garantir la transparence, la confiance et la responsabilité dans l’utilisation des technologies d’IA.

1.7.3 Les avantages de l’IA Explicable

Dans le paysage de l’intelligence artificielle moderne, l’accent est de plus en plus mis sur la transparence, la compréhensibilité et la responsabilité des modèles d’IA. L’IA Explicable (XAI) émerge comme une solution cruciale pour répondre à ces impératifs, offrant une compréhension claire et transparente des décisions prises par les systèmes d’Intelligence Artificielle. En explorant les avantages de l’IA explicable, nous plongeons dans un domaine où la confiance, l’efficacité et la gouvernance se rejoignent pour façonner l’avenir de l’Intelligence Artificielle.

— **Confiance dans l’utilisation de l’IA :**

L'IA explicable renforce la confiance des utilisateurs dans l'IA en fournissant une compréhension claire et transparente des décisions prises par les modèles. En garantissant l'interprétabilité et l'explicabilité des modèles d'IA, les utilisateurs sont plus enclins à utiliser ces modèles en toute confiance, ce qui permet un déploiement plus rapide des modèles en production. De plus, en simplifiant le processus d'évaluation des modèles et en augmentant la transparence et la traçabilité, l'IA explicable aide à créer un climat de confiance durable dans l'utilisation de l'Intelligence Artificielle.

— **Accélération des résultats de l'IA :**

En suivant et en gérant systématiquement les modèles, l'IA explicable permet d'optimiser les résultats métier. En évaluant continuellement et en améliorant les performances des modèles, les organisations peuvent affiner leurs efforts de développement et accélérer l'obtention de résultats pertinents et de haute qualité. L'évaluation continue des modèles permet également d'identifier rapidement les domaines nécessitant des ajustements, ce qui contribue à accélérer le processus de développement et de déploiement des modèles.

— **Risques et coûts de gouvernance réduits :**

L'IA explicable aide à réduire les risques et les coûts de gouvernance en rendant les modèles d'IA plus explicables et transparents. En garantissant la conformité réglementaire et la gestion efficace des risques, les organisations peuvent minimiser les frais liés aux inspections manuelles et aux erreurs coûteuses. De plus, en atténuant les risques de biais involontaires grâce à une compréhension approfondie des décisions prises par les modèles, les entreprises peuvent renforcer leur conformité et leur responsabilité tout en réduisant les coûts associés à la correction des erreurs.

1.7.4 Cas d'utilisation pour l'IA Explicable

L'IA explicable trouve des applications variées et essentielles dans de nombreux secteurs, où la transparence et la compréhension des décisions prises par les systèmes d'IA sont cruciales. Grâce à son approche, les entreprises et les institutions peuvent non seulement améliorer l'efficacité et la

précision de leurs processus, mais aussi instaurer un climat de confiance et de responsabilité. Voici quelques exemples marquants de cas d'utilisation de l'IA explicable dans divers domaines.

— **Santé :**

Dans le domaine de la santé, l'XAI peut être utilisée pour expliquer les décisions prises par des systèmes d'IA utilisés dans le diagnostic médical. Cela permet aux médecins de mieux comprendre les raisons derrière les recommandations d'un système d'IA et de prendre des décisions plus éclairées.

Par exemple, Holzinger et al. (2019) ont discuté de la causabilité et de l'explicabilité de l'IA en médecine, soulignant l'importance de la transparence pour la confiance des praticiens. [25]

Tjoa et Guan (2020) ont également effectué une enquête sur l'IA explicable dans le contexte médical, mettant en avant les bénéfices pour la précision des diagnostics. [26]

— **Services financiers :**

Dans le secteur financier, l'XAI peut être utilisée pour expliquer les décisions prises par des systèmes d'IA dans la gestion des risques, la détection de fraude et la prise de décisions d'investissement. Cela permet aux utilisateurs de mieux comprendre les raisons derrière les décisions prises par l'IA et de prendre des décisions plus informées.

Kuiper et al. (2021) ont exploré l'importance de l'IA explicable dans le secteur financier, en soulignant la nécessité de transparence, d'équité et de responsabilité pour assurer la confiance dans les institutions financières. [27]

Quant à Benhamou et al. (2023), ils ont proposé un cadre commun pour la description et la prédiction des décisions financières via des modèles d'IA explicables appliqués aux marchés financiers. [28]

— **Transport :**

Dans le domaine des transports, l'XAI peut être utilisée pour expliquer les décisions

prises par des véhicules autonomes. Cela permet aux passagers et aux piétons de mieux comprendre pourquoi un véhicule autonome prend une certaine décision et de se sentir en sécurité lorsqu'ils utilisent ces véhicules.

Par exemple, Anderson et al. (2014) ont fourni des directives pour les décideurs politiques concernant la technologie des véhicules autonomes, en soulignant la nécessité d'une IA explicable pour la sécurité des utilisateurs. [29]

Biran et Cotton (2017) ont aussi exploré l'importance des explications et justifications dans le Machine Learning pour les systèmes de transport. [30]

— **Justice pénale :**

Dans le domaine de la justice pénale, l'XAI peut optimiser les processus de prédiction et d'évaluation des risques. Elle peut également accélérer les résolutions en utilisant l'IA explicable pour l'analyse de l'ADN, l'analyse de la population carcérale et les prévisions criminelles.

Par exemple, Rudin et Ustun (2018) ont discuté des systèmes de score optimisés pour une confiance accrue dans le Machine Learning pour la santé et la justice pénale. [31]

— **Industrie :**

Dans l'industrie, l'XAI peut être utilisée pour expliquer les décisions prises par des systèmes d'IA dans la gestion de la chaîne d'approvisionnement, la maintenance prédictive et la gestion de la qualité. Cela permet aux entreprises de mieux comprendre les raisons derrière les décisions prises par l'IA et de prendre des mesures appropriées.

Dans une publication publiée en juin 2019, Gunning et Aha (2019) ont discuté du programme XAI de la DARPA, en mettant en évidence les défis et les solutions pour l'explicabilité dans les systèmes industriels complexes. [32]

1.7.5 Quelques méthodes de l'IA Explicable

L'intelligence artificielle explicable (XAI) repose sur un ensemble diversifié de méthodes conçues pour rendre les modèles de machine learning et de deep learning plus interprétables et compréhensibles. Ces techniques permettent non seulement de démystifier le processus de prise de décision des modèles d'IA, mais aussi d'assurer la transparence et la confiance dans les systèmes automatisés.

En utilisant ces méthodes, les utilisateurs peuvent mieux comprendre les mécanismes internes des algorithmes complexes, identifier les biais potentiels et garantir que les modèles fonctionnent de manière équitable et responsable. Il existe de nombreuses méthodes et outils pour interpréter les décisions prises par un algorithme, à savoir :

— LIME (Local Interpretable Model-agnostic Explanations) :

LIME (Local Interpretable Model-agnostic Explanations) est l'une des méthodes les plus populaires pour interpréter les décisions prises par les algorithmes d'Intelligence Artificielle. Développée par trois chercheurs de l'Université de Washington, dont Marco Tulio Ribeiro [33], actuellement chercheur chez Microsoft Research, LIME est une librairie open-source qui propose des explications pour les prédictions des algorithmes de classification.

LIME s'inscrit dans une approche visant à identifier pas à pas les variables d'entrée qui ont conduit aux résultats produits par l'algorithme. Cette méthode est particulièrement utile en computer vision, où elle permet de remonter aux pixels ou aux régions d'une image ayant joué un rôle déterminant dans la reconnaissance de formes. LIME peut être appliqué à divers types de données, y compris les données tabulaires, textuelles et les images.

Dans un article publié en Aout 2016 [34], Marco Tulio Ribeiro et ses collègues ont bien expliqué le fonctionnement de la méthode LIME en présentant quelques exemples bien expliqués de cette dernière dans différents contextes.

— SHAP (Shapley Additive exPlanations) :

La méthode SHAP (Shapley Additive exPlanations) repose sur la théorie des jeux

et permet d'expliquer les résultats des modèles d'apprentissage automatique. En utilisant SHAP, il est possible d'obtenir une explication locale de l'allocation optimale des crédits grâce aux valeurs de Shapley. Cette approche tire parti de la théorie des jeux pour fournir des explications efficaces et précises sur les prédictions des modèles.

En nous concentrant sur les valeurs de Shapley, une méthode clé de l'intelligence artificielle explicable (XAI), nous pouvons mieux comprendre comment chaque facteur contribue aux résultats d'un modèle. Les valeurs de Shapley sont basées sur des concepts mathématiques rigoureux de la théorie des jeux. Elles évaluent les contributions marginales de chaque facteur en considérant l'ordre dans lequel ils interviennent dans les permutations de groupe lors d'une compétition.

Cette approche permet de déterminer précisément la contribution de chaque facteur à l'obtention d'un résultat, offrant ainsi une compréhension approfondie des processus décisionnels des modèles d'apprentissage automatique.

Par exemple, un article publié en janvier 2022 [35] offre un aperçu approfondi de l'utilisation des valeurs de Shapley (SHAP) pour expliquer le fonctionnement des modèles d'apprentissage automatique.

— **LRP (Layer-wise Relevance Propagation) :**

La Layer-wise Relevance Propagation (LRP), également connue sous le nom de "transmission de l'importance couche par couche", a été introduite pour la première fois en 2015. Cette méthode vise à décomposer les prédictions des réseaux de neurones en redistribuant la sortie finale vers les caractéristiques d'entrée. En attribuant une "pertinence" à chaque caractéristique, la LRP permet de visualiser quelles parties de l'entrée ont le plus contribué à la décision du modèle.

LRP est une technologie essentielle pour comprendre les spécificités des vecteurs d'entrée dans les modèles de réseaux de neurones. Applicable à divers types de données, tels que les images, les vidéos ou le texte, elle offre un moyen puissant d'analyser et d'interpréter le fonctionnement des modèles d'Intelligence Artificielle.

Les expérimentations réalisées sur une approche basée sur la méthode LRP pour l'ex-

plication des Réseaux de Neurones Convolutifs appliqués à la classification des textes [36], ont bien expliqué et décrit le fonctionnement de la méthode LRP et sa performance.

Il était question de combiner les méthodes LRP et Jacovi afin d’obtenir une méthode d’explicabilité des Text-CNN plus générique. Les résultats obtenus montrent que leur méthode permet effectivement d’expliquer plusieurs architectures de Text-CNN en conservant les propriétés de la méthode LRP et ayant une fidélité supérieure à celle de LIME.

1.8 Conclusion

La classification de la maladie d’Alzheimer à l’aide du traitement d’images et des techniques d’apprentissage automatique constitue un domaine de recherche crucial pour la santé publique, avec des avancées significatives réalisées au fil des décennies, passant des méthodes classiques aux techniques avancées de Machine Learning et de Deep Learning.

Les méthodes classiques, telles que l’extraction de caractéristiques manuelles et les algorithmes de classification classiques, ont été les premières à être explorées. Cependant, ces approches présentent des limitations en termes de capacité à capturer la complexité des images cérébrales et de performances limitées.

Les avancées récentes en Deep Learning ont révolutionné ce domaine en exploitant la puissance des réseaux neuronaux profonds pour apprendre des représentations hiérarchiques complexes à partir des données. Des architectures telles que les réseaux de neurones convolutionnels et les auto-encodeurs ont permis une extraction automatique de caractéristiques pertinentes et une amélioration significative des performances de classification.

L’intégration de techniques de Machine Learning et de Deep Learning, combinée à des méthodes avancées de prétraitement d’images et d’extraction de caractéristiques, a permis d’obtenir des résultats encore plus prometteurs. Cette approche combinatoire exploite au mieux les avantages respectifs de ces techniques et améliore la robustesse et la généralisation des modèles de classification.

Malgré ces avancées, des défis subsistent, notamment la nécessité de collecter et d’annoter de grandes quantités de données, la garantie de la généralisation des modèles à des populations diverses, et la compréhension approfondie des mécanismes sous-jacents de la maladie.

En conclusion, la recherche continue dans ce domaine est essentielle pour relever ces défis et réaliser pleinement le potentiel des avancées récentes en Deep Learning. Ces efforts peuvent non seulement améliorer la précision des outils de diagnostic de la maladie d'Alzheimer, mais également ouvrir la voie à de nouvelles stratégies de traitement et de prévention, offrant ainsi un espoir renouvelé aux millions de personnes touchées par cette maladie dans le monde.

Chapitre 2

Approche de la classification des images TEP de la maladie d'Alzheimer

2.1 Introduction

Dans ce chapitre, nous présentons notre approche pour la prédiction de la maladie d'Alzheimer à partir des images cérébrales. Notre méthodologie se distingue par l'utilisation d'un modèle autoencodeur pour le débruitage des images, suivi par la classification de ces images débruitées à l'aide d'un réseau de neurones convolutionnels (CNN) au cas où les images sont bruitées. Enfin, nous intégrons des techniques d'intelligence artificielle explicable (XAI) pour interpréter et expliquer les résultats obtenus.

L'objectif principal de notre approche est d'améliorer la précision et la robustesse de la classification des images cérébrales en réduisant le bruit présent dans les données initiales, ce qui permet une meilleure extraction des caractéristiques pertinentes par les CNN. En expliquant les décisions prises par notre modèle de classification, nous visons également à renforcer la transparence et la confiance dans les systèmes d'intelligence artificielle utilisés pour le diagnostic médical.

2.2 Base de Données

La base de données utilisée pour notre étude, est constitué d'images TEP scans (Tomographie par Émission de Positrons) issues de l'ensemble de données de base ADNI (Alzheimer's Disease

Neuroimaging Initiative). [37]

Ce type des bases de données est essentielle pour la classification de la maladie d'Alzheimer, car il fournit des informations visuelles détaillées sur l'état du cerveau à travers des images axiales 2D. (Voir figure 2.1)

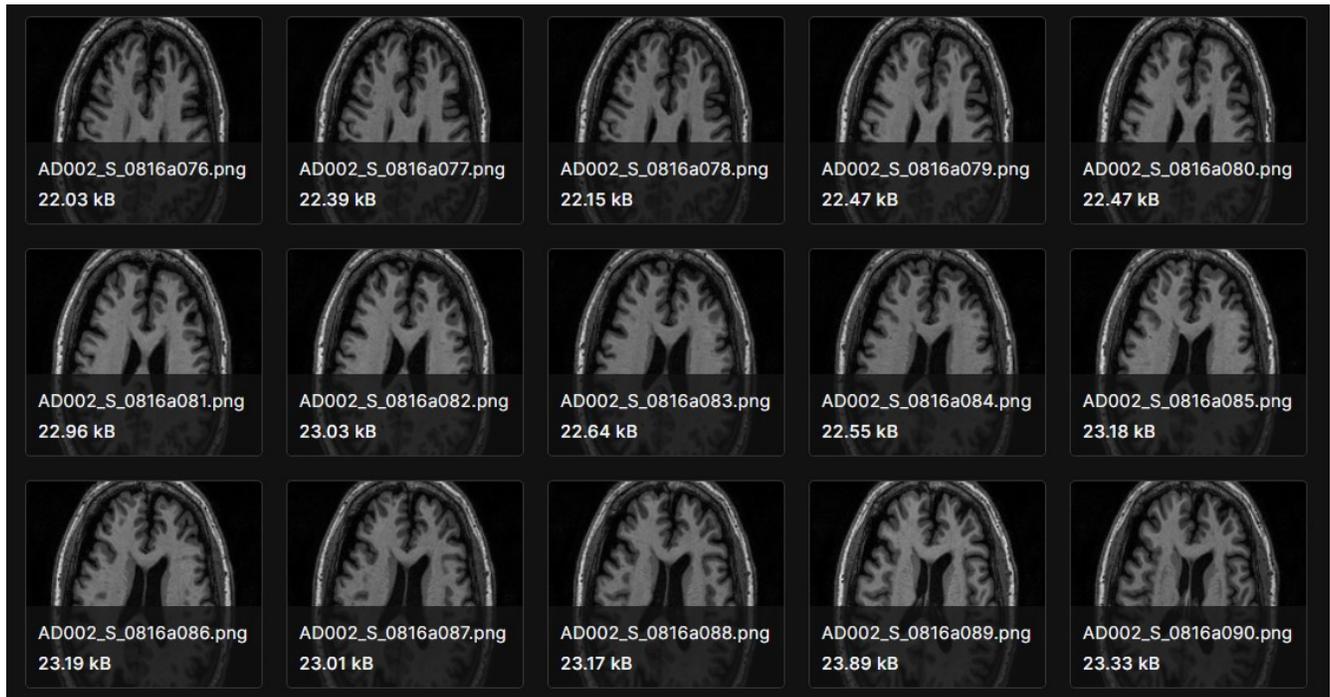


FIGURE 2.1 – Base de Données

2.2.1 Description de la Base de Données

L'ensemble de données ADNI est une ressource largement utilisée dans la recherche sur la maladie d'Alzheimer. Il contient diverses modalités d'images cérébrales qui sont particulièrement utiles pour détecter les anomalies métaboliques associées à la maladie d'Alzheimer. Les images TEP permettent de visualiser l'activité métabolique du cerveau, fournissant des indices précieux sur les changements pathologiques.

2.2.2 Format des Données

Les images TEP de notre base de données sont extraites du format NIfTI (Neuroimaging Informatics Technology Initiative) [38], un format de fichier standard pour les données de neuroimagerie. Le format NIfTI est conçu pour répondre aux besoins de la communauté de recherche en imagerie par résonance magnétique fonctionnelle (IRMf) et autres techniques de neuroimagerie. Il permet de stocker des images cérébrales volumineuses de manière structurée, facilitant ainsi leur traitement et analyse.

Pour notre étude, nous avons utilisé des images NIfTI converti en images axiales 2D, ce qui simplifie l'entraînement des modèles de classification tout en conservant des informations critiques sur les structures cérébrales.

2.2.3 Les Classes de la Base de Données

La base de données utilisée dans notre étude est divisée en trois classes principales, permettant une classification multicatégorielle des images cérébrales. On a :

- **AD (Alzheimer's Disease)** : Cette classe contient des images de patients diagnostiqués avec la maladie d'Alzheimer. Les images montrent généralement des réductions métaboliques dans des régions spécifiques du cerveau, caractéristiques de la maladie.
- **CI (Cognitive Impairment)** : Cette classe inclut des images de patients présentant une déficience cognitive légère, un état intermédiaire entre le vieillissement normal et la maladie d'Alzheimer. Les images CI sont cruciales pour la détection précoce et l'intervention avant que la maladie d'Alzheimer ne se développe pleinement.
- **CN (Cognitively Normal)** : Cette classe comprend des images de sujets cognitivement normaux, sans signes de déficience cognitive ou de maladie d'Alzheimer. Ces images servent de référence pour comparer les changements métaboliques dans les autres classes.

2.2.4 Source et pertinence de la Base de Données

L'ensemble de données ADNI est collecté et maintenu par l'Initiative de Neuroimagerie de la Maladie d'Alzheimer (ADNI) [39], un projet de recherche collaboratif majeur visant à développer

des biomarqueurs pour la progression de la maladie d'Alzheimer. La disponibilité de cette base de données facilite la recherche et le développement de nouvelles techniques de diagnostic et de traitement.

La pertinence de cette base de données pour notre étude réside dans sa richesse en données d'imagerie de haute qualité et bien annotées, couvrant un large éventail de stades de la maladie d'Alzheimer. Son utilisation permet de développer des modèles robustes et généralisables, essentiels pour des applications cliniques fiables.

En résumé, la base de données utilisée dans notre étude est une composante cruciale, fournissant les bases nécessaires pour entraîner, évaluer et interpréter nos modèles de classification de la maladie d'Alzheimer.

2.3 Collection des Données

La collecte des données est une étape cruciale dans notre processus de création de modèles de classification de la maladie d'Alzheimer. En utilisant un ensemble de données provenant de l'ADNI, nous avons accès à une variété d'images TEP scans représentant différents stades de la maladie et des individus en bonne santé. Ces données fournissent une base solide pour entraîner et évaluer nos modèles de classification.

2.3.1 Méthodologie

Pour garantir la cohérence et la facilité de traitement des données, nous avons adopté une méthodologie rigoureuse pour le chargement et la préparation des images à des fins d'analyse et de classification. Les étapes sont les suivantes :

- **Chargement des images :**

Pour simplifier la manipulation des images TEP scans provenant du répertoire spécifié, nous avons développé une fonction sur mesure en Python. Cette fonction est conçue pour parcourir le répertoire et charger chaque image individuellement à l'aide de la bibliothèque *OpenCV (cv2)*.

— **Redimensionnement des images :**

Afin d'assurer une cohérence dans les dimensions des images et de faciliter le traitement ultérieur, nous avons redimensionné chaque image à une taille spécifique de 128x128 pixels. Ce processus de redimensionnement est essentiel pour garantir que toutes les images ont la même taille, ce qui est crucial pour l'entraînement efficace des modèles de classification.

Ce redimensionnement uniforme garantit que toutes les images ont des dimensions compatibles, ce qui simplifie le processus d'entraînement des modèles et permet une analyse cohérente des caractéristiques visuelles des images.

— **Collecte des Étiquettes :**

En plus de charger les images, notre fonction personnalisée collecte également les étiquettes correspondantes pour chaque image. Ce processus est essentiel pour étiqueter correctement chaque image avec sa classe respective. Les étiquettes sont dérivées des noms de répertoires dans lesquels les images sont stockées.

Par exemple, si une image est située dans le répertoire "AD", elle est automatiquement étiquetée comme appartenant à la classe "AD" (maladie d'Alzheimer). Cette approche assure une association précise entre les images et leurs étiquettes respectives, ce qui est indispensable pour l'entraînement supervisé des modèles de classification.

En collectant les étiquettes de cette manière, nous établissons une correspondance directe entre les images et les catégories auxquelles elles appartiennent, fournissant ainsi les données d'entraînement nécessaires pour construire des modèles de classification précis et fiables.

— **Structuration des données :**

Après le chargement des images et la collecte de leurs étiquettes respectives, nous avons structuré les données de manière organisée. Les images sont stockées dans un tableau multidimensionnel, où chaque dimension représente une caractéristique de l'image, telle que sa hauteur, sa largeur et ses canaux de couleur.

Parallèlement, les étiquettes correspondantes sont stockées dans un tableau unidimensionnel aligné avec les images, garantissant ainsi que chaque étiquette est associée à la bonne image. Cette structuration des données permet une manipulation efficace lors des étapes suivantes de l'analyse et de la classification.

Elle crée également une cohérence dans la représentation des données, facilitant ainsi leur traitement par les algorithmes d'apprentissage automatique. En structurant les données de cette manière, nous nous assurons que chaque image est correctement associée à sa classe correspondante, ce qui est crucial pour l'entraînement et l'évaluation précis des modèles de classification.

2.3.2 Vue d'Ensemble des Données Collectées

Une fois toutes les images chargées, elles sont classées en trois catégories distinctes, chacune représentant une condition cérébrale spécifique.

Cette distribution des classes dans notre ensemble de données reflète la diversité des conditions cérébrales représentées dans les images TEP scans collectées. Chaque classe joue un rôle essentiel dans la compréhension et la classification des différentes conditions neurologiques, offrant ainsi une base solide pour notre analyse et nos modèles de classification. (Voir figure 2.2)

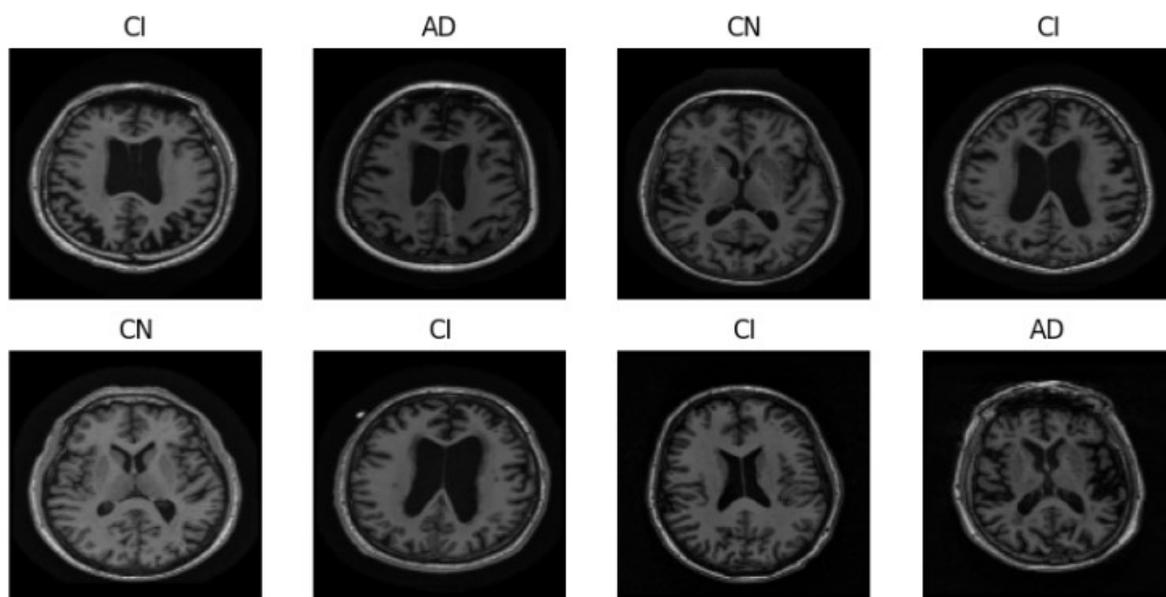


FIGURE 2.2 – Images collectées

Cette visualisation offre un aperçu immédiat de la variété des images et des différentes classes présentes dans notre ensemble de données. En visualisant ces images, nous pouvons apprécier les variations visuelles entre les différentes classes et obtenir une première impression des défis posés par la classification de ces images.

2.4 Prétraitement

Le prétraitement des données est une étape cruciale pour améliorer la qualité des images et faciliter l'entraînement des modèles de classification. Dans notre étude, nous avons appliqué plusieurs techniques de prétraitement aux images TEP scans pour optimiser leur utilisation.

2.4.1 Conversion en niveau de gris des images TEP scans

La conversion en niveaux de gris est la première étape du prétraitement de nos images TEP scans. Cette transformation joue un rôle crucial dans l'optimisation de notre base de données pour l'entraînement des modèles d'apprentissage profond (Deep Learning). En convertissant les images de couleur en niveaux de gris, nous réduisons la complexité des données tout en conservant les informations essentielles nécessaires pour la classification.

Bien que dans notre cas, les images TEP scans soient initialement en niveaux de gris, elles sont chargées avec trois canaux (rouge, vert, bleu) par défaut. Cette redondance de canaux n'apporte aucune information supplémentaire pertinente pour l'analyse. La conversion en niveaux de gris était donc nécessaire pour réduire le nombre de canaux à un seul, simplifiant ainsi les données.

Les images médicales, en particulier les TEP scans, se concentrent principalement sur les variations d'intensité plutôt que sur les variations des canaux de couleurs RGB. Les variations d'intensité dans ces images représentent des différences essentielles dans l'activité métabolique du cerveau, ce qui est fondamental pour détecter et analyser les anomalies associées à la maladie d'Alzheimer. La couleur, en revanche, n'apporte pas d'informations supplémentaires pertinentes dans ce contexte.

2.4.2 Normalisation des images TEP scans

Après la conversion en niveaux de gris, nous procédons à la normalisation des images. La normalisation consiste à ajuster les valeurs des pixels pour qu'elles se situent dans une plage uniforme, généralement entre 0 et 1. Cette étape est cruciale pour garantir que les variations d'intensité des pixels soient comparables à travers toutes les images, ce qui aide à stabiliser et accélérer l'entraînement des modèles d'apprentissage profond (**Deep Learning**).

La normalisation permet de rendre les valeurs de pixel homogènes, indépendamment des variations de contraste et de luminosité entre les images. En ajustant les valeurs de pixel à une plage standardisée, nous atténuons les effets des fluctuations de conditions d'éclairage ou de qualité d'image, ce qui est particulièrement important dans les analyses médicales où la précision est essentielle.

2.4.3 Débruitage des images TEP scans avec un Autoencodeur

Dans le cas où nous avons des images bruitées, il est crucial d'améliorer leur qualité avant de les utiliser pour l'entraînement du modèle de classification. L'objectif de cette étape est de garantir que le modèle de classification reçoit des données les plus claires et précises possibles, afin d'améliorer ses performances et la fiabilité des prédictions.

Pour ce faire, nous avons entraîné un autoencodeur, qui est une technique efficace pour débruiter les images en réduisant le bruit et en préservant les détails essentiels. L'autoencodeur apprend à reconstruire les images en se concentrant sur les caractéristiques importantes tout en éliminant les artefacts indésirables. En débruitant les images bruitées, nous augmentons la qualité des données d'entrée, ce qui est fondamental pour un apprentissage supervisé précis et robuste.

2.5 Construction du modèle CNN pour la classification

2.5.1 Architecture du Modèle

Après le prétraitement des images, la prochaine étape consiste à construire un modèle de réseau de neurones convolutifs (CNN) pour classer les images TEP en trois catégories : Maladie d'Alzheimer (AD), Trouble Cognitif Léger (CI), et Normal (CN). Un CNN est particulièrement

adapté pour la tâche de classification d'images en raison de sa capacité à capturer les caractéristiques spatiales et à réduire la complexité des données d'entrée grâce à ses couches convolutives et de pooling.

La conception de notre modèle CNN repose sur plusieurs couches, chacune jouant un rôle spécifique dans l'extraction des caractéristiques et la classification :

— **Couche Convulsive (Conv2D) :**

La première couche de notre modèle est une couche Conv2D avec 200 filtres, une taille de noyau de $(3, 3)$ et une fonction d'activation ReLU. Cette couche extrait des caractéristiques importantes des images d'entrée en appliquant des filtres convolutifs qui détectent des motifs locaux, comme les bords et les textures.

Ensuite la fonction ReLU introduit de la non-linéarité, accélère la convergence de l'entraînement et atténue le problème de gradient qui disparaît. Ainsi, cette couche permet au modèle de capturer une grande variété de motifs et de caractéristiques essentiels pour différencier les images PET en fonction des catégories cibles.

— **Couche de Pooling (MaxPooling2D) :**

La couche MaxPooling2D utilise une taille de pool de $(3, 3)$ pour réduire la dimensionnalité spatiale des cartes de caractéristiques générées par la couche convulsive précédente. Cette réduction diminue le nombre de paramètres et la charge de calcul tout en conservant les informations essentielles.

En sélectionnant les valeurs maximales dans chaque région de pooling, cette couche aide à extraire les caractéristiques les plus importantes et à rendre le modèle plus robuste aux variations et aux translations dans les images d'entrée.

— **Deuxième couche Convulsive (Conv2D) :**

La seconde couche Conv2D utilise 100 filtres avec une taille de noyau de $(3, 3)$ et une fonction d'activation ReLU. Cette couche continue le processus d'extraction de caractéris-

tiques, en s'appuyant sur les cartes de caractéristiques réduites produites par la première couche de pooling. En détectant des motifs plus complexes, cette couche permet au modèle de capturer des informations plus abstraites et détaillées des images d'entrée.

— **Deuxième couche de Pooling (MaxPooling2D) :**

La deuxième couche MaxPooling2D utilise une taille de pool de (3, 3). Cette opération réduit davantage la dimensionnalité spatiale des cartes de caractéristiques extraites par la couche précédente de convolution. En diminuant la résolution spatiale tout en préservant les caractéristiques les plus saillantes, cette couche permet au réseau de neurones de se concentrer sur des motifs plus larges et significatifs dans les images d'entrée.

Cette réduction de dimensionnalité est cruciale pour limiter le nombre de paramètres du modèle, améliorant ainsi l'efficacité computationnelle tout en conservant les informations essentielles pour la classification.

— **Couches Entièrement Connectées (Dense) :**

Les couches entièrement connectées (ou Dense) jouent un rôle crucial dans les réseaux de neurones convolutifs (CNN) pour la classification de nos images TEP.

La première couche Dense comprend 100 unités avec une fonction d'activation ReLU, permettant d'apprendre des combinaisons non linéaires des caractéristiques extraites par les couches convolutives précédentes. Cette capacité est essentielle pour capturer des relations complexes et des motifs abstraits dans les données d'entrée, améliorant ainsi la capacité du modèle à discriminer entre les différentes classes.

La deuxième couche Dense, avec 50 unités également activées par ReLU, poursuit ce processus en raffinant les représentations apprises par la couche précédente. Ensemble, ces couches entièrement connectées transforment les caractéristiques spatiales en des représentations plus abstraites et hiérarchiques, facilitant ainsi une classification précise et robuste des images TEP basée sur les caractéristiques extraites par les convolutions.

— Couche de Sortie (Dense) :

La couche de sortie (Dense) dans notre réseau de neurones convolutifs (CNN) est une couche Dense avec 3 unités et une fonction d'activation softmax. Cette configuration spécifique est cruciale pour résoudre notre problème de classification des images PET en trois classes distinctes.

La fonction d'activation softmax est choisie pour cette couche finale car elle transforme les logits en probabilités normalisées, permettant ainsi de déterminer la probabilité que chaque image appartienne à chaque classe respective. En produisant une distribution de probabilités sur les trois classes, cette couche nous permet de choisir la classe ayant la probabilité la plus élevée comme prédiction finale pour chaque image.

Ainsi, la couche de sortie joue un rôle crucial dans la décision finale de classification du modèle CNN, en convertissant les activations des couches précédentes en prédictions probabilistes claires et interprétables.

2.5.2 Compilation et Entraînement du modèle CNN pour la classification

Pour compiler et entraîner le modèle CNN pour la classification des images TEP, plusieurs choix stratégiques sont effectués afin d'optimiser les performances du modèle et de faciliter le processus d'entraînement.

— Optimiseur d'Adam :

L'optimiseur Adam est sélectionné pour ajuster les poids du réseau neuronal de manière efficace. Il combine les avantages des algorithmes AdaGrad et RMSProp, ce qui lui permet de gérer dynamiquement le taux d'apprentissage pour chaque poids du réseau. Cela est crucial pour converger rapidement vers une solution optimale tout en minimisant les risques de rester piégé dans des minima locaux.

— **Fonction de Perte (Sparse Categorical Crossentropy) :**

La fonction de perte “Sparse Categorical Crossentropy” est choisie spécifiquement pour les problèmes de classification multi-classes où les étiquettes sont des entiers plutôt que des vecteurs one-hot. Cette fonction mesure la dissimilarité entre la distribution de probabilité prédite par le modèle et la distribution réelle des étiquettes.

Contrairement aux vecteurs one-hot, où chaque classe est représentée par un vecteur binaire avec une seule valeur "1" (correspondant à la classe) et les autres valeurs à "0", la “Sparse Categorical Crossentropy” optimise directement la prédiction en utilisant les indices entiers des classes. Cela simplifie le calcul de la perte en évitant la comparaison entre des vecteurs de dimension égale, améliorant ainsi l’efficacité du processus de formation pour les tâches de classification multi-classes.

— **Callbacks :**

Le callback “*ModelCheckpoint*” est essentiel pour sauvegarder le modèle avec les meilleures performances sur l’ensemble de validation. Cela permet de conserver le modèle optimal obtenu pendant l’entraînement, ce qui est crucial pour garantir des performances optimales sur de nouvelles données après l’entraînement complet. En enregistrant automatiquement le modèle à chaque amélioration des performances de validation, on évite ainsi de perdre les progrès réalisés et on assure la disponibilité du meilleur modèle.

D’autre part, le callback “*EarlyStopping*” surveille les performances de validation du modèle à chaque époque et arrête l’entraînement prématurément si ces performances ne s’améliorent pas pendant un nombre spécifié d’époques consécutives (dans notre cas, 5 époques). En restaurant les poids du modèle à ceux correspondant aux meilleures performances de validation observées, ce callback prévient le surapprentissage en empêchant le modèle de s’adapter excessivement aux données d’entraînement et en favorisant une meilleure généralisation sur de nouvelles données.

Ces choix de configuration de **callbacks**, combinés avec l’utilisation de l’**optimiseur Adam** et de la fonction de perte “**Sparse Categorical Crossentropy**”, sont cruciaux pour maximiser la précision du modèle CNN dans la classification des images TEP en trois

catégories distinctes. Cette approche optimise l'entraînement du modèle tout en assurant des performances robustes et généralisables, particulièrement adaptées aux applications médicales nécessitant une précision et une fiabilité élevées pour la classification des maladies neurologiques à partir d'images médicales.

2.6 Explicabilité du Modèle avec LIME : IA Explicable

Une fois notre modèle de classification CNN construit et entraîné, il est crucial d'expliquer et de comprendre ses décisions, en particulier dans le contexte médical où les implications des prédictions peuvent être significatives. Pour cette raison, nous avons utilisé LIME (Local Interpretable Model-agnostic Explanations), une méthode qui aide à interpréter les prédictions des modèles complexes.

— Chargement du Modèle de classification et de l'Image :

Pour expliquer les prédictions de notre modèle, nous avons commencé par charger le modèle de classification CNN préalablement sauvegardé. Ensuite, nous avons sélectionné une image à analyser et nous l'avons prétraité pour correspondre aux attentes du modèle. Ce prétraitement est essentiel pour assurer que l'image d'entrée est traitée de manière cohérente par le modèle de classification, ce qui permet des prédictions précises et fiables.

— Prédiction du Modèle de classification :

Après avoir préparé l'image, nous l'avons passée à travers le modèle pour obtenir des prédictions de classe. Le modèle produit des probabilités pour chaque classe (Maladie d'Alzheimer, Trouble Cognitif Léger, et Normal), nous indiquant ainsi la classe la plus probable pour l'image donnée. Cette étape permet de vérifier que le modèle fonctionne correctement et d'identifier les scores de probabilité pour chaque classe, fournissant un aperçu initial des résultats de classification.

— Utilisation de LIME pour l’Explicabilité :

Pour générer des explications locales des prédictions du modèle, nous avons utilisé LIME. LIME fonctionne en perturbant l’image originale en générant des versions légèrement modifiées, et en observant comment les prédictions du modèle changent en réponse à ces perturbations. En analysant ces variations, LIME identifie les parties de l’image qui influencent le plus les décisions du modèle, permettant ainsi de comprendre les éléments clés qui conduisent à la classification.

— Visualisation des Résultats de LIME :

Les résultats de LIME sont visualisés en superposant une carte de chaleur sur l’image d’origine. Cette visualisation met en évidence les régions de l’image qui ont le plus contribué à la prédiction du modèle. La carte de chaleur permet de voir quelles parties de l’image sont les plus importantes pour le modèle lorsqu’il prend une décision, offrant une explication visuelle claire et intuitive de la décision du modèle.

2.7 Conclusion

Le prétraitement des images TEP et la construction d’un modèle CNN optimisé sont des étapes essentielles pour la classification précise des maladies neurologiques telles que la maladie d’Alzheimer. En convertissant les images en niveaux de gris, en normalisant les valeurs de pixels, et en utilisant un autoencodeur pour débruiter les images en cas de nécessité, nous allons améliorer la qualité des données d’entrée.

La conception d’un modèle CNN avec des couches convolutives et de pooling, suivie de couches entièrement connectées, permet une extraction et une interprétation efficaces des caractéristiques essentielles des images. Enfin, en utilisant des techniques de compilation et d’entraînement avancées, incluant l’optimiseur Adam et des callbacks stratégiques, nous avons maximisé les performances et la robustesse de notre modèle.

Ces étapes garantissent des prédictions fiables et interprétables, essentielles pour les applications médicales sensibles.

Chapitre 3

Résultats et Discussion

3.1 Introduction

Ce chapitre présente les résultats obtenus à partir de notre modèle de classification des images TEP pour la détection de la maladie d'Alzheimer. Après avoir appliqué les étapes de prétraitement des données et entraîné notre modèle CNN, nous évaluons ses performances à travers diverses métriques et visualisations.

Nous discuterons les principaux résultats, y compris la précision globale du modèle ainsi que son efficacité dans la classification des différentes classes : Maladie d'Alzheimer (AD), Trouble Cognitif Léger (CI) et Normal (CN). Les explications fournies par LIME seront également examinées pour comprendre les décisions prises par le modèle.

Cette analyse nous permettra de vérifier l'efficacité de notre approche et d'identifier les aspects nécessitant d'éventuels ajustements pour améliorer les performances du modèle dans la détection précoce de la maladie d'Alzheimer à partir d'images TEP.

3.2 Résultats

3.2.1 Approche 1 : Débruitage des images suivi de la classification

3.2.1.1 Description de l'Approche et Résultats

Dans cette première approche, nous avons entraîné un modèle autoencodeur variationnel pour débruiter notre base de données d'images TEP avant de les soumettre à un modèle CNN pour la

classification. L'objectif principal était d'améliorer la qualité des images en réduisant le bruit avant leur classification.

Par la suite, nous suivons les étapes ci-dessous :

1. L'entraînement de l'autoencodeur variationnel :

L'entraînement de l'autoencodeur variationnel pour débruiter les images avant la classification a été réalisé avec une approche stratégique utilisant une partie spécifique de notre base de données. Voici les détails de ce processus :

— Sélection des Données d'Entraînement :

Pour garantir une efficacité maximale de l'autoencodeur variationnel, nous avons sélectionné une portion représentative mais distincte de notre base de données d'images TEP. Cette sélection a été effectuée pour inclure une variété suffisante d'exemples tout en assurant une représentation presque équilibrée des classes : Maladie d'Alzheimer (AD), Trouble Cognitif Léger (CI), et Normal (CN). (Voir figure 3.1)

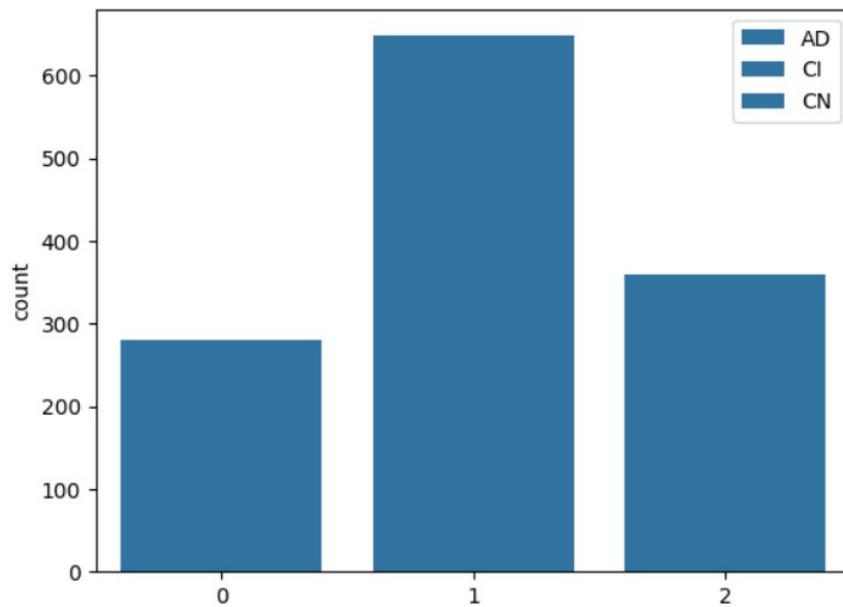


FIGURE 3.1 – Repartition des Images Sélectionnées

— Ajout de Bruit Gaussien :

Avant d'entraîner l'autoencodeur variationnel, un bruit gaussien a été artificiellement ajouté aux images sélectionnées. Ce processus de perturbation contrôlée a permis de simuler divers niveaux de bruit que l'autoencodeur devrait apprendre à reconnaître et à réduire pendant l'entraînement. L'objectif était de rendre le modèle robuste aux variations et aux imperfections présentes dans les images TEP réelles. (Voir figure 3.2 et 3.3)

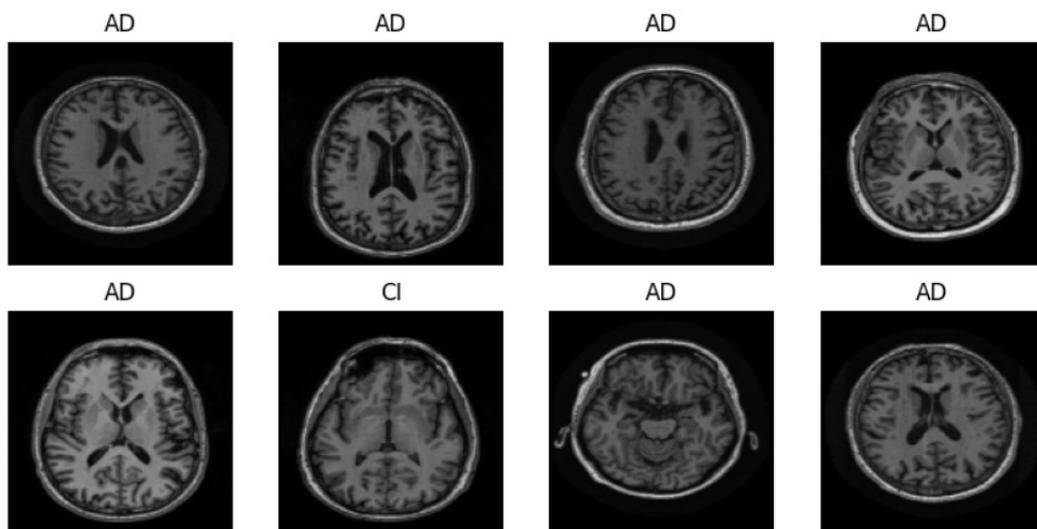


FIGURE 3.2 – Images Non Bruitées

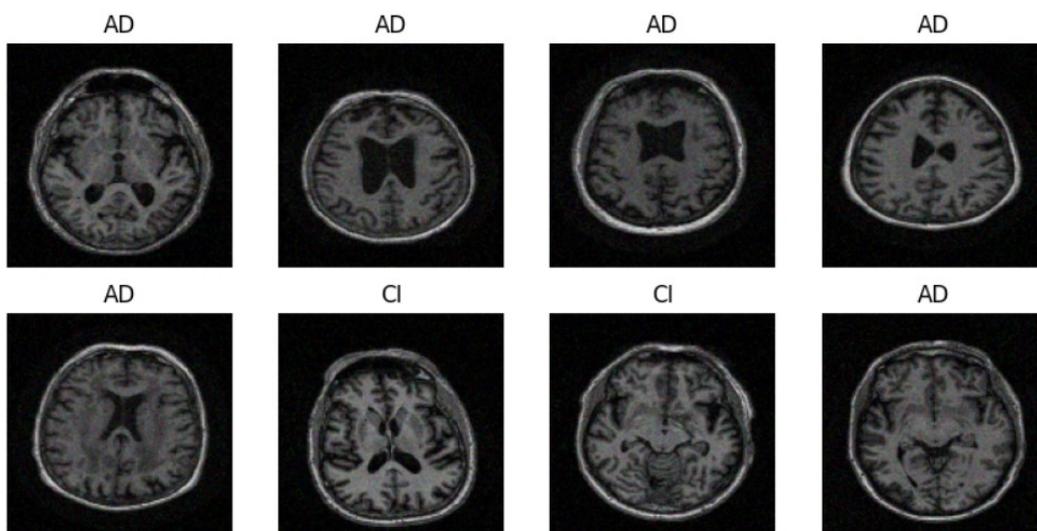


FIGURE 3.3 – Images Bruitées

— Architecture de l'Autoencodeur :

L'autoencodeur variationnel utilisé dans notre étude pour le débruitage des images TEP a été configuré avec des couches convolutionnelles adaptées aux caractéristiques complexes de ces images. Cette architecture a été optimisée pour extraire efficacement les caractéristiques essentielles tout en préservant la dimensionnalité des données d'entrée. (Voir figure 3.4)

```
Model: "model"
```

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 128, 128, 1)]	0
conv2d (Conv2D)	(None, 128, 128, 16)	160
max_pooling2d (MaxPooling2D)	(None, 64, 64, 16)	0
conv2d_1 (Conv2D)	(None, 64, 64, 8)	1160
max_pooling2d_1 (MaxPooling2D)	(None, 32, 32, 8)	0
conv2d_2 (Conv2D)	(None, 32, 32, 8)	584
up_sampling2d (UpSampling2D)	(None, 64, 64, 8)	0
conv2d_3 (Conv2D)	(None, 64, 64, 16)	1168
up_sampling2d_1 (UpSampling2D)	(None, 128, 128, 16)	0
conv2d_4 (Conv2D)	(None, 128, 128, 1)	145

FIGURE 3.4 – Architecture de l'autoencodeur

— Paramétrage et Validation :

Pendant l'entraînement, notre modèle a été guidé par une combinaison de la perte de reconstruction, typiquement l'erreur quadratique moyenne (MSE), qui mesure l'erreur de reconstruction entre les images débruitées et les originales. Cette approche a permis de structurer efficacement la représentation latente, assurant une reconstruction précise des images débruitées. Au cours des 200 epochs d'entraînement, l'erreur de

reconstruction a été progressivement réduite, démontrant l'efficacité continue de notre modèle dans la capture des caractéristiques importantes des images TEP.

En parallèle, des techniques de validation croisée ont été utilisées pour évaluer et ajuster régulièrement le modèle sur un ensemble de validation distinct, prévenant le surapprentissage et garantissant sa capacité à généraliser sur de nouvelles données TEP. (Voir figure 3.5)

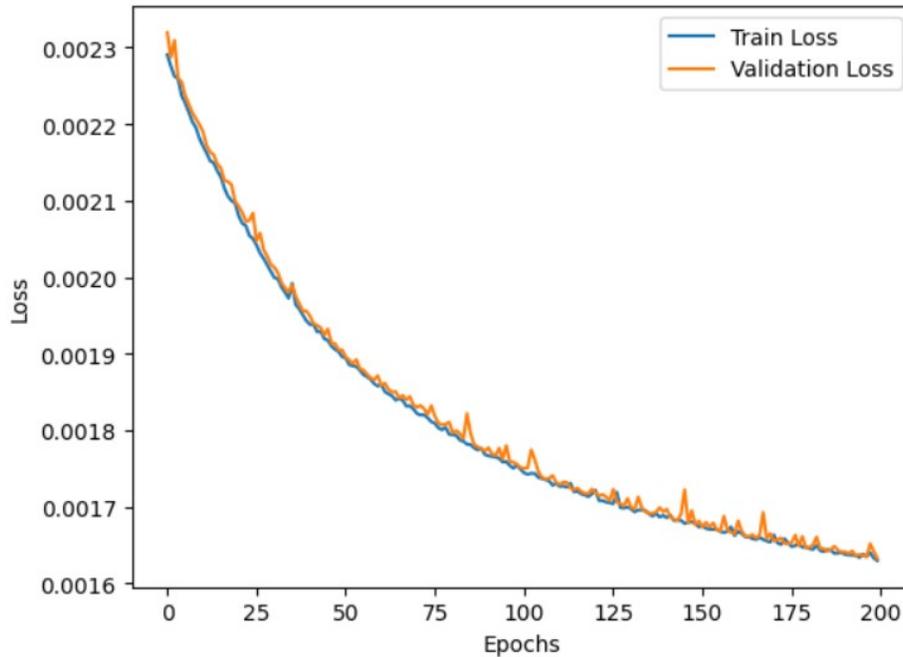


FIGURE 3.5 – Progression de l'erreur quadratique moyenne

2. Classification des images débruitées avec CNN :

Après avoir débruité les images à l'aide de l'autoencodeur variationnel, nous avons utilisé un réseau de neurones convolutifs (CNN) pour classer les images débruitées en trois catégories : Maladie d'Alzheimer (AD), Trouble Cognitif Léger (CI), et Normal (CN). Les détails de cette approche sont les suivants :

— Prétraitement des Données :

Avant la classification des images TEP, un prétraitement complet a été appliqué pour optimiser leur qualité et leur format d'entrée dans notre modèle CNN. Tout d'abord, les

images ont été converties en niveaux de gris et normalisées pour standardiser les valeurs de pixel, facilitant ainsi le traitement par la suite.

Ensuite, l'ensemble total des données a été soumis à un processus de débruitage utilisant l'autoencodeur préalablement entraîné. Cette étape visait à améliorer la clarté et la qualité des images en réduisant le bruit artificiel ajouté lors de l'acquisition des données TEP. (Voir figure 3.6, 3.7 et 3.8)

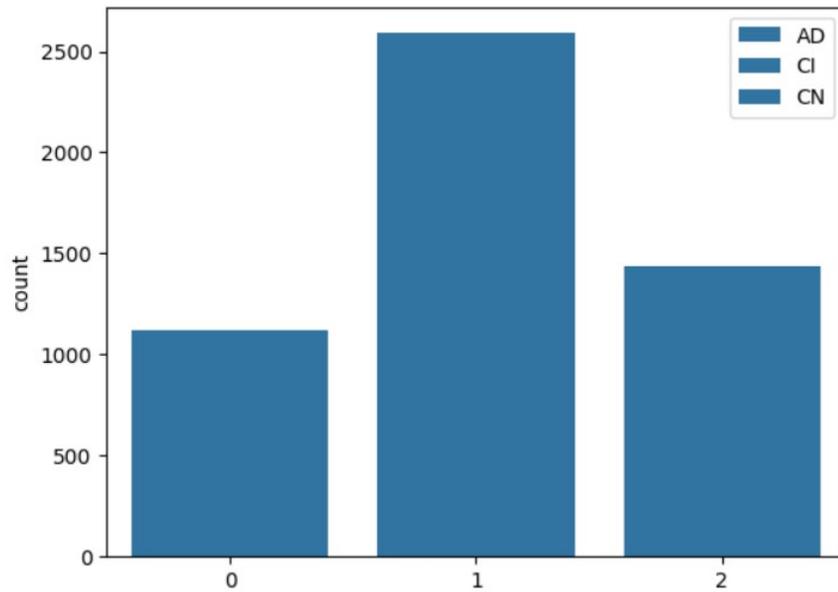


FIGURE 3.6 – L'ensemble des images TEP

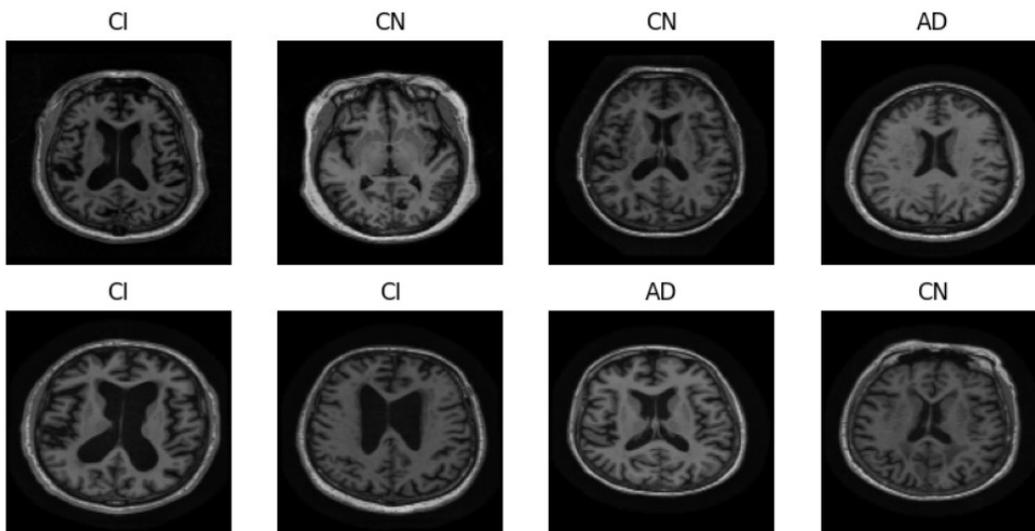


FIGURE 3.7 – Images TEP avant débruitage

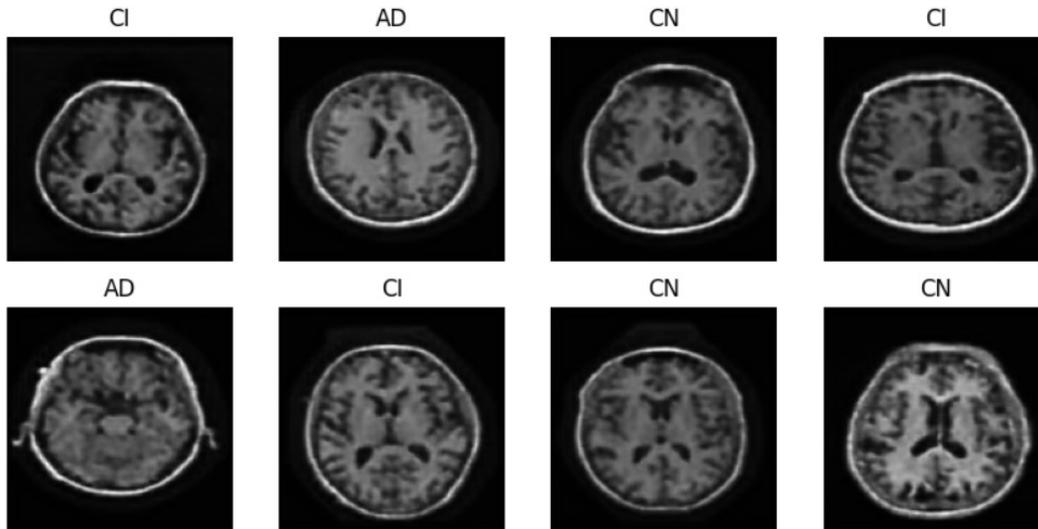


FIGURE 3.8 – Images Débruitées

On peut remarquer qu’après le débruitage, nos images TEP ont perdu en qualité. En comparant les images avant et après le débruitage, il est évident que les images TEP originales étaient déjà suffisamment claires. Le processus de débruitage a introduit des artefacts et dégradé la qualité des images, ce qui pourrait potentiellement affecter les performances de classification du modèle CNN.

Cependant, malgré cette dégradation, les détails essentiels des images TEP ont été préservés. Cette observation souligne l’importance d’évaluer soigneusement la nécessité du débruitage pour chaque ensemble de données spécifique et de s’assurer que cette étape améliore réellement la qualité des images plutôt que de la diminuer. Une analyse approfondie est essentielle pour déterminer si le débruitage est bénéfique ou s’il introduit des perturbations qui pourraient compromettre la précision de la classification.

— Architecture du Modèle :

Le CNN utilisé dans cette étude est configuré avec des couches convolutionnelles pour extraire les caractéristiques importantes des images débruitées, suivies de couches de pooling pour réduire la dimensionnalité et augmenter la robustesse du modèle. En fin de réseau, des couches entièrement connectées ont été employées pour effectuer la classification finale en trois classes : Maladie d’Alzheimer (AD), Trouble Cognitif Léger (CI), et Normal (CN). (Voir figure 3.9)

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 126, 126, 200)	2000
max_pooling2d (MaxPooling2D)	(None, 42, 42, 200)	0
conv2d_1 (Conv2D)	(None, 40, 40, 100)	180100
max_pooling2d_1 (MaxPooling2D)	(None, 13, 13, 100)	0
flatten (Flatten)	(None, 16900)	0
dense (Dense)	(None, 100)	1690100
dense_1 (Dense)	(None, 50)	5050
dense_2 (Dense)	(None, 3)	153

FIGURE 3.9 – Architecture du CNN

— Entraînement et Optimisation :

Le modèle CNN a été entraîné sur les images débruitées en utilisant une fonction de perte adaptée à la classification multi-classe, en l'occurrence la "SparseCategoricalCrossentropy", qui est efficace pour ce type de tâche. L'optimisation a été réalisée avec l'algorithme Adam, connu pour son efficacité dans l'entraînement de réseaux de neurones profonds.

Pour améliorer la performance du modèle et éviter le surapprentissage, plusieurs techniques de régularisation ont été mises en place. Cela inclut l'ajustement dynamique du taux d'apprentissage et l'utilisation de callbacks comme l'arrêt précoce ("EarlyStopping") pour interrompre l'entraînement si les performances sur les données de validation ne s'améliorent plus après 5 époques, et la sauvegarde du meilleur modèle avec "ModelCheckpoint". Ces stratégies ont permis d'assurer que le modèle puisse apprendre efficacement à partir des données débruitées tout en maintenant sa capacité à généraliser sur de nouvelles données.

— **Évaluation des Performances :**

- **Courbe de Perte (Entraînement et Validation) :** La courbe de perte montre la manière dont la perte du modèle diminue au cours de l'entraînement. Une perte plus faible indique que le modèle s'adapte bien aux données. La comparaison entre la perte d'entraînement et la perte de validation permet également d'identifier des problèmes potentiels de surapprentissage. (Voir figure 3.10)

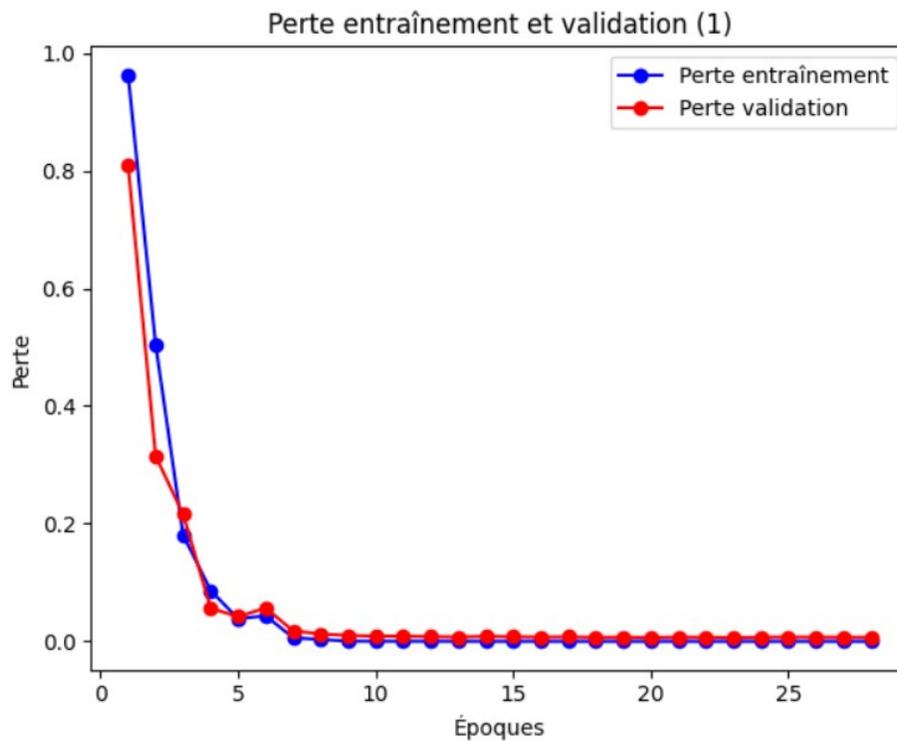


FIGURE 3.10 – Courbe de Perte de l'approche 1

- **Courbe d'Exactitude (Entraînement et Validation) :** La courbe d'exactitude, quant à elle, illustre la précision du modèle. Une exactitude plus élevée signifie que le modèle classe correctement un plus grand nombre d'images. Comme pour la perte, il est important de comparer l'exactitude d'entraînement et l'exactitude de validation pour s'assurer que le modèle ne surapprend pas. (Voir figure 3.11)

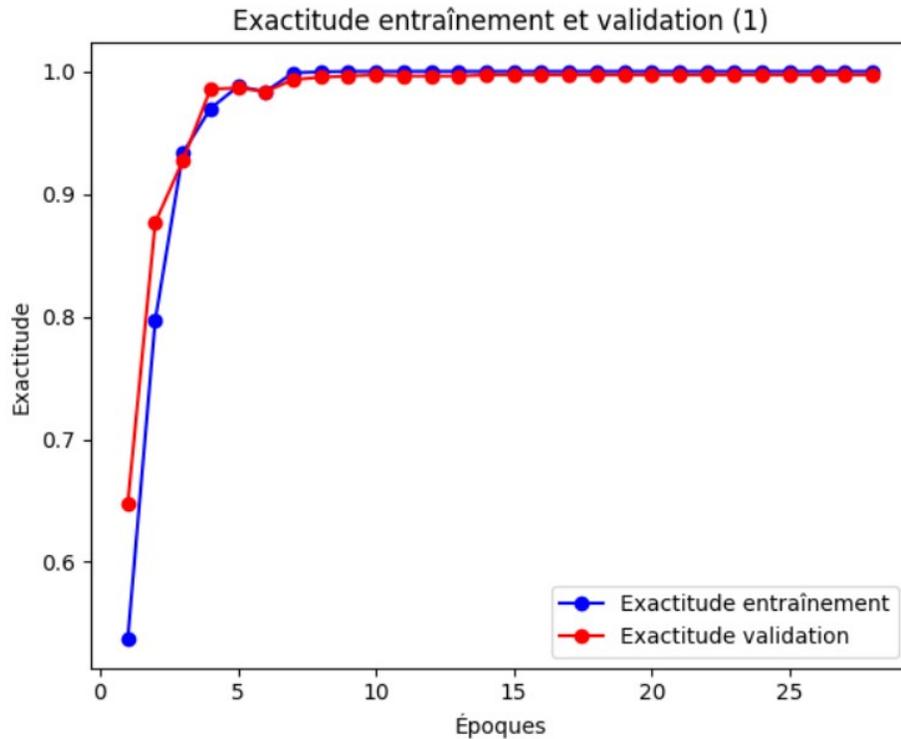


FIGURE 3.11 – Courbe de Précision de l’approche 1

- **Précision** : Sur les données d’entraînement, notre modèle a atteint une précision de 100%, soulignant sa capacité à apprendre les caractéristiques des données d’entraînement avec une grande exactitude. En revanche, sur les données de test, la précision s’est élevée à 99.71%, ce qui démontre la robustesse du modèle à généraliser sur de nouvelles données tout en conservant des performances élevées. Ces résultats mettent en évidence la capacité exceptionnelle du modèle CNN à maintenir une précision élevée à la fois sur les données utilisées pour l’entraînement et sur des données qu’il n’a pas encore vues.
- **Matrice de Confusion** : La matrices de confusion du test fournit une visualisation des résultats de classification du modèle, montrant les vrais positifs, les vrais négatifs, les faux positifs et les faux négatifs pour chaque classe. Elle est essentielle pour évaluer les performances du modèle de manière détaillée. (Voir figure 3.12)

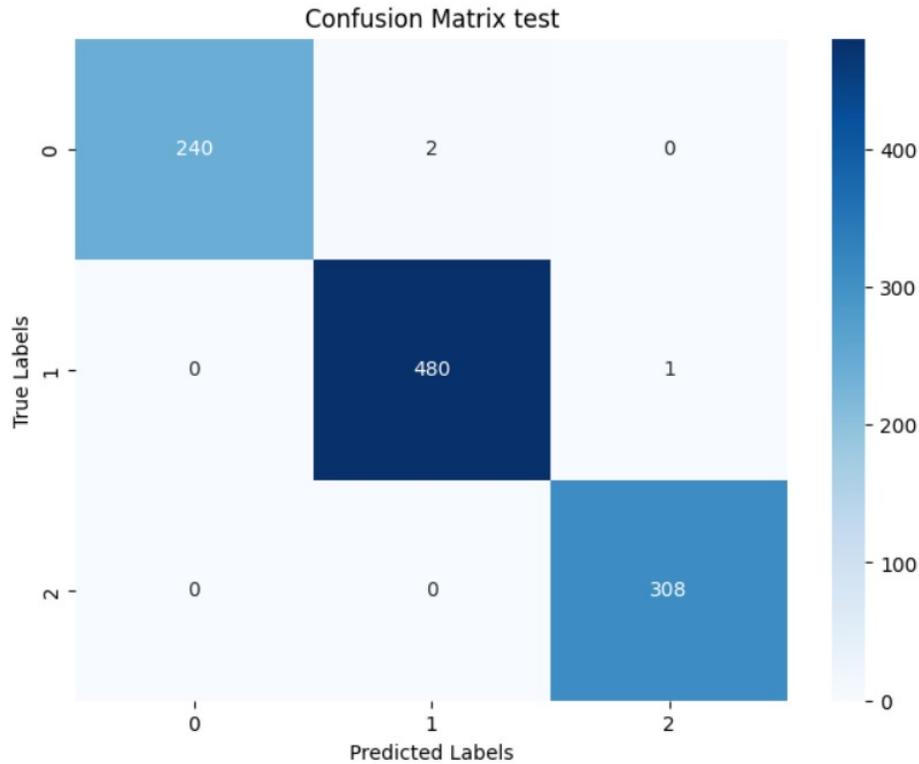


FIGURE 3.12 – Matrice de confusion du test de l’approche 1

— Résultats par Classe :

Classe	True Positives	True Negatives	False Positives	False Negatives
AD	240	789	0	2
CI	480	548	2	1
CN	308	722	1	0

TABLE 3.1 – Interpretation de la Matrice de Confusion de l’approche 1

Ce premier tableau présente la matrice de confusion, qui montre les vrais positifs (True Positives, TP), vrais négatifs (True Negatives, TN), faux positifs (False Positive, FP) et faux négatifs (False Negative, FN) pour chaque classe (Maladie d’Alzheimer (AD), Trouble Cognitif Léger (CI), et cas Normaux (CN)).

Maladie d'Alzheimer (AD) : Le modèle a correctement identifié 240 images comme appartenant à la classe AD (TP) et a correctement rejeté 789 images comme n'appartenant pas à la classe AD (TN). Il n'y a eu aucun faux positif (FP), mais 2 images ont été incorrectement classées comme n'étant pas AD alors qu'elles l'étaient (FN).

Trouble Cognitif Léger (CI) : Le modèle a correctement identifié 480 images comme CI (TP) et 548 images comme n'étant pas CI (TN). Il y a eu 2 images incorrectement identifiées comme CI (FP) et 1 image qui n'a pas été reconnue comme CI alors qu'elle l'était (FN).

Cas Normaux (CN) : Le modèle a correctement identifié 308 images comme CN (TP) et 722 images comme n'étant pas CN (TN). Il y a eu 1 image incorrectement identifiée comme CN (FP) et aucune image n'a été manquée (FN).

Classe	Précision	Rappel (Sensibilité)	F1-Score
AD	1.0	0.9917	0.9958
CI	0.9958	0.9979	0.9968
CN	0.9967	1.0	0.9983

TABLE 3.2 – Métriques de Performance de l'approche 1

Ce deuxième tableau présente les métriques de performance : précision, rappel (sensibilité) et F1-score pour chaque classe.

Maladie d'Alzheimer (AD) : La précision est de 1.0, ce qui signifie que toutes les images classées comme AD étaient correctes. Le rappel est de 0.9917, indiquant que le modèle a identifié 99.17% des images réellement AD. Le F1-score est de 0.9958, montrant un équilibre élevé entre précision et rappel.

Trouble Cognitif Léger (CI) : La précision est de 0.9958, ce qui signifie que 99.58% des images classées comme CI étaient correctes. Le rappel est

de 0.9979, indiquant que 99.79% des images réellement CI ont été correctement identifiées. Le F1-score est de 0.9968, reflétant une performance équilibrée.

Cas Normaux (CN) : La précision est de 0.9967, signifiant que 99.67% des images classées comme CN étaient correctes. Le rappel est de 1.0, indiquant que toutes les images réellement CN ont été correctement identifiées. Le F1-score est de 0.9983, démontrant une très bonne performance globale.

3.2.1.2 Test

Ici, on a les résultats de la classification de quelques images, issues de l'ensemble d'images TEP débruitées, par notre modèle CNN. Chaque image est accompagnée de son étiquette originale et de la classe prédite par le modèle, ainsi que de la précision de cette prédiction. (Voir figure 3.13)

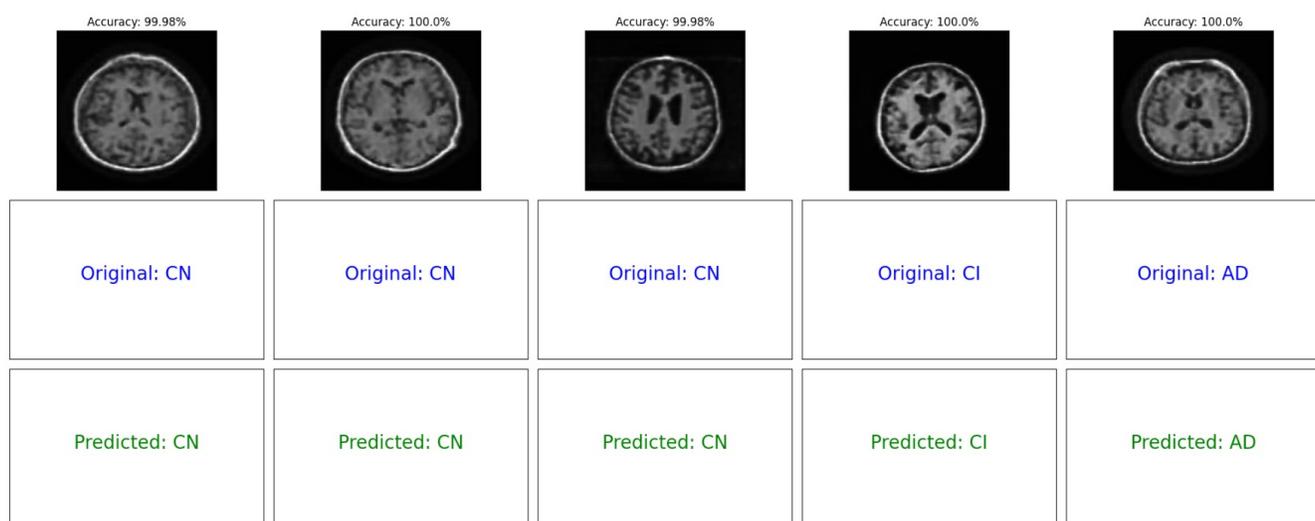


FIGURE 3.13 – Test de l'approche 1

Les résultats démontrent que notre modèle CNN, après un débruitage effectué par un auto-encodeur, atteint des niveaux de précision très élevés, souvent proches de 100 %. La majorité des prédictions sont correctes, comme l'illustrent les correspondances entre les étiquettes originales et les étiquettes prédites, qui sont affichées en vert pour les prédictions exactes et en rouge pour celles

erronées.

Cependant, il y a des cas où des erreurs de classification peuvent se produire malgré une haute précision de prédiction. Cela souligne la nécessité de continuer à optimiser et à valider le modèle pour garantir sa robustesse et sa fiabilité.

3.2.1.3 Discussion

Pour évaluer rigoureusement les performances de notre modèle CNN dans le cadre de l'approche de débruitage suivi de la classification des images TEP, nous avons analysé les courbes de perte et d'exactitude générées pendant les phases d'entraînement et de validation. Ces visualisations dynamiques illustrent la progression du modèle au fil des époques et sa capacité à maintenir une bonne généralisation sur les données de validation, fournissant ainsi des perspectives claires sur l'évolution des performances et la robustesse du modèle pour la classification des images TEP dans la détection de la maladie d'Alzheimer.

Après avoir appliqué le débruitage aux images TEP à l'aide d'un modèle autoencodeur, notre modèle CNN a été évalué pour classer les données en trois catégories distinctes : Maladie d'Alzheimer (AD), Trouble Cognitif Léger (CI), et cas Normaux (CN). Les résultats obtenus indiquent que notre modèle atteint des performances exceptionnelles sur toutes les classes, avec des scores de précision, de rappel (sensibilité) et de F1-score très proches de 1.0. Cela témoigne de l'efficacité du modèle à discriminer avec précision entre les différentes conditions cliniques à partir des images TEP.

En termes de précision, notre modèle a atteint une précision de 100% sur l'ensemble d'entraînement et de 99.81% sur l'ensemble de test, soulignant sa capacité à classer avec une grande précision les images TEP débruitées.

Malgré l'impact du processus de débruitage sur la clarté initiale des images, notre modèle a maintenu des performances élevées en matière de classification. Cette observation souligne l'importance cruciale de trouver un équilibre entre l'amélioration des données par débruitage et la préservation des détails diagnostiquement pertinents. Bien que le débruitage ait introduit des artefacts et réduit la netteté des images, notre modèle a démontré une robustesse notable dans

l'exploitation des caractéristiques essentielles malgré ces perturbations.

En conclusion, l'approche combinée de débruitage suivi de classification avec un modèle CNN présente un potentiel prometteur pour améliorer la précision et la fiabilité du diagnostic de la maladie d'Alzheimer à partir des images TEP. Ces résultats soulignent l'importance continue de l'optimisation des méthodes de prétraitement et de classification pour maximiser l'utilité clinique des données d'imagerie médicale.

3.2.2 Approche 2 : Classification Directe des images sans Débruitage

3.2.2.1 Description de l'Approche et Résultats

Dans cette deuxième approche, nous avons opté pour une classification directe des images TEP avec le même modèle CNN sans prétraitement préalable par débruitage à l'aide d'un autoencodeur. Contrairement à l'approche précédente qui visait à améliorer la qualité des images en réduisant le bruit avant la classification, cette méthode a pour objectif d'évaluer la performance du modèle CNN en utilisant les images TEP brutes, avec un prétraitement limité à la conversion en niveaux de gris et à la normalisation.

— Évaluation des Performances :

- **Courbe de Perte (Entraînement et Validation) :** Dans cette approche, nous observons aussi que la perte diminue progressivement pour les deux ensembles, indiquant que le modèle apprend bien les caractéristiques des images TEP brutes. La convergence de la perte entre l'entraînement et la validation suggère que le modèle ne souffre pas d'overfitting, même sans débruitage préalable. (Voir figure 3.14)

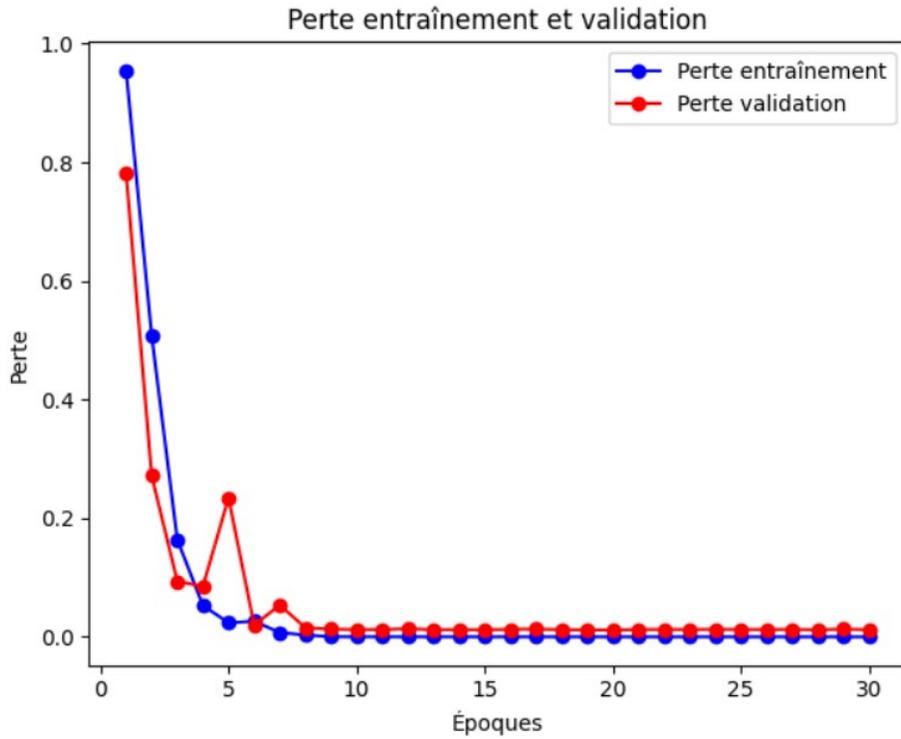


FIGURE 3.14 – Courbe de Perte de l’approche 2

- **Courbe d’Exactitude (Entraînement et Validation)** : La courbe d’exactitude montre que le modèle atteint des niveaux de précision élevés, similaires à ceux observés avec la première approche. Les courbes pour l’entraînement et la validation suivent une tendance ascendante, indiquant une amélioration continue des performances au fil des époques. (Voir figure 3.15)

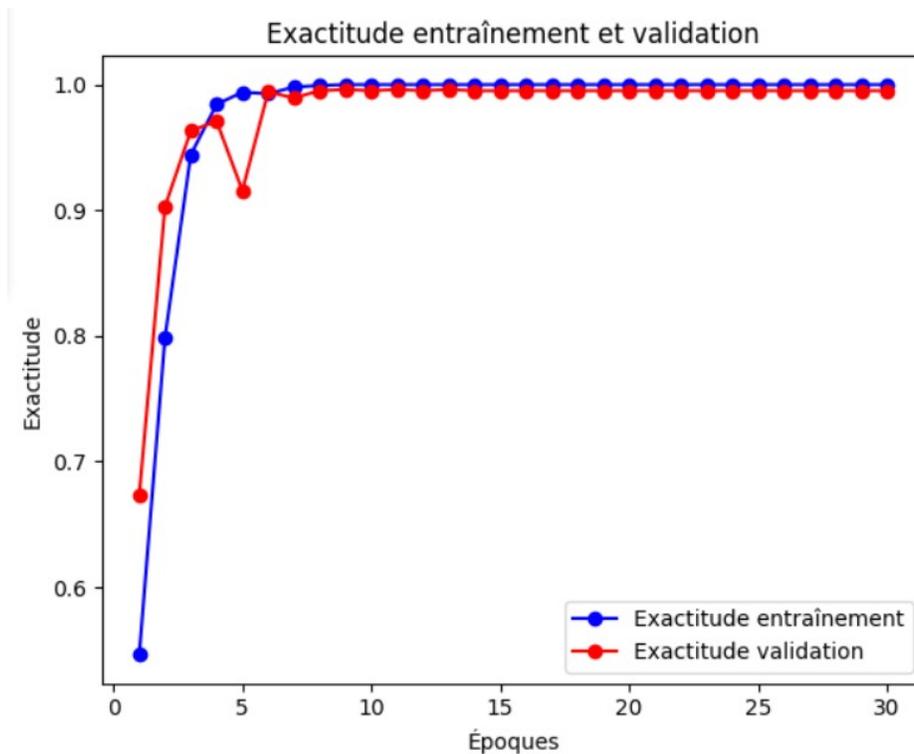


FIGURE 3.15 – Courbe de Précision de l’approche 2

- **Précision** : Le modèle a atteint une précision de 100% sur l’ensemble d’entraînement et de 99,52% sur l’ensemble de test. Ces résultats démontrent que, même sans débruitage, le modèle est capable de classer les images TEP avec une très haute précision. Cependant, cette légère différence entre les ensembles d’entraînement et de test pourrait indiquer que le modèle est bien généralisable.
- **Matrice de confusion** : La matrices de confusion du test fournit une visualisation des résultats de classification du modèle et elle est essentielle pour évaluer les performances du modèle de manière détaillée. (Voir figure 3.16)

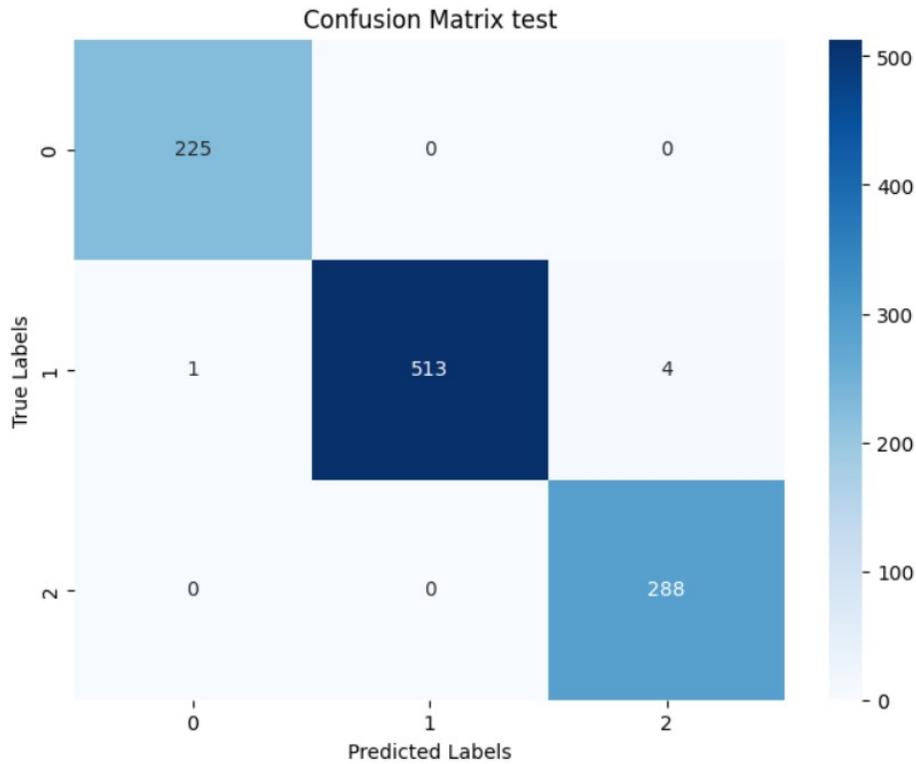


FIGURE 3.16 – Matrice de confusion du test de l’approche 2

— Résultats par Classe :

Classe	True Positives	True Negatives	False Positives	False Negatives
AD	225	805	1	0
CI	513	513	0	5
CN	288	739	4	0

TABLE 3.3 – Interpretation de la Matrice de Confusion de l’approche 2

Ce tableau présente l’interprétation de la matrice de confusion pour la deuxième approche.

Maladie d’Alzheimer (AD) : Le modèle a correctement identifié 225 images comme appartenant à la classe AD (True Positives, TP) et a correctement rejeté 805 images comme n’appartenant pas à la classe AD (True Negatives, TN). Il y a eu 1 faux positif (False Positive, FP), c’est-à-dire une image incorrectement identifiée

comme AD, et aucune image n'a été manquée (False Negative, FN).

Trouble Cognitif Léger (CI) : Le modèle a correctement identifié 513 images comme CI (TP) et 513 images comme n'étant pas CI (TN). Il n'y a eu aucun faux positif (FP), mais 5 images n'ont pas été reconnues comme CI alors qu'elles l'étaient (FN).

Cas Normaux (CN) : Le modèle a correctement identifié 288 images comme CN (TP) et 739 images comme n'étant pas CN (TN). Il y a eu 4 images incorrectement identifiées comme CN (FP) et aucune image n'a été manquée (FN).

Classe	Précision	Rappel (Sensibilité)	F1-Score
AD	0.9955	1.0	0.9977
CI	1.0	0.9903	0.9951
CN	0.9863	1.0	0.9931

TABLE 3.4 – Métriques de Performance de l'approche 2

Ce deuxième tableau présente les métriques de performance : précision, rappel (sensibilité) et F1-score pour chaque classe.

Maladie d'Alzheimer (AD) : La précision est de 0.9955, ce qui signifie que 99.55% des images classées comme AD étaient correctes. Le rappel est de 1.0, indiquant que le modèle a identifié 100% des images réellement AD. Le F1-score est de 0.9977, montrant un équilibre élevé entre précision et rappel.

Trouble Cognitif Léger (CI) : La précision est de 1.0, ce qui signifie que 100% des images classées comme CI étaient correctes. Le rappel est de 0.9903, indiquant que 99.03% des images réellement CI ont été correctement identifiées. Le F1-score est de 0.9951, reflétant une performance équilibrée.

Cas Normaux (CN) : La précision est de 0.9863, signifiant que 98.63% des

images classées comme CN étaient correctes. Le rappel est de 1.0, indiquant que toutes les images réellement CN ont été correctement identifiées. Le F1-score est de 0.9931, démontrant une très bonne performance globale.

3.2.2.2 Test

Ici, nous présentons les résultats de la classification de quelques images, issues de l'ensemble d'images TEP non débruitées, par notre modèle CNN. Chaque image est accompagnée de son étiquette originale et de la classe prédite par le modèle, ainsi que de la précision de cette prédiction.

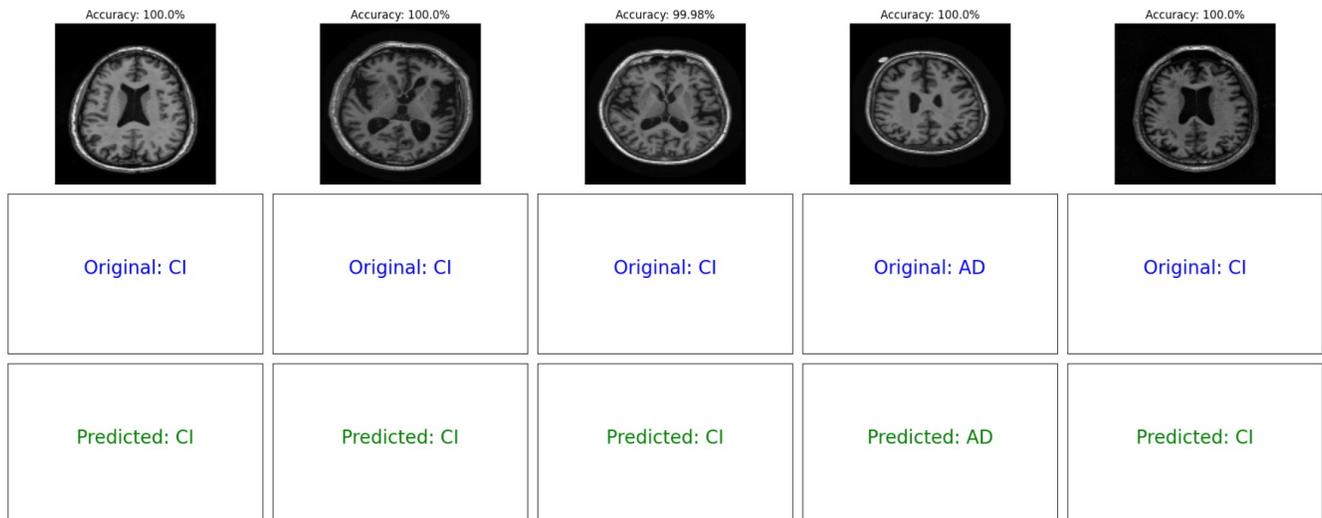


FIGURE 3.17 – Test de l'approche 2

Les résultats démontrent que notre modèle CNN atteint des niveaux de précision très élevés, souvent proches de 100%. La majorité des prédictions sont correctes, comme l'illustrent les correspondances entre les étiquettes originales et les étiquettes prédites.

Bien que des erreurs de classification puissent survenir malgré une haute précision de prédiction, ces résultats illustrent le potentiel de notre modèle pour améliorer la détection et le suivi de la maladie d'Alzheimer à partir des images TEP non débruitées.

3.2.2.3 Discussion

Dans cette deuxième approche de classification directe des images TEP sans prétraitement préalable par débruitage, notre modèle CNN a démontré des performances significatives et des résultats encourageants pour la détection de la maladie d'Alzheimer.

Les résultats révèlent que le modèle a atteint une précision de 100% sur l'ensemble d'entraînement et de 99,52% sur l'ensemble de test. Ces chiffres témoignent de la capacité du modèle à classifier avec une très haute précision les images TEP brutes, confirmant son efficacité même sans traitement préalable pour réduire le bruit.

La courbe de perte montre une diminution progressive et similaire pour les ensembles d'entraînement et de validation, indiquant que le modèle apprend efficacement à partir des caractéristiques intrinsèques des images TEP brutes. Cette convergence suggère également que le modèle ne présente pas de surapprentissage, ce qui est crucial pour sa généralisabilité dans des contextes cliniques réels.

L'analyse détaillée de la matrice de confusion révèle une capacité élevée du modèle à classifier correctement les trois classes (AD, CI, CN). Par exemple, pour la classe AD, le modèle a identifié correctement 225 images comme AD, avec seulement 1 faux positif. De même, pour les classes CI et CN, les performances en termes de précision, de rappel et de F1-score montrent une robustesse significative du modèle dans l'identification des cas positifs et négatifs.

Ces résultats soulignent l'efficacité du modèle CNN dans la classification directe des images TEP brutes pour la détection précoce et le suivi de la maladie d'Alzheimer. Ils ouvrent la voie à des applications cliniques potentielles où une haute précision et une fiabilité sont essentielles pour un diagnostic précis et efficace.

En conclusion, cette approche sans débruitage préalable confirme la pertinence et l'efficacité de notre modèle dans le domaine de l'imagerie médicale pour la détection précoce et la surveillance de la maladie d'Alzheimer à partir d'images TEP brutes. Ces résultats encouragent la poursuite de la validation et de l'optimisation du modèle pour répondre aux exigences rigoureuses des environnements cliniques réels.

3.2.3 Explicabilité du Modèle

L'explicabilité du modèle joue un rôle crucial dans la validation et l'adoption de modèles d'IA dans des applications médicales telles que la détection précoce de la maladie d'Alzheimer à partir d'images TEP. Cette section explore comment notre modèle CNN peut être interprété à l'aide de techniques d'explicabilité, telles que LIME (Local Interpretable Model-agnostic Explanations), pour identifier les zones d'intérêt et les caractéristiques importantes utilisées pour la classification.

3.2.3.1 Utilisation de LIME pour l'Explicabilité

L'application de LIME à notre modèle de classification basé sur les images TEP pour la détection de la maladie d'Alzheimer a révélé des informations significatives sur le fonctionnement du modèle et son interprétabilité. Les étapes clés sont les suivantes :

— **Génération de Superpixels :**

Après avoir chargé l'image, on utilise l'algorithme quickshift pour générer des superpixels. Cet algorithme (quickshift) segmente l'image en régions homogènes, ce qui permet une analyse plus fine des zones d'intérêt sur l'image.

En utilisant les superpixels générés, nous pouvons visualiser les frontières de segmentation marquées sur l'image d'origine, facilitant ainsi l'interprétation des résultats de notre modèle de manière visuelle et détaillée.

— **Perturbations et Interprétations :**

Lors de la phase de perturbations et d'interprétations, on crée plusieurs variations de l'image en masquant différents superpixels. Chaque superpixel représente une région homogène de l'image, et en les masquant sélectivement, on évalue l'impact de chaque région sur la décision de classification du modèle. Les prédictions associées à ces perturbations aident à estimer l'importance relative des superpixels. Plus le nombre de perturbations est élevé, plus les explications fournies sont fiables. (Voir figure 3.18)

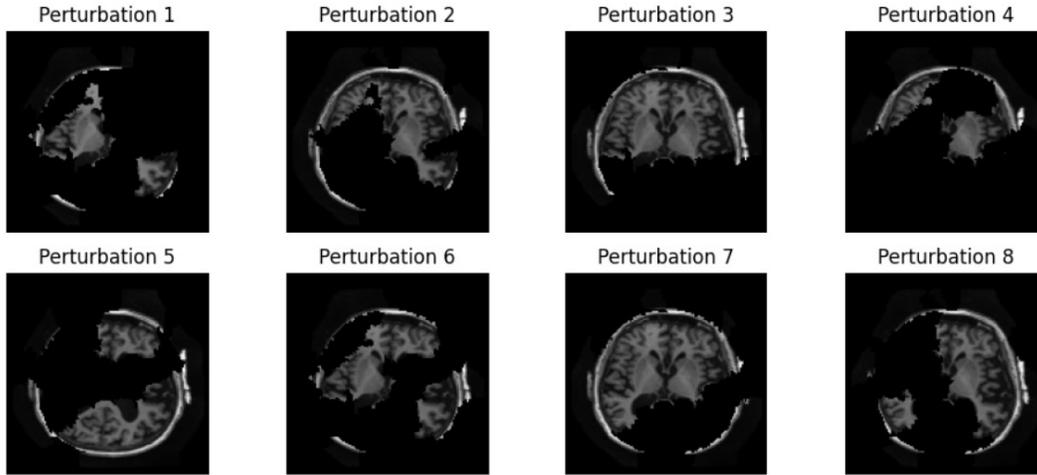


FIGURE 3.18 – Perturbations

— **Modèle Linéaire pour l’Explication :**

Lorsque nous utilisons un modèle linéaire pour l’explication, nous ajustons ce modèle aux prédictions du modèle principal pour chaque perturbation de l’image. Chaque perturbation modifie l’image en masquant certains superpixels, permettant ainsi d’évaluer leur impact sur la prédiction. En utilisant la distance cosinus comme poids, nous attribuons des scores d’importance aux superpixels lors de cet ajustement. Cette approche quantifie la contribution de chaque région de l’image à la décision de classification finale. Les formules mathématiques utilisées sont les suivantes :

— **Calcul des distances cosinus :**

Pour chaque perturbation \mathbf{P}_j et l’image originale \mathbf{X}_i , la distance cosinus est calculée comme suit :

$$\text{distance}_{\text{cosinus}}(\mathbf{P}_j, \mathbf{X}_i) = 1 - \frac{\mathbf{P}_j \cdot \mathbf{X}_i}{\|\mathbf{P}_j\| \|\mathbf{X}_i\|} \quad (3.1)$$

où :

- $\mathbf{P}_j \cdot \mathbf{X}_i$ représente le produit scalaire entre les vecteurs \mathbf{P}_j et \mathbf{X}_i ,
- $\|\mathbf{P}_j\|$ et $\|\mathbf{X}_i\|$ représentent les normes de \mathbf{P}_j et \mathbf{X}_i respectivement.

— **Transformation des distances en poids :**

Les distances cosinus calculées sont ensuite transformées en poids w_j comme suit :

$$w_j = \exp \left(- \left(\frac{\text{distance}_{\text{cosinus}}(\mathbf{P}_j, \mathbf{X}_i)}{\sigma} \right)^2 \right) \quad (3.2)$$

où :

- $\text{distance}_{\text{cosinus}}(\mathbf{P}_j, \mathbf{X}_i)$ est la distance cosinus entre la perturbation \mathbf{P}_j et l'image originale \mathbf{X}_i ,
- σ est la largeur du noyau (*kernel_width*).

— **Visualisation des Caractéristiques Importantes :**

Une fois les superpixels les plus influents identifiés à l'aide du modèle linéaire, on les met en évidence visuellement sur l'image d'origine. Cette étape de visualisation est cruciale car elle permet de comprendre quelles caractéristiques spécifiques de l'image ont conduit le modèle à prendre sa décision de classification. En marquant ces superpixels sur l'image, nous offrons une représentation graphique intuitive des zones qui ont le plus contribué à l'analyse du modèle. (Voir figure 3.19)

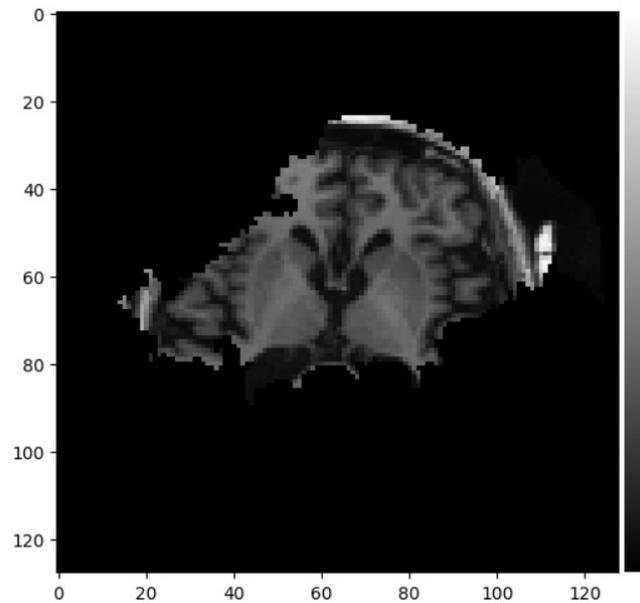


FIGURE 3.19 – Superpixels

3.2.3.2 Discussion

L'application de LIME à notre modèle de classification TEP pour la détection de la maladie d'Alzheimer a révélé des informations importantes sur son fonctionnement. La segmentation en superpixels avec quickshift a permis d'identifier les régions de l'image influençant la décision du modèle, alignées avec les attentes cliniques. En créant des perturbations par masquage des superpixels, nous avons évalué leur impact sur la classification, ce qui a révélé que certaines régions ont un effet significatif sur la prédiction.

L'ajustement d'un modèle linéaire aux prédictions perturbées, en utilisant la distance cosinus comme poids, a permis de quantifier l'importance relative des superpixels. Cette méthode a rendu les décisions du modèle plus transparentes pour les experts médicaux. La visualisation des superpixels influents a facilité la compréhension des caractéristiques importantes pour la classification.

Toutefois, la qualité des explications dépend des paramètres de segmentation et du nombre de perturbations. LIME a été efficace mais doit être complété par d'autres méthodes d'explicabilité pour une compréhension plus complète du modèle. En somme, LIME a amélioré la transparence et la compréhension des décisions du modèle, validant des hypothèses cliniques et suggérant des pistes pour des recherches futures en explicabilité des IA médicales.

3.3 Comparaison avec d'autres résultats

Afin de démontrer l'importance de notre approche, il est essentiel de comparer nos résultats avec ceux obtenus par d'autres études similaires dans le domaine de la classification des images TEP pour la détection de la maladie d'Alzheimer. Cette comparaison permet de situer notre modèle par rapport à l'état de l'art et d'évaluer sa performance relative.

3.3.1 Études comparatives

Nous avons identifié deux études récentes qui se concentrent sur des approches similaires utilisant l'apprentissage profond pour la classification des images TEP :

- Une étude de Cheng et Liu (2017), intitulée "**Classification of Alzheimer's Disease by Cascaded Convolutional Neural Networks Using PET Images**", rapporte une précision de 92.2% [40]. Cette étude utilise des réseaux de neurones convolutifs en cascade pour classifier les images TEP afin de détecter la maladie d'Alzheimer. Les auteurs ont développé un modèle CNN en plusieurs étapes pour améliorer la précision de la classification. Leur approche, bien qu'efficace, reste inférieure en termes de précision comparée à notre approche.
- Une autre étude de Amini et al. (2014), intitulée "**GC-CNNnet : Diagnosis of Alzheimer's Disease with PET Images Using Genetic and Convolutional Neural Network**", rapporte une précision de 91.1% [41]. Cette étude propose une méthode combinant des réseaux de neurones convolutifs et des algorithmes génétiques pour diagnostiquer la maladie d'Alzheimer à partir des images TEP. Le modèle GC-CNNnet a été conçu pour optimiser les paramètres du réseau à l'aide d'algorithmes génétiques afin d'améliorer les performances. Malgré cette approche innovante, la précision obtenue reste en deçà de celles de notre modèle.

3.3.2 Discussion

Les résultats obtenus montrent que nos approches de classification dépassent les performances rapportées dans les études comparatives en termes de précision. Nos approches atteignent des précisions de 99.71% et 99.52%, ce qui est significativement supérieur aux précisions de 92.2% et 91.1% rapportées respectivement par les études de Cheng et Liu et de Amini et al.

Cette performance accrue peut être attribuée à la robustesse de notre modèle CNN et à l'efficacité de notre pipeline de prétraitement des données. La comparaison avec d'autres résultats souligne l'importance de notre travail et justifie l'adoption de notre méthode pour une application clinique. Nos résultats suggèrent que nos approches permettent une classification plus précise et fiable, ce qui est crucial pour le diagnostic précoce et le suivi de la maladie d'Alzheimer.

3.4 Conclusion

En conclusion, cette étude a exploré deux approches distinctes pour la détection de la maladie d'Alzheimer à partir d'images TEP, en mettant en œuvre un modèle CNN. La première approche a intégré un prétraitement par débruitage avec un autoencodeur, visant à améliorer la qualité des images avant la classification. La seconde approche a adopté une classification directe des images brutes, sans prétraitement préalable, montrant une capacité remarquable à atteindre des niveaux élevés de précision sur les ensembles d'entraînement et de test.

Les résultats ont démontré que même sans débruitage, le modèle CNN était capable de classifier efficacement les images TEP, avec des performances comparables voire supérieures dans certains aspects à celles de l'approche avec débruitage. La convergence des courbes de perte et d'exactitude, ainsi que l'analyse détaillée des matrices de confusion et des métriques de performance, ont confirmé la robustesse et la généralisabilité du modèle dans des contextes cliniques réels.

En outre, l'application de techniques d'explicabilité comme LIME a renforcé la transparence des décisions du modèle, permettant une meilleure interprétation des caractéristiques importantes utilisées pour la classification.

Ces résultats encouragent la poursuite des efforts pour optimiser et valider ces modèles dans des environnements cliniques, ouvrant la voie à des applications potentielles pour améliorer la détection précoce et le suivi de la maladie d'Alzheimer.

Conclusion Générale et Perspectives

Ce mémoire a exploré des approches novatrices pour la détection et le suivi de la maladie d'Alzheimer à partir d'images TEP, en mettant en œuvre des modèles de réseaux de neurones convolutifs (CNN) et des techniques d'intelligence artificielle explicable (XAI). Nous avons développé et comparé deux approches distinctes : une première intégrant un prétraitement par débruitage des images avec un autoencodeur, et une seconde utilisant directement les images TEP, sans prétraitement par débruitage, pour la classification.

Les résultats obtenus ont montré que les deux approches offrent des performances de classification remarquables, avec une légère différence mais une précision presque équivalente. Cette observation souligne la robustesse et la capacité des CNN à extraire des caractéristiques discriminantes même à partir de données brutes.

L'incorporation de techniques d'explicabilité, notamment LIME, a permis d'améliorer la transparence des décisions du modèle. En rendant visibles les superpixels influents, LIME a offert des explications visuelles claires, facilitant ainsi l'interprétation des caractéristiques importantes pour la classification. Cette transparence est cruciale dans un contexte clinique, car elle renforce la confiance des professionnels de santé dans les prédictions du modèle et ouvre des pistes pour de futures améliorations et validations.

Ce travail a démontré l'efficacité et la robustesse des modèles CNN pour la détection précoce de la maladie d'Alzheimer à partir d'images TEP, tout en soulignant l'importance des techniques d'explicabilité pour une adoption clinique. Les résultats encourageants obtenus incitent à poursuivre les efforts d'optimisation et de validation dans des environnements cliniques réels, afin de maximiser l'impact positif de ces technologies sur la gestion et le suivi de la maladie d'Alzheimer.

Les perspectives de ce travail sont nombreuses et prometteuses. Tout d'abord, il serait bénéfique d'élargir le jeu de données en incluant des images provenant de diverses sources et populations pour mieux vérifier la généralisation du modèle. L'amélioration continue des algorithmes de débruitage et de classification pourrait également améliorer les performances et la précision.

Ensuite, intégrer des approches multimodales en combinant les données TEP avec d'autres types d'imagerie médicale ou des données cliniques pourrait rendre les modèles encore plus robustes et précis. La validation du modèle dans des contextes cliniques réels et l'évaluation de son impact sur les pratiques médicales seraient cruciales pour son adoption.

Enfin, l'implémentation et l'évaluation d'autres techniques d'explicabilité aideraient à mieux comprendre les décisions des modèles, ce qui améliorerait leur acceptation par les utilisateurs. La collaboration avec les professionnels de santé sera essentielle pour s'assurer que les solutions répondent aux besoins médicaux.

Ces perspectives ouvrent de nombreuses opportunités pour la recherche future et l'amélioration des technologies d'IA pour la détection et le suivi de la maladie d'Alzheimer.

Bibliographie

- [1] Inserm. Maladie d'alzheimer. une maladie neurodégénérative complexe mais de mieux en mieux comprise. 2017.
- [2] J Freymond. *Maladie d'Alzheimer : Etude des marqueurs pathologiques dans le cortex frontal humain et corrélations anatomo-cliniques*. PhD thesis, Université de Lausanne, Faculté de biologie et médecine, 2012.
- [3] Agneta Nordberg, Juha O Rinne, Ahmadul Kadir, and Bengt Långström. The use of pet in alzheimer disease. *Nature Reviews Neurology*, 6(2) :78–87, 2010.
- [4] VIDAL. Tomographie par émission de positons (tep ou petscan). 2022.
- [5] Luc Zimmer. L'imagerie tep pour une meilleure compréhension de la neurotransmission normale et pathologique. *Biologie Aujourd'hui*, 213(3-4) :109–120, 2019.
- [6] Cap Retraite. Alzheimer : un pet scan pour prédire l'évolution de la maladie. 2020.
- [7] Fondation Vaincre Alzheimer. Diagnostic précoce alzheimer : pourquoi est-il si important ? 2022.
- [8] Ali Amad, Aïda Cancel, and Thomas Fovet. L'imagerie cérébrale en psychiatrie clinique : du diagnostic différentiel au machine learning. *L'Information psychiatrique*, (2) :277–284, 2016.
- [9] Jedha. Algorithmes de machine learning : présentation et fonctionnement. 2024.
- [10] RabehAmira Ben, Benzarti Faouzi, and Amiri Hamid. Diagnosis of alzheimer diseases in early step using svm (support vector machine). In *IEEE 13th International Conference on, Computer Graphics, Imaging and Visualization (CGiV)*, pages 364–367, 2016.
- [11] RA Saputra, C Agustina, D Puspitasari, R Ramanda, D Pribadi, K Indriani, et al. Detecting alzheimer's disease by the decision tree methods based on particle swarm optimization. In *Journal of Physics : Conference Series*, volume 1641, page 012025. IOP Publishing, 2020.

- [12] PJ Moore, TJ Lyons, John Gallacher, and Alzheimer’s Disease Neuroimaging Initiative. Random forest prediction of alzheimer’s disease using pairwise selection from time series data. *PloS one*, 14(2) :e0211558, 2019.
- [13] Thema Radiologie. Détecter précocement la maladie d’alzheimer grâce au deep learning des données de pet-fdg. 2018.
- [14] Yiming Ding, Jae Ho Sohn, Michael G Kawczynski, Hari Trivedi, Roy Harnish, Nathaniel W Jenkins, Dmytro Lituiev, Timothy P Copeland, Mariam S Aboian, Carina Mari Aparici, et al. A deep learning model to predict a diagnosis of alzheimer disease by using 18f-fdg pet of the brain. *Radiology*, 290(2) :456–464, 2019.
- [15] Valeuriad. Qu’est-ce qu’un réseau de neurones convolutifs ? 2021.
- [16] IBM. Qu’est-ce qu’un réseau de neurones convolutifs ? 2024.
- [17] Marc Mercier. Qu’est-ce qu’un réseau neuronal profond ? 2022.
- [18] Heung-Il Suk and Dinggang Shen. Deep learning-based feature representation for ad/mci classification. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013 : 16th International Conference, Nagoya, Japan, September 22–26, 2013, Proceedings, Part II 16*, pages 583–590. Springer, 2013.
- [19] Adrien Payan and Giovanni Montana. Predicting alzheimer’s disease : a neuroimaging study with 3d convolutional neural networks. *arXiv preprint arXiv :1502.02506*, 2015.
- [20] Heung-Il Suk, Seong-Whan Lee, Dinggang Shen, Alzheimer’s Disease Neuroimaging Initiative, et al. Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. *NeuroImage*, 101 :569–582, 2014.
- [21] Lovedeep Gondara. Medical image denoising using convolutional denoising autoencoders. In *2016 IEEE 16th international conference on data mining workshops (ICDMW)*, pages 241–246. IEEE, 2016.
- [22] Inam Rafia, Shorov Andrey, Ustundag Soykan Elif, Reno Jim, and Berlanga Raquel. Quatre avantages de l’ia pour la sécurité, la sûreté et la transparence dans les télécommunications. 2024.
- [23] The Royal Society. Explainable ai. 2024.
- [24] Le Pont. Tout savoir sur l’xai ou l’intelligence artificielle explicable. 2024.

- [25] Andreas Holzinger, Georg Langs, Helmut Denk, Kurt Zatloukal, and Heimo Müller. Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews : Data Mining and Knowledge Discovery*, 9(4) :e1312, 2019.
- [26] Erico Tjoa and Cuntai Guan. A survey on explainable artificial intelligence (xai) : Toward medical xai. *IEEE transactions on neural networks and learning systems*, 32(11) :4793–4813, 2020.
- [27] Ouren Kuiper, Martin van den Berg, Joost van der Burgt, and Stefan Leijnen. Exploring explainable ai in the financial sector : Perspectives of banks and supervisory authorities. In *Artificial Intelligence and Machine Learning : 33rd Benelux Conference on Artificial Intelligence, BNAIC/Benelearn 2021, Esch-sur-Alzette, Luxembourg, November 10–12, 2021, Revised Selected Papers 33*, pages 105–119. Springer, 2022.
- [28] Eric Benhamou, Jean-Jacques Ohana, David Saltiel, Beatrice Guez, and Steve Ohana. Explainable ai (xai) models applied to planning in financial markets. *SSRN Electron. J. doi*, 10, 2021.
- [29] James M Anderson, Kalra Nidhi, Karlyn D Stanley, Paul Sorensen, Constantine Samaras, and Oluwatobi A Oluwatola. *Autonomous vehicle technology : A guide for policymakers*. Rand Corporation, 2014.
- [30] Or Biran and Courtenay Cotton. Explanation and justification in machine learning : A survey. In *IJCAI-17 workshop on explainable AI (XAI)*, volume 8, pages 8–13, 2017.
- [31] Cynthia Rudin and Berk Ustun. Optimized scoring systems : Toward trust in machine learning for healthcare and criminal justice. *Interfaces*, 48(5) :449–466, 2018.
- [32] Gunning D Aha DW. Darpa’s explainable artificial intelligence program. *AI Mag*, 40(2) :44, 2019.
- [33] University of Washington. Marco tulio ribeiro bibliography.
- [34] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. " why should i trust you ?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2016.
- [35] V Trevisan. Using shap values to explain how your machine learning model works. *Medium*, 2022.

- [36] Florentin Jiechieu and Norbert Tsopze. Une approche basée sur la méthode lrp pour l'explication des réseaux de neurones convolutifs appliqués à la classification des textes. In *CARI 2022*, 2022.
- [37] Kaggle. "adni extracted axial. In *Alzheimer's Disease Multiclass imaging data*, 2021.
- [38] Sara Samuel, Michael Moore, Brandon Patterson, Helenmary Sheridan, and Chris Sorensen. Neuroimaging data primer : A resource for curating digital imaging and communications in medicine (dicom) and neuroimaging informatics technology initiative (nifti) files. 2020.
- [39] Ronald Carl Petersen, Paul S Aisen, Laurel A Beckett, Michael C Donohue, Anthony Collins Gamst, Danielle J Harvey, CR Jack Jr, William J Jagust, Leslie M Shaw, Arthur W Toga, et al. Alzheimer's disease neuroimaging initiative (adni) clinical characterization. *Neurology*, 74(3) :201–209, 2010.
- [40] Danni Cheng and Manhua Liu. Classification of alzheimer's disease by cascaded convolutional neural networks using pet images. In *Machine Learning in Medical Imaging : 8th International Workshop, MLMI 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 10, 2017, Proceedings 8*, pages 106–113. Springer, 2017.
- [41] Morteza Amini, Mir Mohsen Pedram, AliReza Moradi, Mahdieh Jamshidi, and Mahshad Ouchani. Gc-cnnnet : Diagnosis of alzheimer's disease with pet images using genetic and convolutional neural network. *Computational Intelligence and Neuroscience*, 2022(1) :7413081, 2022.

Annexe

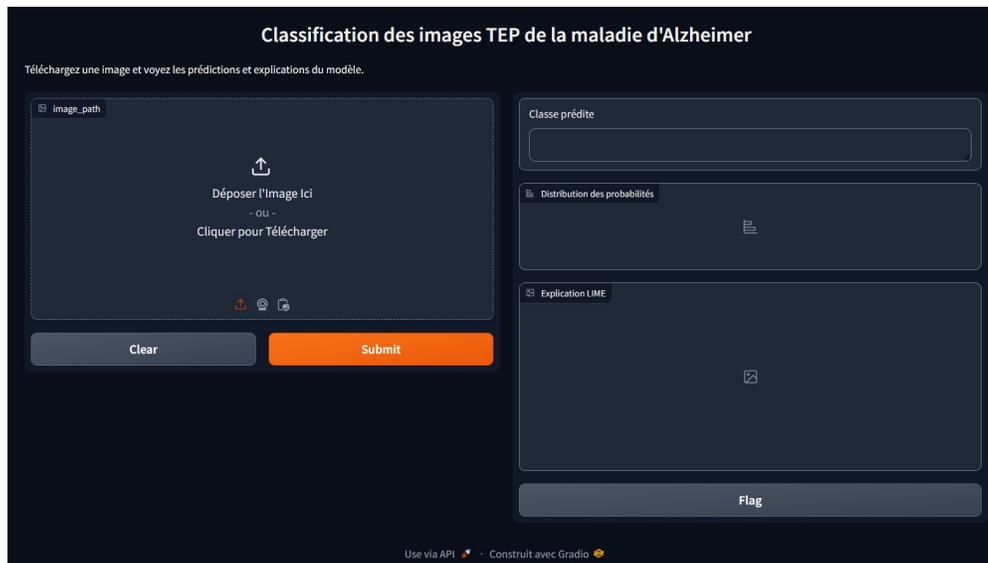


FIGURE 3.20 – Interface Graphique

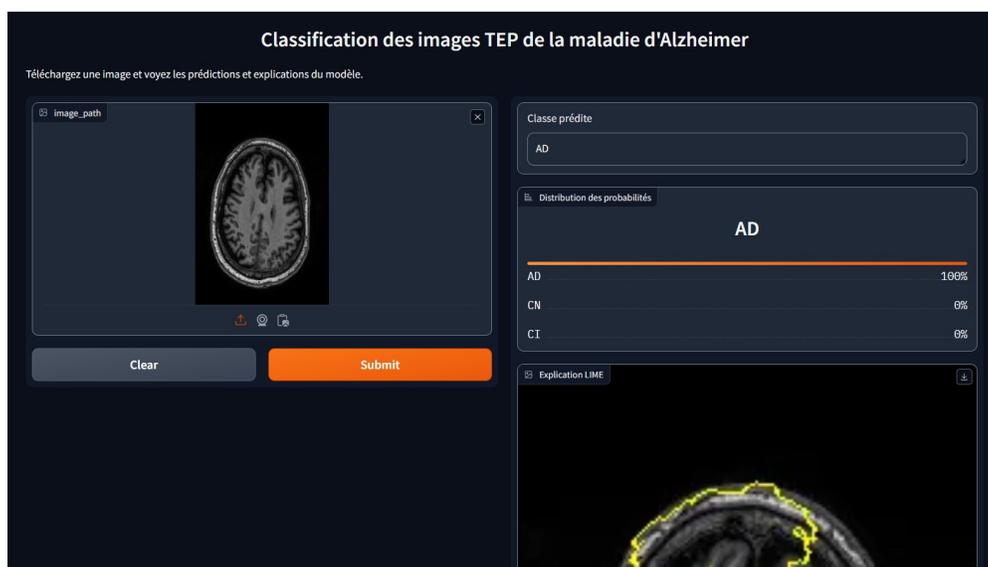


FIGURE 3.21 – Test avec l'interface graphique