

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة ابوبكر بلقايد - تلمسان

Université Aboubakr Belkaïd – Tlemcen – Faculté de TECHNOLOGIE



MEMOIRE

Présenté pour l'obtention du **diplôme** de **MASTER**

En : Génie Biomédical

Spécialité : Imagerie Médicale

Par : Zenatri Anfel & Kherroubi Hizia.

Sujet

Système Intelligent de Classification des pathologies médicales par Deep Learning Avancé Intégrant l'Explicabilité et l'Analyse Statistique

Soutenu publiquement, le 18/06/2026, devant le jury composé de :

Mr MESSADI Mahammed	Pr	Université de Tlemcen	Président
Mme Lazouni Sihem	MCB	Université de Tlemcen	Examinatrice
Mme BENSMAIL Ilham	MCB	Université de Tlemcen	Encadrante
Mlle ELAOUABER Zineb Aziza	Dr	Université de Tlemcen	Co-Encadrante

Année universitaire : 2025 /2026

Remerciements

Avant toute chose, nous rendons grâce à Allah le Tout-Puissant pour nous avoir accordé la force, la persévérance et les capacités nécessaires afin de mener à bien ce travail. Sa bénédiction nous a accompagnés tout au long de ce parcours.

Nous souhaitons exprimer notre profonde gratitude à Madame **BENSMAIL Ilham**, notre encadrante, pour son accompagnement constant, son écoute attentive et la confiance qu'elle nous a témoignée. Ses orientations éclairées, ses remarques constructives et son engagement ont constitué un soutien précieux dans l'élaboration de ce mémoire.

Nos remerciements les plus sincères s'adressent également à Madame **Zineb Aziza ELAOUABER**, notre co-encadrante, pour son implication, sa disponibilité et la richesse de ses conseils. Son expertise scientifique, sa rigueur et son soutien permanent nous ont permis d'avancer avec sérénité et confiance dans la réalisation de ce projet.

Nous sommes particulièrement honorés par l'intérêt porté à notre travail par les membres du jury, **Pr. MESSADI Mahammed** et **Dr. BENSEGHIR Sihem**. Nous les remercions chaleureusement d'avoir accepté d'évaluer ce mémoire et de contribuer, par leurs observations et recommandations, à son enrichissement.

Nous adressons également nos remerciements à l'ensemble des enseignants et responsables pédagogiques de l'Université de Tlemcen. Grâce à la qualité de leur enseignement, à leur dévouement et au partage de leurs connaissances, nous avons pu acquérir les compétences nécessaires à notre formation académique et professionnelle.

Enfin, nous exprimons notre reconnaissance à toutes les personnes qui, par leur soutien moral, leurs encouragements ou leur aide, ont contribué de près ou de loin à l'aboutissement de ce travail. Qu'elles trouvent ici l'expression de notre sincère gratitude.

Dédicaces

﴿وَأَخِرُ دَعْوَاهُمْ أَنْ الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ﴾

الْحَمْدُ لِلَّهِ الَّذِي مَا تَمَّ جُهْدٌ وَلَا خُتِمَ سَعْيٌ إِلَّا بِفَضْلِهِ، فَالْحَمْدُ لِلَّهِ عَلَى الْبُلُوغِ، ثُمَّ الْحَمْدُ لِلَّهِ عَلَى التَّكْمَلِ

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

A mon cher père en signe de ma profonde reconnaissance pour sa veille sur mon éducation, ses encouragements et pour m'avoir permis de réaliser mes études dans les meilleures conditions.

A ma chère mère, Un merci infini et particulier. Ton amour, ta tendresse et tes prières ont été le cœur battant de mes réussites. Tu es mon refuge, ma lumière, et la plus belle bénédiction que la vie m'ait offerte. Que Dieu te protège et te garde pour nous.

A toute ma famille, avec qui je partage les bons et difficiles moments de ma vie. Pour son soutien moral tout au long de mon travail

Mes frères et sœurs : Feriel, Latifa, Ikbel, Seif Eddine, Fidae Eddine

A mes amis avec qui nous avons partagé pleins de belles aventures tout au long de nos 5 années d'études universitaires.

À ma chère binôme Hizia, pour sa compréhension, sa patience, son esprit d'équipe et son soutien moral tout au long de ce projet.

À toutes les personnes que j'aime et qui m'aiment, merci d'être dans ma vie.

Enfin, et surtout... Je veux me remercier. Je veux me remercier d'avoir cru en moi. Je veux me remercier d'avoir accompli tout ce travail. Je veux me remercier de n'avoir jamais abandonné.

Anfel

Dédicaces

À l'issue de ce parcours, je ressens le besoin profond de dédier ce travail à celles et ceux qui ont été ma lumière dans les moments d'obscurité, ma force dans les instants de doute, et mon repère tout semblait incertain.

Je dédie tout d'abord ce mémoire à ma famille :

À ma mère, Un merci infini et particulier. Ton amour, ta tendresse et tes prières ont été le cœur battant de mes réussites. Tu es mon refuge, ma lumière, et la plus belle bénédiction que la vie m'ait offerte. Que Dieu te protège et te garde pour nous.

À mon cher papa,

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

À mes frères, Abdel Kader, Abdel Aziz, Ahmed, Ayoub, Sadik et balmhel, merci pour votre amour, votre soutien et vos encouragements continus. Merci d'être toujours là pour moi, dans les moments de doute comme dans les moments de réussite.

À Anfel, ma binôme,

Merci pour ta collaboration sincère, ton sérieux, ton soutien constant et ta bonne humeur. Nous avons traversé cette aventure ensemble, et ton implication a été précieuse du début jusqu'à la fin.

À mes amis,

Pour leur soutien indéfectible, leur encouragement et leur bonne humeur. Merci d'avoir partagé avec moi les joies et les défis de ce parcours, et d'avoir toujours cru en moi.

À moi-même,

Pour avoir persévéré malgré les obstacles, pour être restée fidèle à mes objectifs, et pour ne pas avoir baissé les bras. Cette réussite, je me la dois aussi. Merci à celle que je suis devenue.

Hizia

Résumé

L'Accident Vasculaire Cérébral (AVC) constitue une urgence neurologique majeure dont le diagnostic rapide est essentiel pour améliorer la prise en charge des patients et réduire les risques de séquelles permanentes. Dans ce travail, nous proposons un système intelligent d'aide au diagnostic des AVC à partir d'images de scanner CT cérébral basé sur les techniques d'apprentissage profond. Quatre architectures pré-entraînées, à savoir EfficientNet-B0, MobileNet-V2, ViT-B/16 et Swin-T, ont été développées et comparées dans différents scénarios de classification binaire et multiclassés visant à distinguer les cas normaux, ischémiques et hémorragiques. Afin d'améliorer les performances des modèles, plusieurs étapes de prétraitement ont été appliquées, incluant le nettoyage des images, la normalisation, l'augmentation de données et le transfert d'apprentissage. Les performances ont été évaluées à l'aide de métriques standard telles que l'Accuracy, la Sensibilité, la Spécificité, le F1-score et l'AUC, puis validées statistiquement à travers les tests de McNemar, Bootstrap et DeLong. L'interprétabilité des prédictions a été assurée grâce à la méthode Grad-CAM, permettant de visualiser les régions anatomiques ayant le plus contribué aux décisions du modèle et de vérifier leur cohérence clinique. En outre, une validation externe réalisée sur la base de données RSNA a permis d'évaluer la capacité de généralisation des modèles face à un changement de domaine. Les résultats obtenus montrent des performances élevées, une bonne robustesse diagnostique et une capacité de généralisation satisfaisante, avec une supériorité globale des architectures basées sur les mécanismes d'attention, notamment le modèle Swin-T. Ces travaux confirment ainsi le potentiel du Deep Learning comme outil d'aide au diagnostic automatique des AVC sur images CT cérébrales et soulignent l'importance de l'explicabilité et de la validation externe pour une future intégration en pratique clinique.

Mots-clés : Accident Vasculaire Cérébral (AVC), Scanner CT cérébral, Apprentissage profond, EfficientNet-B0, Swin Transformer, Vision Transformer, Classification d'images médicales, Grad-CAM, Généralisation, Validation externe, RSNA, Intelligence artificielle.

Abstract

Stroke is a major neurological emergency in which rapid diagnosis is essential to improve patient management and reduce the risk of permanent neurological damage. In this work, we propose an intelligent computer-aided diagnosis system for stroke detection from brain CT images using deep learning techniques. Four pre-trained architectures, namely EfficientNet-B0, MobileNet-V2, ViT-B/16, and Swin-T, were developed and compared in both binary and multiclass classification scenarios aimed at distinguishing normal, ischemic, and hemorrhagic cases. To enhance model performance, several preprocessing steps were applied, including image cleaning, normalization, data augmentation, and transfer learning. Model performance was evaluated using standard metrics such as Accuracy, Sensitivity, Specificity, F1-score, and AUC, and statistically validated using the McNemar, Bootstrap, and DeLong tests. The interpretability of the predictions was ensured through the Grad-CAM method, which enabled visualization of the anatomical regions that most influenced the model decisions and allowed verification of their clinical relevance. In addition, an external validation was performed on the RSNA dataset to assess the generalization capability of the models under domain shift conditions. The obtained results demonstrate high classification performance, strong diagnostic robustness, and satisfactory generalization ability, with overall superiority of attention-based architectures, particularly the Swin-T model. These findings confirm the potential of deep learning as a reliable tool for automated stroke diagnosis from brain CT images and highlight the importance of explainability and external validation for future clinical deployment.

Keywords : Stroke, Brain CT Scan, Deep Learning, EfficientNet-B0, Swin Transformer, Vision Transformer, Medical Image Classification, Grad-CAM, Generalization, External Validation, RSNA, Artificial Intelligence.

الملخص

تُعدّ السكتة الدماغية من أبرز أسباب الوفاة والإعاقة في العالم، مما يجعل التشخيص المبكر والدقيق عاملاً أساسياً لتحسين فرص العلاج وتقليل المضاعفات. يهدف هذا العمل إلى تطوير نظام ذكي لتصنيف السكتات الدماغية اعتماداً على صور التصوير المقطعي المحوسب للدماغ باستخدام تقنيات التعلم العميق. تم في هذه الدراسة توظيف ومقارنة أربع معماريات حديثة في مهام التصنيف الثنائي ومتعدد الفئات. كما تم تطبيق مجموعة من تقنيات المعالجة المسبقة، وزيادة البيانات، والتعلم بالنقل بهدف تحسين أداء النماذج وتعزيز قدرتها على التعميم. ولتفسير قرارات النماذج وزيادة موثوقيتها، تم استخدام إحدى تقنيات الذكاء الاصطناعي القابل للتفسير، بالإضافة إلى إجراء تحليل إحصائي لتقييم دقة النتائج والتحقق من مصداقيتها. أظهرت النتائج المتحصل عليها أداءً مرتفعاً في تصنيف السكتات الدماغية، مع قدرة جيدة على التعميم عند الاختبار على قواعد بيانات خارجية، مما يؤكد فعالية التعلم العميق كأداة مساعدة لدعم التشخيص الطبي.

الكلمات المفتاحية: السكتة الدماغية، التصوير المقطعي المحوسب، التعلم العميق، التصنيف متعدد الفئات، التصنيف الثنائي، التعلم بالنقل، زيادة البيانات، المعالجة المسبقة للصور، الذكاء الاصطناعي القابل للتفسير، التعميم، التشخيص الطبي بمساعدة الحاسوب.

Table de matières

Remerciements.....	VII
Dédicace.....	VII
Dédicace	VII
Résumé.....	VII
Abstract.....	V
المخلص	VII
Table de matières.....	VII
Liste des figures.....	XII
Liste des tableaux	XVII
Liste d'abréviations.....	XVIII
Introduction Générale	XVIII
Chapitre 1 : Contexte médical.	
1. 1 Introduction.....	2
1. 2 Le cerveau humain.....	2
1. 2.1 Rôle et fonctions du cerveau.....	3
1. 2.2 Anatomie générale du cerveau	3
1.2.2.1 L'encéphale	3
1.2.2.1.1 Le cerveau	4
1.2.2.1.2 Le cervelet	4
1.2.2.1.3 Le tronc cérébral	5
1.2.2.2 Les Méninges	5
1.2.2.3 Système ventriculaire et liquide céphalo-rachidien (LCR)	6
1.2.2.4 Substance blanche et substance grise	6
1.2.3 Vascularisation cérébrale	7
1.2.3.1 Principaux axes artériels et territoires	7
1. 3. Pathologies cérébrales.....	8
1.3. 1 Les pathologies non vasculaires cérébrales	8

1. 3. 1. 1 Tumeurs cérébrales	8
1.3.1. 2 Maladies neurodégénératives	9
1.3. 2 Les pathologies vasculaires cérébrales	9
1. 3.2.1 L'Accident Vasculaire Cérébral (AVC).....	10
1. 3.2.1.1 L'AVC Ischémique	10
1. 3.2.1.2 L'AVC Hémorragique	11
1.3.2.1.3 Signes cliniques et score de gravité	12
1. 4. Epidémiologie des AVC.....	12
1.5. Facteurs de risque des AVC.....	13
1.5.1 Facteurs de risque non modifiables.....	13
1.5.2 Facteurs de risque modifiables.....	13
1.6. Hémorragie intracrânienne (HIC)	14
1.6.1 Définition et Problématique	14
1.6.2 Étude des différents types d'HIC	14
1.7. Imagerie médicale pour le diagnostic des HIC	16
1.7.1 La Tomodensitométrie (Scanner CT)	16
1.7.2 L'Imagerie par Résonance Magnétique (IRM)	18
1.7.3 Importance du CT-scan pour les hémorragies	19
1.8. Conclusion.....	20
Chapitre 2 : Fondements théoriques et État de l'art	
2.1 Introduction.....	23
2.2 Intelligence artificielle.....	23
2.3 Apprentissage automatique.....	24
2.4 Réseaux de neurones artificiels.....	24
2.4.1 Neurone artificiel.....	25
2.4.2 Fonction d'activation.....	25

2.4.3 Réseau multicouche(MLP).....	25
2.4.4 Propagation avant.....	25
2.4.5 Rétro-propagation.....	26
2.5 Apprentissage profond.....	26
2.5.1 Importance de l'apprentissage profond.....	26
2.5.2 Fonctionnement de l'apprentissage profond	27
2.5.3 Différences entre l'Apprentissage profond et l'Apprentissage automatique.....	28
2.5.4 Critères de choix entre l'Apprentissage profond et automatique	29
2.5.5 Création et entraînement des modèles de l'Apprentissage profond.....	30
2.6 Réseaux de neurones convolutifs (CNN).....	31
2.6.1 Principe des CNN.....	32
2.6.2 Couche de convolution.....	32
2.6.3 Fonction ReLU.....	32
2.6.4 Pooling.....	33
2.6.5 Fully connected Layer.....	33
2.6.6 Avantages des CNN en imagerie médicale.....	33
2.7 Architectures étudiées.....	33
2.7.1 MobileNet.....	34
2.7.2 EfficientNet.....	35
2.7.3 Vision Transformer.....	36
2.7.4 Swin Transformer.....	37
2.8 Intelligence artificielle explicable (XAI)	37
2.8.1 L'importance de l'XAI en santé.....	38
2.8.2 Grad-CAM.....	38
2.8.3 Avantages et limites.....	38
2.9 État de l'art.....	38
2.9.1 Analyse critique de l'état de l'art.....	42
2.10 Evaluation des modèles de classification.....	44
2.11 Conclusion.....	46
Chapitre 3 : Méthodologie Expérimentale.	
3.1 Introduction.....	49
3.2 Description de la base des données.....	51
3.2.1 Brain Stroke CT Dataset.....	51

3.2.2 RSNA Intracranial Hemorrhage Detection Dataset.....	52
3.3 Méthodes proposées	54
3.3.1 Prétraitement des images CT.....	55
3.3.2 Division des données.....	57
3.3.3 Équilibrage des données.....	58
3.3.4 Augmentation des données.....	59
3.4 Classification.....	60
3.5 Modèles d'apprentissage profond utilisés.....	61
3.5.1 Architecture EfficientNet-B0.....	62
3.5.2 Architecture MobileNet-V2.....	63
3.5.3 Architecture ViT-B/16 (Vision Transformer)	64
3.5.4 Architecture Swin-T (Swin Transformer)	65
3.6 Hyperparamètres d'Optimisation.....	66
3.7 Métriques d'évaluation des modèles	68
3.8 Explicabilité des modèles	69
3.9 Analyse statistique	69
3.9.1 Comparaison des modèles (Test de McNemar)	69
3.9.2 Estimation de l'incertitude (Bootstrap)	70
3.9.3 Évaluation de la généralisation (Test de DeLong)	70
3.10 Conclusion.....	70
Chapitre 4 : Résultats et discussion.	
4.1 Introduction.....	73

4.2 Classification automatique des accidents vasculaires cérébraux sur scanner CT cérébral.....	73
4.2. 1 Évaluation comparative de la classification multi-classes	74
4. 2. 2 Évaluation comparative de la classification binaire (normal/ pathologie)	76
4. 2. 3 Évaluation comparative de la classification binaire (normal/ bleeding)	79
4.2.4 Évaluation comparative de la classification binaire (normal/ Ischemia)	82
4.3 La généralisation.....	84
4. 3. 1 Évaluation de la robustesse et de la généralisation.....	85
4.4. Comparaison entre test interne et test externe	88
4.5. Explicabilité des modèles	89
4. 6. Analyse statistique	91
4.7 Comparaison avec l'état de l'art	92
4.7. 1 Brain Stroke CT Dataset.....	93
4. 7.2 RSNA.....	94
4.8 Conclusion	95
Conclusion générale.....	97

Liste des figures

Figure 1.1 : les composants du cerveau.

Figure 1.2 : Anatomie de l'encéphale montrant l'hémisphère cérébral, le cervelet et le tronc cérébral.

Figure 1.3 : Les différents lobes cérébraux et leurs fonctions principales.

Figure 1.4 : Représentation des trois couches méningées et des espaces péri-cérébraux.

Figure 1.5 : Distribution de la substance grise et de la substance blanche dans le cerveau.

Figure 1.6 : Architecture artérielle de la base du cerveau : le Polygone de Willis

Figure 1.7 : Accident Vasculaire Cérébral Hémorragique.

Figure 1.8 : Diagnostic radiologique des différents types d'hémorragies intracrâniennes sur des coupes axiales de scanner (TDM).

Figure 1.9 : Représentation schématique des composants d'un scanner CT et principe d'acquisition rotative.

Figure 1.10 : Architecture interne d'un système IRM montrant les différentes bobines et l'aimant principal.

Figure 2.1 : Architecture générale d'un réseau de neurones profond avec plusieurs couches cachées.

Figure 2.2 : Comparaison entre le Machine Learning et le Deep Learning pour le diagnostic médical.

Figure 2.3 : Schéma du processus de Transfer Learning : passage d'une architecture pré-entraînée (Backbone) vers une tâche de classification médicale.

Figure 2.4 : Mécanisme des convolutions séparables en profondeur (Depthwise Separable Convolution) utilisé dans l'architecture MobileNet.

Figure 2.5 : Architecture détaillée d'EfficientNet.

Figure 2.6 : Architecture détaillée d'un Vision Transformer (ViT).

Figure 2.7 : L'architecture d'un transformateur Swin (Swin-T).

Figure 2.8 : Structure d'une matrice de confusion illustrant les quatre types de prédictions (VP, VN, FP, FN) pour l'évaluation du diagnostic médical.

Figure 3.1 : pipeline générale de notre travail.

Figure 3.2 : Exemples d'images CT représentant les trois classes de la base Brain Stroke CT Dataset.

Figure 3.3 : Exemples d'images CT de la base RSNA Intracranial Hemorrhage Detection.

Figure 3. 4 : Organigramme des différents phases de notre approche.

Figure 3. 5 : Pipeline de prétraitement appliqué aux images CT.

Figure 3. 6 : Comparaison avant/après prétraitement.

Figure 3. 7 : Répartition des données pour la classification multiclasse dans les ensembles d'entraînement, de validation et de test.

Figure 3. 8 : Répartition des données pour la classification binaire dans les ensembles d'entraînement, de validation et de test.

Figure 3.9 : distribution des trois classes avant et après l'équilibrage.

Figure 3.10 : distribution des deux classes avant et après l'équilibrage.

Figure 3. 11 : Exemples d'images après l'augmentation des données.

Figure 3.12 : Notre architecture pour EfficientNet-B0.

Figure 3.13 : Notre architecture pour MobileNet-V2.

Figure 3.14 : Notre architecture pour ViT-B/16

Figure 3.15 : Notre architecture pour Swin-T.

Figure 4. 1 : Matrices de confusion des quatre modèles d'apprentissage profond pour la classification multiclasse (Bleeding, Ischemia, Normal).

Figure 4. 2 : Matrices de confusion des quatre modèles d'apprentissage profond pour la classification binaire (Pathologie, Normal).

Figure 4. 3 : Matrices de confusion des quatre modèles d'apprentissage profond pour la classification binaire (Bleeding, Normal).

Figure 4. 4 : Matrices de confusion des quatre modèles d'apprentissage profond pour la classification binaire (Ischemia, Normal).

Figure 4.5 : Matrices de confusion comparatives des performances de classification pour EfficientNet-B0 et Swin-T

La figure 4.6 : Visualisation de l'interprétabilité des prédictions du modèle Swin-T à l'aide de la méthode Grad-CAM pour la classe Bleeding.

Figure 4.7 : Analyse statistique comparative entre EfficientNet-B0 et Swin-T sur les datasets interne et RSNA externe.

Liste des tableaux

Tableau 1.1 : Tableau comparatif AVC ischémique / AVC hémorragique

Tableau 1. 2 : Différents types des HIC.

Tableau 1.3 : Comparaison synthétique entre les caractéristiques du Scanner (TDM) et de l'IRM dans le contexte des HIC.

Tableau 2.1 : Synthèse des travaux récents sur la classification des AVC à partir d'images CT.

Tableau 3.1 : Classification de la base de données Brain Stroke CT Dataset

Tableau 3.2 : Répartition des classes du dataset RSNA Intracranial Hemorrhage Detection utilisé pour le test externe.

Tableau 3. 3 : les différentes classes de BDD en classification multiclasse.

Tableau 3. 4 : les deux classes de BDD en classification binaire.

Tableau 3.5 : Motivation du choix des modèles de DL.

Tableau 3.6 : Configuration synthétique des hyperparamètres d'entraînement.

Tableau 3.7 : Hyperparamètres par modèle.

Tableau 4. 1 : Résultats comparatifs des architectures d'apprentissage profond pour la classification multiclassées des AVC.

Tableau 4. 2 : Résultats comparatifs des architectures d'apprentissage profond pour la classification binaire (normal pathologie) des AVC.

Tableau 4. 3 : Résultats comparatifs des architectures d'apprentissage profond pour la classification binaire des AVC.

Tableau 4. 4 : Synthèse comparative des métriques de performance des quatre architectures d'apprentissage profond pour la classification binaire (Normal vs Ischemia).

Le tableau 4.5 : Analyse comparative des performances des modèles testés.

Tableau 4.6 : Comparaison des performances des modèles (EfficientNet-B0 et Swin-T) entre le test interne et l'évaluation externe (RSNA).

Tableau 4.7 : Comparaison des performances des modèles de classification des AVC ischémiques et Hémorragiques.

Tableau 4.8 : Comparaison des performances des modèles de détection des hémorragies intracrâniennes sur le dataset RSNA

Liste des abréviations

AVC : Accident Vasculaire Cérébral.

DL : Deep Learning.

AI : Intelligence Artificielle

ROC : Receiver Operating Characteristic Curve.

FP : Faux positif.

FN : Faux négatif.

VP : Vrai positif.

VN : Vrai négatif

ACC : Accuracy (Exactitude).

CNN : Réseaux de neurones convolutifs.

AUC : Area under the Curve.

XAI : Intelligence Artificielle explicable.

ReLU : Rectified Linear Unit.

GPU : Unité de traitement graphique (Graphics Processing Unit)

Grad-CAM : Gradient-weighted Class Activation Mapping

EfficientNetB0 : Architecture de réseaux de neurones convolutifs de la famille EfficientNet, zéro variant

IRM : Imagerie par Résonance Magnétique

SHAP : SHapley Additive exPlanations

LIME : Explications locales interprétables du modèle diagnostic

Introduction Générale

L'accident vasculaire cérébral (AVC) constitue l'une des principales causes de décès et de handicap dans le monde. Cette pathologie survient lorsqu'une région du cerveau est privée de son apport sanguin, soit à la suite de l'occlusion d'un vaisseau sanguin (AVC ischémique), soit en raison de sa rupture (AVC hémorragique). En l'absence d'une prise en charge rapide, les cellules cérébrales sont privées d'oxygène, ce qui peut entraîner des lésions irréversibles, voire le décès du patient.

Le diagnostic précoce des AVC représente un enjeu majeur en neurologie, où chaque minute gagnée peut améliorer considérablement le pronostic du patient. Pour cette raison, les techniques d'imagerie médicale, notamment la tomodensitométrie (CT-scan) et l'imagerie par résonance magnétique (IRM), jouent un rôle essentiel dans l'identification des lésions cérébrales et l'orientation des décisions thérapeutiques. Toutefois, l'interprétation de ces examens demeure une tâche complexe qui nécessite une expertise spécialisée et peut être affectée par la charge de travail ainsi que par la variabilité des observations entre praticiens.

Au cours des dernières années, l'intelligence artificielle, et plus particulièrement le Deep Learning, a démontré un fort potentiel dans l'analyse automatique des images médicales. Grâce à leur capacité à apprendre des caractéristiques complexes à partir des données, ces méthodes offrent des perspectives prometteuses pour assister les radiologues dans la détection et la classification des AVC, tout en améliorant la rapidité et la fiabilité du diagnostic.

Dans ce contexte, ce projet de fin d'études vise à développer un système intelligent de classification automatique des AVC à partir d'images médicales. L'approche proposée repose sur l'apprentissage par transfert et s'appuie sur quatre architectures de Deep Learning : MobileNet, EfficientNet, Vision Transformer (ViT) et Swin Transformer. L'objectif est de comparer leurs performances afin d'identifier le modèle le plus adapté à cette tâche de classification.

Au-delà de la recherche de performances élevées, ce travail s'intéresse également à des problématiques essentielles telles que la généralisation des modèles, leur robustesse face à des données externes ainsi que l'interprétabilité des décisions produites. Pour cela, des techniques d'intelligence artificielle explicable (XAI) et une analyse statistique des résultats sont intégrées afin d'évaluer de manière rigoureuse la pertinence des modèles développés.

Ce mémoire est organisé en quatre chapitres. Le premier chapitre présente le contexte médical des AVC ainsi que les techniques d'imagerie utilisées pour leur diagnostic. Le deuxième chapitre est consacré aux fondements théoriques du Deep Learning, de

l'apprentissage par transfert, de l'intelligence artificielle explicable et à l'étude des travaux connexes. Le troisième chapitre décrit la méthodologie adoptée, incluant les bases de données, les étapes de prétraitement et les modèles utilisés. Enfin, le quatrième chapitre présente les résultats obtenus, leur analyse comparative, l'étude de généralisation ainsi que les résultats des méthodes d'explicabilité.

Chapitre 1

Contexte médical

1. 1 Introduction

L'Accident Vasculaire Cérébral (AVC) représente une urgence médicale majeure et demeure l'une des principales causes de mortalité et de handicap à l'échelle mondiale. La rapidité du diagnostic est déterminante afin de réduire les lésions cérébrales et limiter les séquelles neurologiques.

Ce chapitre est consacré à la présentation du contexte médical de notre étude. Il débute par un rappel anatomique du cerveau, puis aborde les principales pathologies cérébrales, en mettant l'accent sur la distinction entre l'AVC ischémique et l'AVC hémorragique. Enfin, une attention particulière est accordée aux techniques d'imagerie médicale, notamment la tomodensitométrie (TDM), qui constitue l'outil de référence en situation d'urgence pour le diagnostic des AVC.

1. 2 Le cerveau humain

Le cerveau constitue l'organe central du système nerveux. Il est responsable du contrôle des fonctions cognitives telles que la pensée, la mémoire et les émotions, ainsi que des fonctions sensorielles (vision, toucher, etc.) et motrices. Il assure également la régulation de plusieurs fonctions vitales, notamment la respiration, la température corporelle et la sensation de faim. Avec la moelle épinière, le cerveau forme le système nerveux central (SNC), qui assure la coordination et la transmission des informations entre les différentes parties du corps. [1]

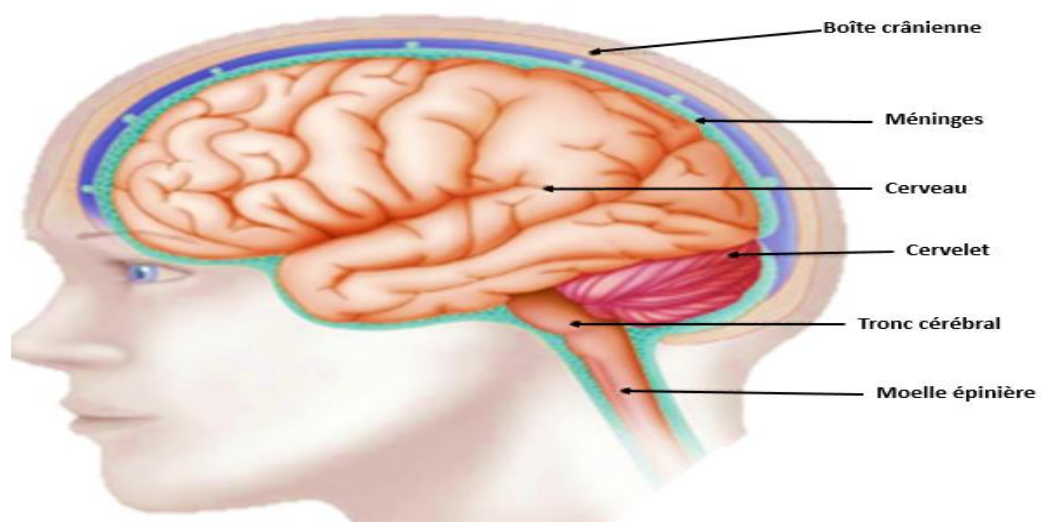


Figure 1.1 : Composants du cerveau

1. 2.1 Rôle et fonctions du cerveau

Le cerveau assure l'intégration et le traitement des informations sensorielles, tout en coordonnant les fonctions motrices et cognitives. Il participe également à la régulation des fonctions physiologiques vitales et permet à l'organisme de s'adapter aux stimuli internes et externes, contribuant ainsi au maintien de l'homéostasie.

En tant que centre de commande du système nerveux central (SNC), il élabore et transmet les commandes motrices vers les muscles et les organes, assurant ainsi le bon fonctionnement et l'adaptation de l'organisme. [1], [2]

1. 2.2 Anatomie générale du cerveau

Pour mieux comprendre l'impact d'un AVC, il est essentiel de décrire l'anatomie générale du cerveau ainsi que ses principales structures.

1.2.2.1 L'encéphale

L'encéphale correspond à la partie du système nerveux central située à l'intérieur de la boîte crânienne, reposant sur la base du crâne. Il est entouré et protégé par le liquide céphalo-rachidien (LCR), qui joue un rôle de protection mécanique et de soutien.

L'encéphale comprend trois structures principales : le cerveau, le tronc cérébral et le cervelet. [3] (Voir Figure 1.2)

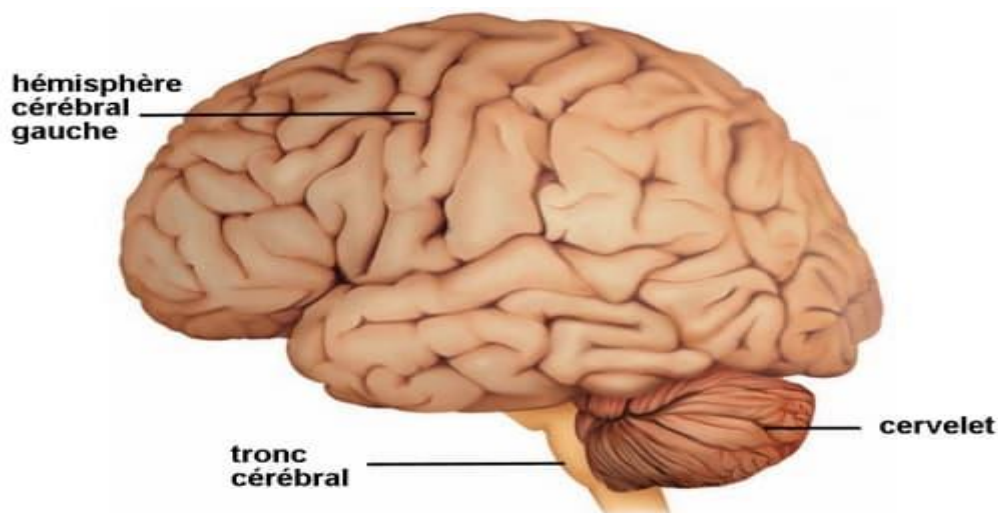


Figure 1.2 : Anatomie de l'encéphale montrant l'hémisphère cérébral, le cervelet et le tronc cérébral [03].

1 .2.2.1.1 Le cerveau

Le cerveau constitue la plus grande partie de l'encéphale. Il est formé de deux hémisphères cérébraux, l'hémisphère droit et l'hémisphère gauche, séparés par la scissure inter-hémisphérique.

Chaque hémisphère est subdivisé en quatre lobes principaux :

- **Le lobe frontal** : il est impliqué dans les fonctions cognitives supérieures telles que la pensée, la planification, l'organisation des actions et la régulation des émotions.
- **Le lobe pariétal** : il intervient dans le traitement des informations sensorielles, la coordination des mouvements volontaires, l'orientation spatiale ainsi que certaines fonctions cognitives comme le calcul.
- **Le lobe occipital** : il est principalement responsable du traitement des informations visuelles.
- **Le lobe temporal** : il joue un rôle essentiel dans l'audition, la compréhension du langage et les mécanismes de la mémoire. [1], [3] (Figure 1.3)

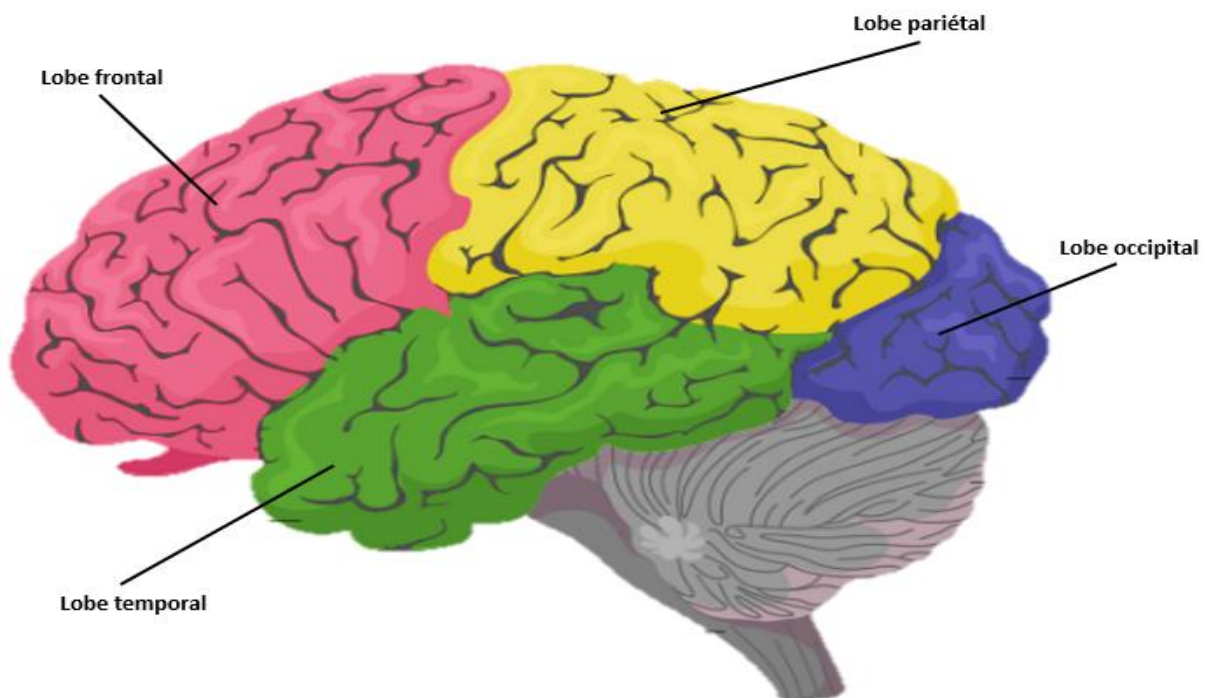


Figure 1.3 : Les différents lobes cérébraux et leurs fonctions principales.

1.1.2.1.2 Le cervelet

Le cervelet est situé en arrière et en dessous du cerveau. Il est relié à ce dernier par des voies nerveuses assurant la transmission des informations sensorielles vers le cortex cérébral ainsi que des commandes motrices vers les muscles.

Il joue un rôle essentiel dans la coordination des mouvements, le maintien de l'équilibre et la précision gestuelle. [3]

1.1.2.1.3 Le tronc cérébral

Le tronc cérébral constitue une structure de liaison fondamentale entre le cerveau, le cervelet et la moelle épinière. Il assure le passage de nombreuses voies nerveuses et participe à la régulation de fonctions vitales essentielles, notamment la respiration et le rythme cardiaque. [3]

1.2.2.2 Les Méninges

Le système nerveux central est protégé par trois membranes superposées appelées méninges. De l'extérieur vers l'intérieur, on distingue la dure-mère, couche fibreuse et résistante, directement accolée à l'os du crâne, l'arachnoïde, puis la pie-mère, qui enveloppe intimement le tissu cérébral.

Entre l'arachnoïde et la pie-mère se trouve l'espace sous-arachnoïdien, dans lequel circule le liquide céphalo-rachidien (LCR). La distinction de ces différentes couches est essentielle, car la localisation d'un saignement (extra- ou sous-méningé) permet d'orienter le diagnostic du type d'hémorragie sur les images scanner . [3] (Voir Figure 1.4)

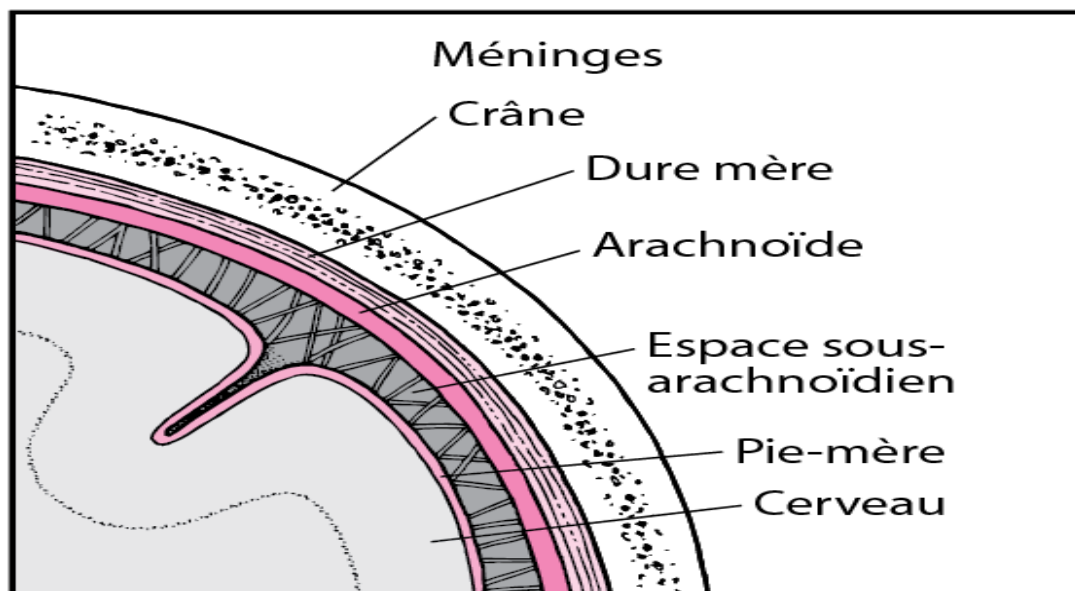


Figure 1.4 : Représentation des trois couches méningées et des espaces péri-cérébraux [03].

1.2.2.3 Système ventriculaire et liquide céphalo-rachidien (LCR)

Le cerveau contient, au sein de son parenchyme, un ensemble de cavités interconnectées appelées ventricules cérébraux. Ce système comprend quatre structures principales : les deux ventricules latéraux, ainsi que les troisième et quatrième ventricules.

Ces cavités sont remplies par le liquide céphalo-rachidien (LCR), un fluide biologique assurant principalement des fonctions de protection mécanique et de drainage du système nerveux central. En imagerie par scanner, la surveillance du système ventriculaire est essentielle, notamment pour détecter des anomalies. La présence de sang au niveau des ventricules peut traduire une hémorragie intraventriculaire, généralement mise en évidence par une hyperdensité visible sur les images [03], [04].

1. 2.2.4 Substance blanche et substance grise

La substance grise et la substance blanche constituent deux composantes essentielles du système nerveux central (SNC), présentant une organisation et des fonctions complémentaires.

Dans le cerveau, la substance grise est localisée en périphérie (cortex cérébral) et assure le traitement et l'intégration de l'information, tandis que la substance blanche, située en profondeur, permet la transmission rapide des signaux nerveux entre les différentes régions cérébrales. À l'inverse, au niveau de la moelle épinière, cette organisation est inversée : la substance blanche est périphérique et la substance grise centrale. Cette complémentarité structurelle garantit le bon fonctionnement global du SNC [3] (Voir Figure 1.5).

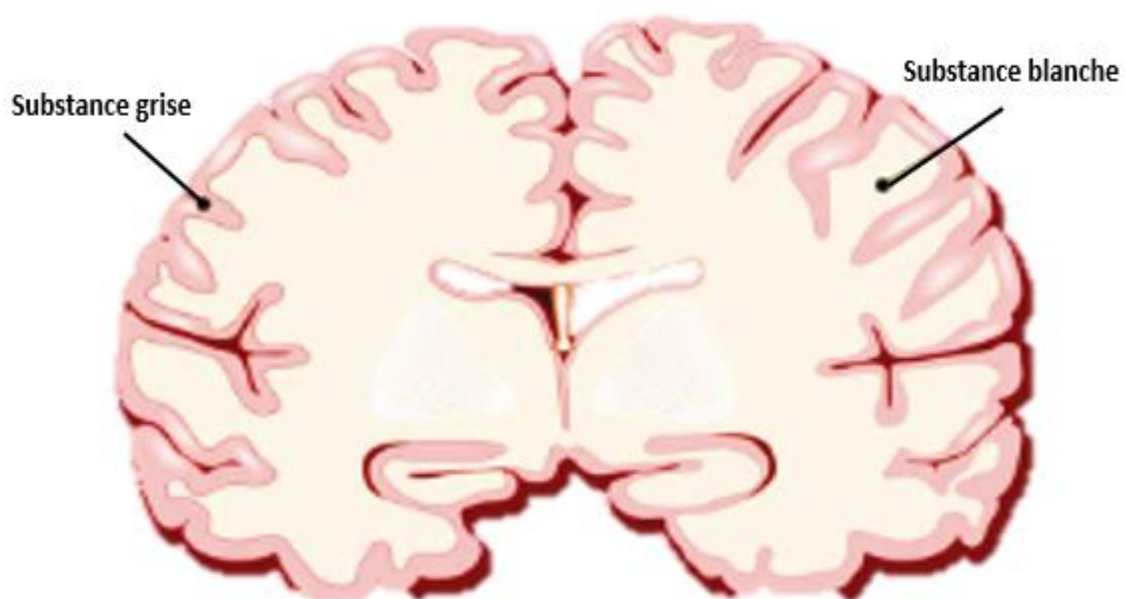


Figure 1.5 : Distribution de la substance grise et de la substance blanche dans le cerveau.

1.2.3 Vascularisation cérébrale

L'apport sanguin au cerveau est assuré par deux systèmes artériels principaux, qui s'anastomosent à la base du crâne pour former un réseau vasculaire complexe garantissant une perfusion continue du tissu cérébral.

1. 2.3.1 Principaux axes artériels et territoires

L'irrigation cérébrale repose essentiellement sur deux systèmes : le système carotidien et le système vertébro-basilaire.

Le système carotidien, principalement via les artères carotides internes, assure environ 80 % du débit sanguin cérébral et vascularise la majorité des hémisphères cérébraux. Parmi ses branches, l'artère cérébrale moyenne (ACM) est la plus fréquemment impliquée dans les accidents vasculaires cérébraux ischémiques.

Le système vertébro-basilaire, quant à lui, assure l'irrigation du tronc cérébral, du cervelet et des régions postérieures du cerveau.

À la base du crâne, ces deux systèmes s'interconnectent pour former le polygone de Willis, une structure anastomotique jouant un rôle de suppléance circulatoire entre les différents territoires cérébraux.

Dans le contexte pathologique, toute obstruction ou rupture d'un vaisseau de ce réseau, notamment au niveau du polygone de Willis ou de ses branches, peut entraîner des déficits neurologiques sévères. En particulier, la rupture vasculaire est responsable d'hémorragies intracrâniennes, caractérisées par une extravasation de sang dans les espaces méningés ou le parenchyme cérébral [03], [04] (Figure 1.6).

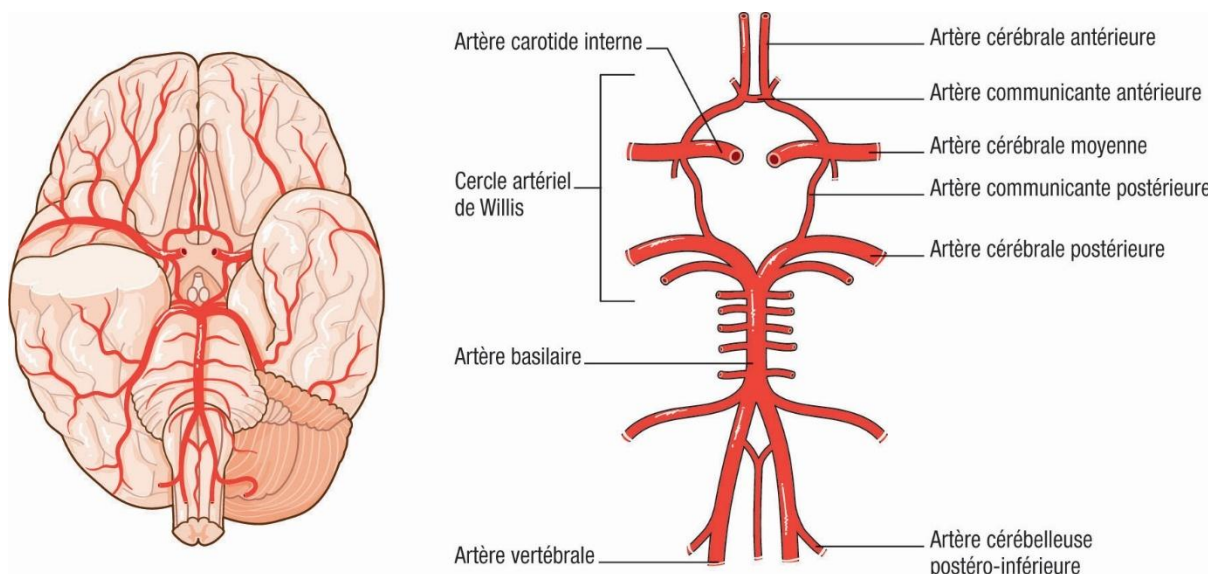


Figure 1.6 : Architecture artérielle de la base du cerveau : le Polygone de Willis [03], [04].

1. 3. Pathologies cérébrales

Le cerveau peut être affecté par diverses pathologies, classées en deux grandes catégories selon leur origine : les pathologies non vasculaires et les pathologies vasculaires.

1.3. 1 Les pathologies non vasculaires cérébrales

Les pathologies non vasculaires du système nerveux central regroupent des affections dont l'origine n'est pas directement liée à une atteinte primaire de la circulation sanguine cérébrale, contrairement aux accidents vasculaires cérébraux. Elles comprennent principalement les tumeurs cérébrales et les maladies neurodégénératives, telles que la maladie d'Alzheimer et la maladie de Parkinson. Ces pathologies présentent des différences importantes sur les plans clinique, radiologique, histopathologique et thérapeutique.

1. 3. 1. 1 Tumeurs cérébrales

Les tumeurs cérébrales, qu'elles soient primitives ou secondaires (métastatiques), résultent d'une prolifération cellulaire anormale. Certaines formes, comme les tumeurs du tronc cérébral associées à la neurofibromatose de type 1, constituent des entités cliniques spécifiques, souvent bien identifiées en imagerie par résonance magnétique [05].

Sur le plan histopathologique, elles peuvent présenter des anomalies microvasculaires telles qu'une hyperplasie endothéliale et des irrégularités de la membrane basale, témoignant

d'une angiogenèse pathologique. Le diagnostic repose principalement sur l'imagerie, complétée par la confirmation histologique. La prise en charge associe généralement neurochirurgie, radiothérapie et chimiothérapie [06].

1.3.1.2 Maladies neurodégénératives

Les maladies neurodégénératives se caractérisent par une perte progressive des neurones, une accumulation anormale de protéines et une altération des fonctions synaptiques, sans origine vasculaire primaire.

- **La maladie d'Alzheimer**

La maladie d'Alzheimer constitue la forme la plus fréquente de démence neurodégénérative. Elle se caractérise principalement par l'accumulation de plaques amyloïdes extracellulaires ainsi que par la présence d'enchevêtrements neurofibrillaires composés de protéine tau hyperphosphorylée [07]. Plusieurs études ont également montré l'existence de mécanismes pathologiques communs avec d'autres maladies neurodégénératives, notamment les maladies à corps de Lewy [07]. Par ailleurs, le vieillissement représente l'un des principaux facteurs de risque associés au développement de cette pathologie, en raison des altérations progressives du métabolisme cérébral et des fonctions neuronales [08].

- **La démence de Parkinson**

La maladie de Parkinson est la deuxième maladie neurodégénérative la plus fréquente. Elle se caractérise par la dégénérescence progressive des neurones dopaminergiques de la substance grise par la présence de corps de Lewy riches en α -synucléine [08]. Toutefois, l'hétérogénéité clinique et moléculaire de cette maladie est aujourd'hui bien reconnue, certaines formes génétiques pouvant évoluer sans dépôts typiques d' α -synucléine, remettant ainsi en question le modèle neuropathologique classique [09].

1.3. 2 Les pathologies vasculaires cérébrales

Les pathologies vasculaires cérébrales regroupent un ensemble de troubles neurologiques aigus ou chroniques liés à une altération du débit sanguin cérébral.

1. 3.2.1. L'Accident Vasculaire Cérébral (AVC)

L'accident vasculaire cérébral (AVC) représente la forme la plus sévère des pathologies vasculaires cérébrales. Il constitue la deuxième cause de mortalité mondiale et la première cause de handicap acquis chez l'adulte.

Selon les définitions épidémiologiques actuelles, l'AVC correspond à un déficit neurologique focal ou global d'installation brutale, d'origine vasculaire, persistant au-delà de 24 heures ou entraînant le décès [1]. Cette définition inclut également l'accident ischémique transitoire (AIT), qui partage un mécanisme physiopathologique similaire, bien que les symptômes soient réversibles [10]

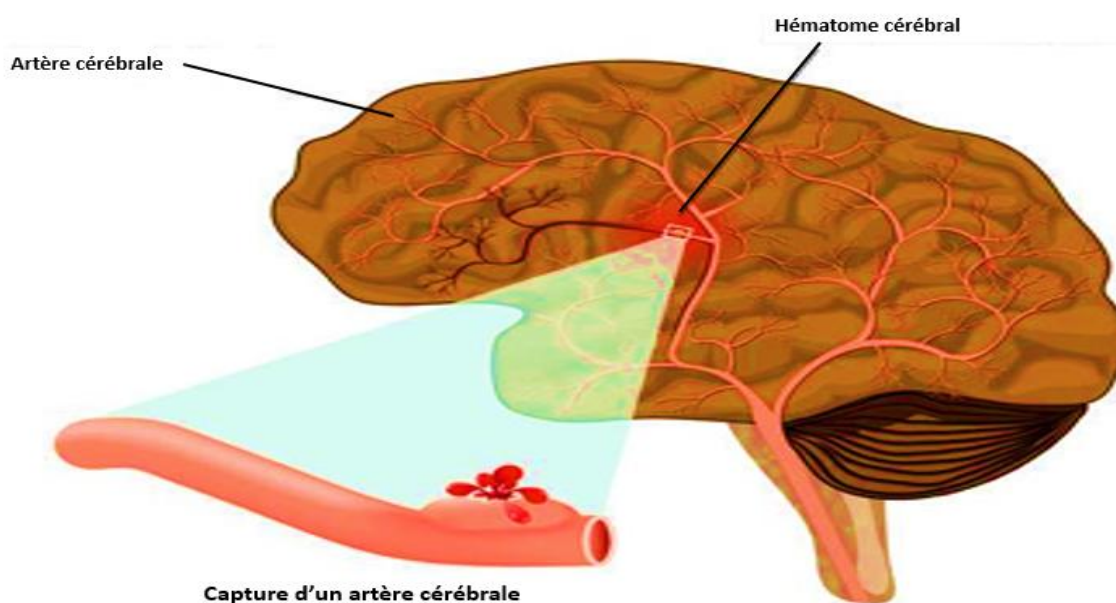


Figure 1.7 : Accident Vasculaire Cérébral Hémorragique.

On distingue deux entités cliniques majeures dont la prise en charge est diamétralement opposée

1. 3.2.1.1 L'AVC Ischémique

Représentant environ 80 % des cas d'AVC [1], l'AVC ischémique résulte de l'occlusion d'une artère cérébrale par un mécanisme thrombotique ou embolique, entraînant une interruption brutale du flux sanguin.

Cette ischémie déclenche une cascade physiopathologique rapide comprenant :

- **Crise bioénergétique** : la privation en oxygène et en glucose entraîne l'arrêt de la production d'ATP, provoquant une défaillance des pompes ioniques et une dépolarisation cellulaire.
- **Excitotoxicité** : la libération excessive de glutamate induit une entrée massive de calcium intracellulaire, activant des enzymes protéolytiques responsables de la mort neuronale.
- **Zone de pénombre** : il s'agit d'un tissu hypo perfusé entourant le cœur de l'infarctus, encore viable mais fonctionnellement altéré. Cette zone constitue la cible privilégiée des stratégies de reperfusion, notamment la thrombolyse et la thrombectomie [11], [12].

1. 3.2.1.2 L'AVC Hémorragique

L'AVC hémorragique représente environ 20 % des cas [1]. Il résulte de la rupture d'un vaisseau cérébral, entraînant une extravasation de sang dans le parenchyme cérébral (hématome intracérébral) ou dans l'espace sous-arachnoïdien [13], [14].

Les lésions neurologiques sont principalement dues à deux mécanismes :

- **Effet de masse** : la collection sanguine exerce une compression directe sur le tissu cérébral, augmentant la pression intracrânienne et perturbant la perfusion cérébrale.
- **Toxicité secondaire** : la dégradation de l'hémoglobine libère du fer, favorisant la production de radicaux libres. Ce stress oxydatif, associé à une réponse inflammatoire intense, contribue au développement d'un œdème périlésionnel.

L'hémorragie intracrânienne (HIC) regroupe l'ensemble de ces saignements intracrâniens, incluant les formes intracérébrales et sous-arachnoïdiennes.

Le tableau I.1 suivant présente une comparaison entre un AVC ischémique et hémorragique.

Critère	AVC ischémique	AVC hémorragique
Fréquence	~80 %	~20 %
Cause	Occlusion	Rupture vasculaire
Scanner	Hypodensité	Hyperdensité
Traitement	Thrombolyse	Contrôle du saignement
Pronostic	Variable	Souvent plus sévère

Tableau 1.1 : Tableau comparatif AVC ischémique / AVC hémorragique

1.3.2.1.3 Signes cliniques et score de gravité

Le diagnostic clinique initial d'un AVC repose sur une reconnaissance rapide des signes neurologiques, notamment à l'aide de la règle FAST (Face, Arm, Speech, Time).

Les manifestations les plus fréquentes incluent une déviation de la bouche, une faiblesse d'un membre et des troubles du langage. La gravité du déficit est quantifiée à l'aide de l'échelle NIHSS (National Institutes of Health Stroke Scale). Un score élevé traduit une atteinte neurologique sévère et nécessite une imagerie en urgence afin de distinguer un infarctus cérébral d'une hémorragie intracrânienne [10] (Voir Figure I.7)

1. 4. Epidémiologie des AVC

L'accident vasculaire cérébral (AVC) constitue un problème majeur de santé publique à l'échelle mondiale. Selon l'Organisation mondiale de la santé (OMS), il représente l'une des principales causes de décès et la première cause de handicap acquis chez l'adulte. Chaque année, plusieurs millions de personnes sont touchées par un AVC, entraînant des conséquences médicales, sociales et économiques considérables.

L'incidence des AVC augmente avec l'âge, bien que cette pathologie puisse également survenir chez des sujets plus jeunes. Le vieillissement de la population, associé à l'augmentation de la prévalence des facteurs de risque cardiovasculaires, contribue à une progression continue du nombre de cas observés dans de nombreux pays. Les AVC ischémiques représentent environ 80 % des cas, tandis que les AVC hémorragiques correspondent à près de 20 % des accidents vasculaires cérébraux.

En Algérie, comme dans de nombreux pays en développement, les AVC constituent une cause importante de mortalité et d'invalidité. L'augmentation de l'espérance de vie, l'urbanisation et les changements du mode de vie ont favorisé la progression des facteurs de risque cardiovasculaires, notamment l'hypertension artérielle, le diabète et l'obésité. Cette situation souligne l'importance du dépistage précoce et de l'amélioration des outils diagnostiques afin d'optimiser la prise en charge des patients.

1.5. Facteurs de risque des AVC

Les facteurs de risque des accidents vasculaires cérébraux peuvent être classés en deux catégories : les facteurs non modifiables et les facteurs modifiables.

1.5.1 Facteurs de risque non modifiables

Certains facteurs sont intrinsèques à l'individu et ne peuvent être modifiés. L'âge constitue le principal facteur de risque, l'incidence des AVC augmentant significativement après 55 ans. Le sexe joue également un rôle, les hommes étant généralement plus exposés aux AVC avant un âge avancé, tandis que le risque devient comparable chez les femmes après la ménopause. Les antécédents familiaux, les prédispositions génétiques ainsi que les antécédents personnels d'AVC ou d'accident ischémique transitoire (AIT) augmentent également le risque de survenue d'un nouvel événement vasculaire cérébral.

1.5.2 Facteurs de risque modifiables

Les facteurs de risque modifiables sont responsables de la majorité des AVC et peuvent être contrôlés par une prise en charge médicale adaptée et l'adoption d'un mode de vie sain.

- **L'hypertension artérielle (HTA)** constitue le principal facteur de risque des AVC ischémiques et hémorragiques. Elle favorise l'altération progressive des parois vasculaires et augmente le risque de rupture ou d'occlusion artérielle.
- **Le diabète** contribue au développement de l'athérosclérose et augmente significativement le risque d'accident vasculaire cérébral.
- **La dyslipidémie**, caractérisée par une élévation du cholestérol sanguin, favorise la formation de plaques athéromateuses pouvant obstruer les artères cérébrales.
- **Le tabagisme** accélère les lésions vasculaires et augmente le risque thrombotique.
- **L'obésité** et la **sédentarité** sont associées à plusieurs facteurs de risque cardiovasculaires, notamment l'HTA et le diabète.
- **La consommation excessive d'alcool** ainsi que certaines maladies cardiovasculaires, telles que la fibrillation auriculaire, augmentent également le risque d'AVC.

L'identification et le contrôle de ces facteurs de risque constituent un enjeu majeur de prévention, permettant de réduire significativement l'incidence des accidents vasculaires cérébraux et leurs conséquences.

1.6. Hémorragie intracrânienne (HIC)

L'hémorragie intracrânienne constitue une urgence neurologique majeure nécessitant un diagnostic rapide et précis, principalement basé sur l'imagerie médicale.

1.6.1 Définition et Problématique

L'hémorragie intracrânienne (HIC) correspond à une rupture de l'homéostasie cérébrale due à l'épanchement de sang à l'intérieur de la boîte crânienne. Ce phénomène entraîne une compression des structures cérébrales, pouvant évoluer vers un œdème ou une ischémie secondaire.

En raison du caractère rigide et inextensible du crâne, toute accumulation sanguine entraîne une augmentation de la pression intracrânienne, faisant du HIC une urgence médico-chirurgicale vitale [05].

1.6.2 Étude des différents types d'HIC

La classification des hémorragies intracrâniennes repose sur la localisation anatomique du sang, critère essentiel pour l'analyse radiologique :

- **Hématome Extra-Dural (HED)** : dû à la rupture d'un vaisseau artériel entre l'os et la dure-mère. En scanner, il apparaît sous forme d'une hyperdensité lenticulaire (biconvexe).
- **Hématome sous-dural (HSD)** : situé sous la dure-mère, il épouse la convexité cérébrale et prend une forme en croissant.
- **Hémorragie sous-arachnoïdienne (HSA)** : le sang diffuse dans les sillons corticaux et les citernes de la base, souvent secondaire à la rupture d'un anévrisme.
- **Hémorragie intra-parenchymateuse (HIP)** : saignement au sein du tissu cérébral, caractérisé par une grande variabilité morphologique, ce qui en fait une cible importante pour les algorithmes de détection automatique.
- **Hémorragie intra-ventriculaire (HIV)** : correspond à une inondation des ventricules cérébraux, facilement identifiable en imagerie par modification de la densité du LCR [04] (Figure 1.8).

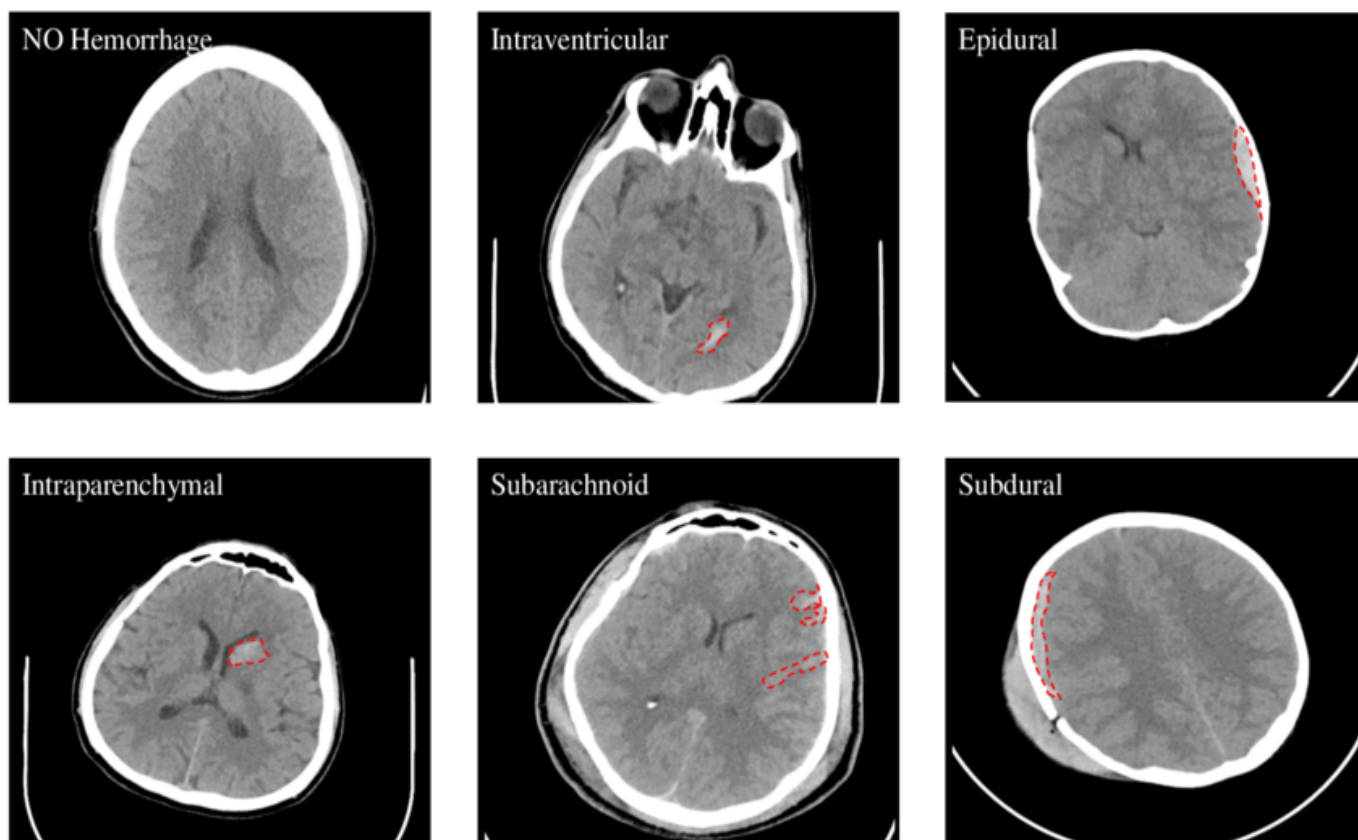


Figure 1.8 : Diagnostic radiologique des différents types d'hémorragies intracrâniennes sur des coupes axiales de scanner (TDM) [04]. (Figure 1.4)

Le tableau I. 2 suivant résume tout ce que nous venons de décrire.

Type	Localisation	Aspect scanner
HED	Extra-Dural	Biconvexe
HSD	Sous-dural	Croissant
HSA	Sous-arachnoïdien	Sillons hyperdenses
HIP	Parenchyme	Hyperdensité focale
HIV	Ventricules	Sang intraventriculaire

Tableau I. 2 : Différents types des HIC.

1. 6.2.1 Étiologies des hémorragies

Chaque type d'hémorragie intracrânienne (HIC) présente une étiologie prédominante permettant d'orienter le diagnostic et la prise en charge thérapeutique.

L'hématome extra-dural (HED) est dans la grande majorité des cas d'origine traumatique, souvent associé à une fracture du crâne entraînant une rupture artérielle.

L'hémorragie sous-arachnoïdienne (HSA) est le plus souvent secondaire à la rupture d'un anévrisme artériel intracrânien.

L'hémorragie intra-parenchymateuse (HIP), quant à elle, est principalement liée à une hypertension artérielle chronique ou à une angiopathie amyloïde cérébrale, notamment chez le sujet âgé.

Cette classification étiologique est essentielle pour orienter la stratégie thérapeutique, qu'elle soit médicale ou chirurgicale [15], [16].

1.7. Imagerie médicale pour le diagnostic des HIC

L'imagerie médicale occupe une place centrale dans le diagnostic, la classification et le suivi des hémorragies intracrâniennes, grâce à plusieurs modalités complémentaires.

1.7.1 La Tomodensitométrie (Scanner CT)

La tomodensitométrie (CT-scan, Computed Tomography) est une technique d'imagerie basée sur l'utilisation des rayons X permettant d'obtenir des coupes axiales du cerveau.

En contexte d'urgence neurologique, elle constitue l'examen de première intention en raison de sa rapidité et de sa grande sensibilité à la détection du sang. En effet, le sang frais apparaît spontanément hyperdense (zones hyperintenses blanches), ce qui facilite l'identification immédiate des hémorragies intracrâniennes.

Cette caractéristique de contraste naturel est également exploitée en intelligence artificielle, notamment pour l'entraînement des modèles de deep learning dédiés à la segmentation et à la détection automatique des lésions [04], [16] (Voir Figure 1.9).

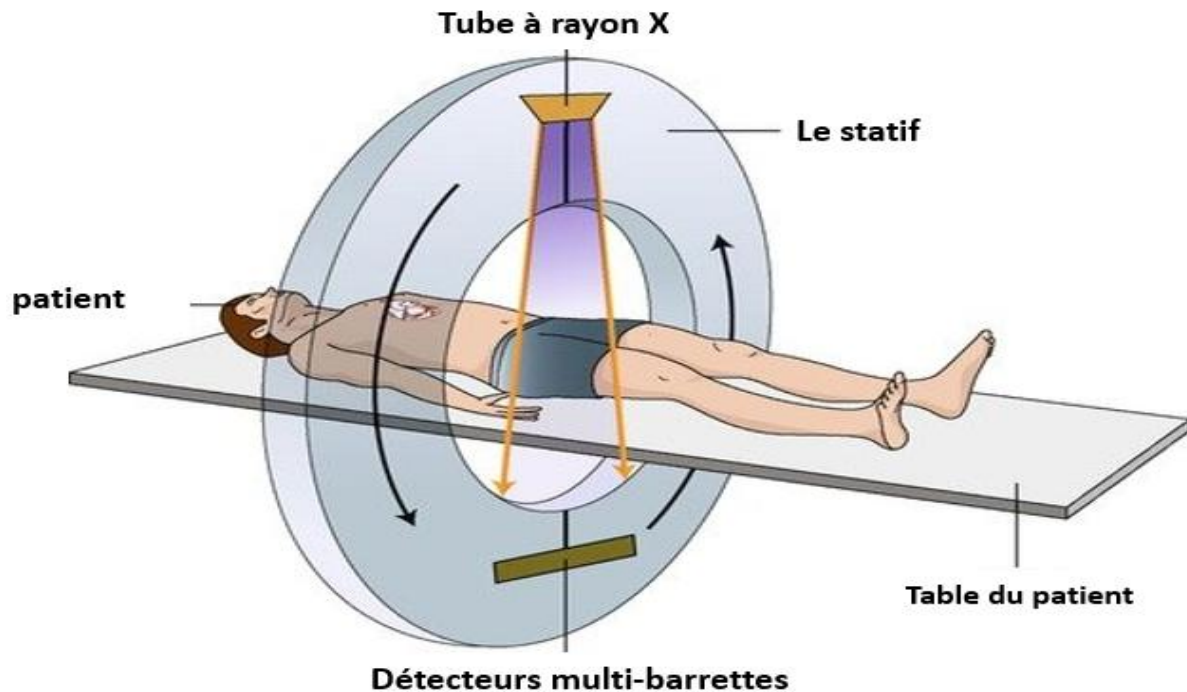


Figure 1.9 : Représentation schématique des composants d'un scanner CT et principe d'acquisition rotative.

1. 7.1.1 Densité et Échelle de Hounsfield (HU)

Le principe de la tomodensitométrie repose sur l'atténuation des rayons X par les tissus biologiques, quantifiée à l'aide de l'échelle de Hounsfield (HU).

Sur cette échelle, l'eau est définie comme référence à 0 HU, l'air à -1000 HU et l'os compact à environ +1000 HU. Le parenchyme cérébral normal présente une densité intermédiaire, généralement comprise entre 20 et 40 HU.

Dans ce contexte, le sang frais apparaît spontanément hyperdense, avec des valeurs situées entre 60 et 80 HU, ce qui lui confère un aspect hyperintense (blanc) nettement distinct des tissus environnants. Cette différence de densité constitue un élément fondamental pour la détection automatisée des hémorragies, notamment dans les approches basées sur la vision par ordinateur et le Deep learning [04], [16].

1.7.2 L'Imagerie par Résonance Magnétique (IRM)

L'imagerie par résonance magnétique (IRM) repose sur l'utilisation de champs magnétiques et d'ondes radio permettant d'obtenir des images de haute résolution des tissus mous.

Bien qu'elle offre une excellente sensibilité pour la détection de lésions cérébrales, notamment chroniques ou subaiguës, elle est moins utilisée en contexte d'urgence pour les hémorragies intracrâniennes. Cette limitation s'explique principalement par un temps d'acquisition plus long et une accessibilité plus contraignante du patient durant l'examen.

Néanmoins, l'IRM reste un outil complémentaire précieux pour préciser l'étiologie et la caractérisation de certaines hémorragies complexes [04] (Voir Figure 1.10).

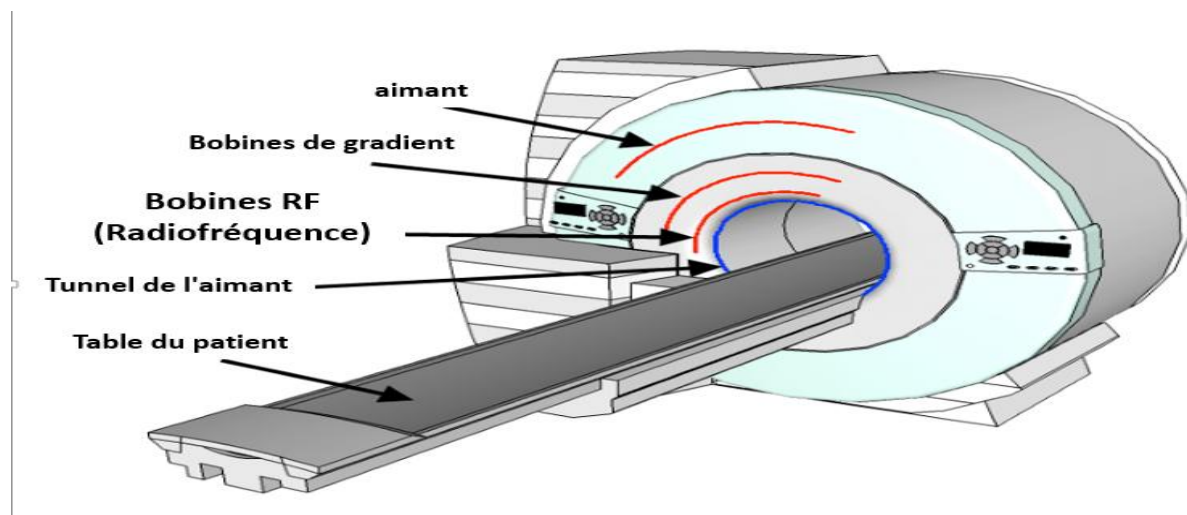


Figure 1.10 : Architecture interne d'un système IRM montrant les différentes bobines et l'aimant principal.

Le tableau 1.3 ci-dessous montre la différence entre les deux modalités d'imagerie, à savoir l'IRM et le CT-Scan

Caractéristiques	IRM	Scanner (CT)
Principe physique	Résonance Magnétique Nucléaire	Atténuation des Rayons X
Temps d'acquisition	Long (20 à 45 minutes)	Très rapide (quelques secondes)
Coût de l'examen	Élevé	Moins coûteux
Disponibilité en urgence	Limitée	Très haute (24h/24)
Détection du sang aigu	Complexe (dépend des séquences)	Excellente (Hyperdensité blanche)
Contre-indications	Pacemakers, implants métalliques	Grossesse, allergie à l'iode

Tableau 1.3 : Comparaison synthétique entre les caractéristiques du Scanner (TDM) et de l'IRM dans le contexte des HIC.

1.7.3 Importance du CT-scan pour les hémorragies

Le recours au scanner (CT-scan) pour le diagnostic des hémorragies intracrâniennes repose principalement sur sa rapidité d'acquisition, réalisée en quelques secondes, ainsi que sur sa large disponibilité dans les services d'urgence.

L'efficacité de cette modalité d'imagerie, associée à l'intelligence artificielle, a été démontrée par plusieurs études récentes. Parmi elles, les travaux de Chilamkurthy et al. [16] ont mis en évidence une performance de détection comparable à celle des radiologues experts. En Algérie, l'automatisation de l'analyse des images médicales constitue un domaine de recherche en plein développement, notamment à travers les travaux de Lamri [15] ainsi que diverses contributions académiques dans les universités, dont l'Université de Tlemcen.

1.7.3.1 Le concept de "Time is Brain"

En neurologie, le principe "Time is Brain" souligne l'importance critique de la rapidité de prise en charge : on estime qu'environ 1,9 million de neurones sont détruits chaque minute lors d'un AVC non traité.

Dans ce contexte, l'intégration de l'intelligence artificielle au scanner permet de réduire significativement les délais diagnostiques, notamment le temps "door-to-needle" (intervalle entre l'arrivée du patient et le début du traitement). En automatisant la détection des lésions et l'estimation du volume hémorragique, l'IA assiste le radiologue dans le triage rapide des cas urgents [15], [16].

1.8. Conclusion

Ce chapitre a permis de présenter le contexte médical de cette étude en abordant les principaux aspects anatomiques, physiopathologiques et diagnostiques des accidents vasculaires cérébraux, avec un intérêt particulier porté aux hémorragies intracrâniennes. L'analyse de l'anatomie cérébrale, de la vascularisation du cerveau ainsi que des différents types d'AVC et d'hémorragies intracrâniennes a mis en évidence la complexité de ces pathologies et la nécessité d'une prise en charge rapide et adaptée.

Par ailleurs, les différentes modalités d'imagerie médicale utilisées dans le diagnostic des lésions cérébrales ont été présentées. Parmi celles-ci, la tomodensitométrie (scanner CT) demeure l'examen de référence en situation d'urgence grâce à sa rapidité d'acquisition, sa large disponibilité et sa grande sensibilité à la détection du sang frais. La précocité du diagnostic constitue en effet un facteur déterminant pour améliorer le pronostic et réduire les séquelles neurologiques des patients.

Toutefois, malgré les performances du scanner cérébral, l'interprétation des images reste une tâche complexe qui dépend fortement de l'expertise du radiologue et du contexte clinique. Face à l'augmentation du nombre d'examens réalisés dans les services d'urgence, le développement d'outils d'aide au diagnostic automatisés apparaît comme une nécessité afin d'améliorer la rapidité, la fiabilité et l'efficacité de la prise en charge.

Dans cette perspective, l'intelligence artificielle, et plus particulièrement les approches basées sur le Deep Learning, suscitent un intérêt croissant dans le domaine de l'imagerie médicale. Ces technologies offrent des perspectives prometteuses pour l'automatisation de la détection et de la classification des hémorragies intracrâniennes à partir des images CT. Le chapitre suivant sera consacré aux fondements théoriques du Deep Learning ainsi qu'à l'état de l'art des travaux récents portant sur l'application de ces méthodes au diagnostic des accidents vasculaires cérébraux.

CHAPITRE 2

État de l'art et Fondements Théoriques du Deep Learning

2. 1 Introduction

Ces dernières années, le secteur de l'intelligence artificielle a connu une progression rapide, en particulier avec l'apparition du Deep Learning (Apprentissage Profond). Cette technologie, qui s'inspire du fonctionnement du cerveau humain, est devenue un instrument indispensable pour l'analyse de données complexes [17].

Dans le domaine de la santé, l'apprentissage profond propose des outils performants pour assister les médecins dans leurs diagnostics, en raison de sa capacité à détecter des motifs et des irrégularités dans les images médicales (telles que la tomodensitométrie et l'IRM) [19].

Ce chapitre vise à présenter les principes de base du Deep Learning. Nous détaillerons son mécanisme, les différentes structures mises en œuvre dans notre recherche (telles que MobileNet, EfficientNet et ViT), ainsi que les techniques d'entraînement et d'évaluation efficaces de ces modèles.

2.2 Intelligence artificielle

L'intelligence artificielle (IA) désigne un domaine scientifique qui vise à développer des systèmes informatiques capables de reproduire certaines capacités intellectuelles habituellement associées à l'être humain. Ces capacités comprennent notamment l'analyse d'informations, l'apprentissage à partir de l'expérience, la résolution de problèmes complexes, l'interprétation du langage naturel ainsi que la reconnaissance d'images, de visages ou de sons [20].

Le fonctionnement de l'IA repose sur l'utilisation de modèles algorithmiques et de méthodes mathématiques permettant aux machines de traiter de grandes quantités de données, d'identifier des relations entre celles-ci et d'améliorer leurs performances au fil du temps. Grâce à ces mécanismes d'apprentissage, les systèmes intelligents peuvent adapter leur comportement en fonction des informations reçues et produire des résultats de plus en plus précis [20].

D'une manière générale, l'intelligence artificielle est classée en deux grandes catégories :

- **L'intelligence artificielle faible**, également appelée IA spécialisée, est développée pour répondre à un objectif précis. Elle excelle dans l'exécution de tâches bien définies telles que la détection d'objets dans une image, la traduction automatique ou encore les assistants conversationnels. Toutefois, ses capacités restent limitées au domaine pour lequel elle a été conçue.

- **L'intelligence artificielle forte**, parfois qualifiée d'IA générale, représente un concept plus avancé dans lequel une machine serait capable de comprendre, raisonner et apprendre de manière comparable à un être humain. Un tel système pourrait accomplir diverses tâches intellectuelles sans être restreint à une application spécifique et disposerait d'une autonomie décisionnelle beaucoup plus importante.

2.3 Apprentissage automatique

L'apprentissage automatique, connu sous le terme anglais *Machine Learning*, constitue l'un des principaux sous-domaines de l'intelligence artificielle. Il s'appuie sur la conception de modèles capables d'extraire des connaissances à partir d'un ensemble de données et d'utiliser ces connaissances pour effectuer des analyses ou prendre des décisions [21].

Contrairement aux approches informatiques traditionnelles, où chaque règle doit être définie à l'avance par le programmeur, les algorithmes d'apprentissage automatique identifient eux-mêmes les relations et les tendances présentes dans les données. Cette capacité leur permet d'améliorer progressivement leur précision et leurs performances lorsqu'ils sont exposés à de nouvelles informations [21].

L'objectif fondamental de cette discipline est de doter les systèmes informatiques d'une certaine autonomie d'apprentissage. En adaptant continuellement leurs paramètres en fonction des expériences acquises, ces systèmes deviennent capables d'optimiser leurs résultats sans intervention humaine directe, se rapprochant ainsi du fonctionnement cognitif observé chez l'être humain [22].

2.4 Réseaux de neurones artificiels

Le réseau de neurones artificiels est constitué d'un ensemble de neurones interconnectés organisés en couches. Le neurone artificiel représente l'élément de base de cette architecture et permet le traitement ainsi que la transmission de l'information au sein du réseau. [17][18]

2.4.1 Neurone artificiel

Le neurone artificiel constitue l'unité de base des réseaux de neurones artificiels. Inspiré du fonctionnement du neurone biologique, il reçoit plusieurs entrées, chacune associée à un poids représentant son importance relative. Les entrées pondérées sont additionnées puis

combinées à un biais avant d'être transmises à une fonction d'activation qui produit la sortie du neurone. Grâce à ce mécanisme, le neurone artificiel est capable d'apprendre des relations complexes entre les données d'entrée et les résultats attendus. Cette structure élémentaire constitue le fondement de la majorité des modèles de Deep Learning utilisés aujourd'hui dans les domaines de la vision par ordinateur et de l'analyse d'images [47], [48].

2.4.2 Fonction d'activation

La fonction d'activation joue un rôle essentiel dans le fonctionnement des réseaux de neurones. Son objectif est d'introduire une non-linéarité dans le modèle afin de permettre l'apprentissage de relations complexes entre les données. Sans fonction d'activation, un réseau composé de plusieurs couches se comporterait comme un simple modèle linéaire, limitant considérablement sa capacité de représentation. Parmi les fonctions les plus utilisées figurent la fonction Sigmoid, la Tangente Hyperbolique (Tanh) et la fonction ReLU (Rectified Linear Unit), cette dernière étant particulièrement populaire dans les architectures modernes en raison de sa simplicité et de son efficacité lors de l'entraînement profond [47], [48].

2.4.3 Réseau multicouche (MLP)

Le perceptron multicouche ou Multi-Layer Perceptron (MLP) représente une extension du neurone artificiel vers une architecture composée de plusieurs couches successives. Un MLP est généralement constitué d'une couche d'entrée, d'une ou plusieurs couches cachées et d'une couche de sortie. Chaque neurone d'une couche est connecté aux neurones de la couche suivante, permettant ainsi au réseau d'extraire progressivement des caractéristiques de plus en plus complexes. Les MLP ont constitué l'une des premières architectures performantes en intelligence artificielle et demeurent à la base des modèles modernes de Deep Learning. Leur capacité à approximer des fonctions non linéaires complexes les rend particulièrement adaptés à diverses tâches de classification et de prédiction [47], [48].

2.4.4 Propagation avant

La propagation avant (Forward Propagation) correspond à l'étape durant laquelle les données traversent le réseau de neurones depuis la couche d'entrée jusqu'à la couche de sortie. À chaque couche, les neurones calculent une combinaison linéaire des entrées pondérées, à laquelle est appliquée une fonction d'activation. Les sorties obtenues deviennent alors les entrées de la couche suivante. Ce processus se poursuit jusqu'à la génération de la prédiction finale du modèle. Dans le contexte de la classification d'images médicales, la propagation avant

permet au réseau d'estimer la probabilité d'appartenance d'une image à une classe donnée, telle qu'un AVC ischémique, hémorragique ou un cas normal [47], [48].

2.4.5 Rétro-propagation

La rétro-propagation (Backpropagation) constitue le mécanisme principal d'apprentissage des réseaux de neurones. Après la propagation avant, une fonction de perte est utilisée pour mesurer l'écart entre les prédictions du modèle et les valeurs réelles. Cette erreur est ensuite propagée en sens inverse à travers le réseau afin de calculer la contribution de chaque paramètre à l'erreur globale. Les poids sont alors ajustés progressivement à l'aide d'un algorithme d'optimisation tel que la descente de gradient ou Adam. Ce processus est répété sur plusieurs itérations jusqu'à ce que le modèle atteigne un niveau de performance satisfaisant. La rétro-propagation est aujourd'hui considérée comme l'un des piliers fondamentaux du Deep Learning moderne [47], [48].

2.5 Apprentissage profond

Le Deep Learning est une branche avancée de l'intelligence artificielle qui permet aux machines d'apprendre de manière autonome, en s'appuyant sur l'expérience plutôt que sur des instructions humaines figées [18]. Contrairement aux méthodes classiques, l'ordinateur n'a pas besoin qu'on lui décrive chaque détail ; il découvre lui-même les caractéristiques importantes des données [17].

Le secret de cette approche réside dans sa structure en "hiérarchie". Concrètement, le modèle apprend d'abord des concepts simples (comme des lignes ou des formes) pour construire progressivement des connaissances beaucoup plus complexes. Cette organisation en plusieurs niveaux, que l'on appelle "couches", est ce qui donne tout son sens au terme "Profond" [23].

2.5. 1 Importance de l'Apprentissage profond

Si le Deep Learning s'est imposé comme une technologie de rupture, c'est avant tout par sa capacité à surpasser les limites de l'apprentissage automatique classique. Là où les algorithmes traditionnels stagnent malgré l'ajout de nouvelles données, le Deep Learning, lui, devient de plus en plus performant à mesure que le volume d'informations augmente. [19]

Sa véritable force réside dans ce qu'on appelle l'apprentissage hiérarchique des représentations. Pour mieux comprendre, imaginez que le modèle ne regarde pas l'image de manière brute ; il la décompose intelligemment. Les premières couches de la structure vont repérer des éléments simples comme des contrastes ou des bords. Plus on s'enfonce dans les

couches "profondes", plus le modèle combine ces petits détails pour reconnaître des formes complexes : des structures anatomiques précises ou, dans notre cas, des signes pathologiques subtils qu'un œil humain pourrait rater sur un scanner cérébral. [18]

Cette autonomie dans l'extraction des caractéristiques (Feature Learning) change tout : on n'a plus besoin d'expliquer manuellement à la machine ce qu'est une anomalie, elle le découvre par elle-même. De plus, grâce au Transfer Learning, nous pouvons exploiter des modèles déjà "intelligents" pour les adapter à la médecine, ce qui permet de gagner un temps précieux et d'obtenir une précision chirurgicale. En résumé, le Deep Learning n'est plus un simple outil informatique, c'est devenu le bras droit technologique de l'expert clinique. [18]

2.5. 2 Fonctionnement de l'Apprentissage profond

Le fonctionnement du Deep Learning repose essentiellement sur les architectures de réseaux de neurones artificiels. C'est d'ailleurs pour cette raison que les termes « Deep Learning » et « Réseaux de neurones profonds » sont souvent utilisés de manière interchangeable.

Ce qui définit l'aspect « profond » d'un réseau, c'est avant tout le nombre de ses couches cachées. Alors que les réseaux de neurones classiques se limitent généralement à deux ou trois niveaux, les architectures modernes peuvent en compter des dizaines, voire plus d'une centaine. [18]

L'avantage majeur de ce mécanisme est qu'il permet au modèle d'apprendre des caractéristiques de plus en plus complexes directement à partir des données brutes. Grâce à l'entraînement sur de vastes ensembles de données labellisées (comme des images médicales annotées), le réseau extrait l'information de manière autonome. Il n'est donc plus nécessaire de procéder à une extraction manuelle des caractéristiques, car le système s'ajuste et se perfectionne de lui-même au fil des couches. [23] (Figure II.1)

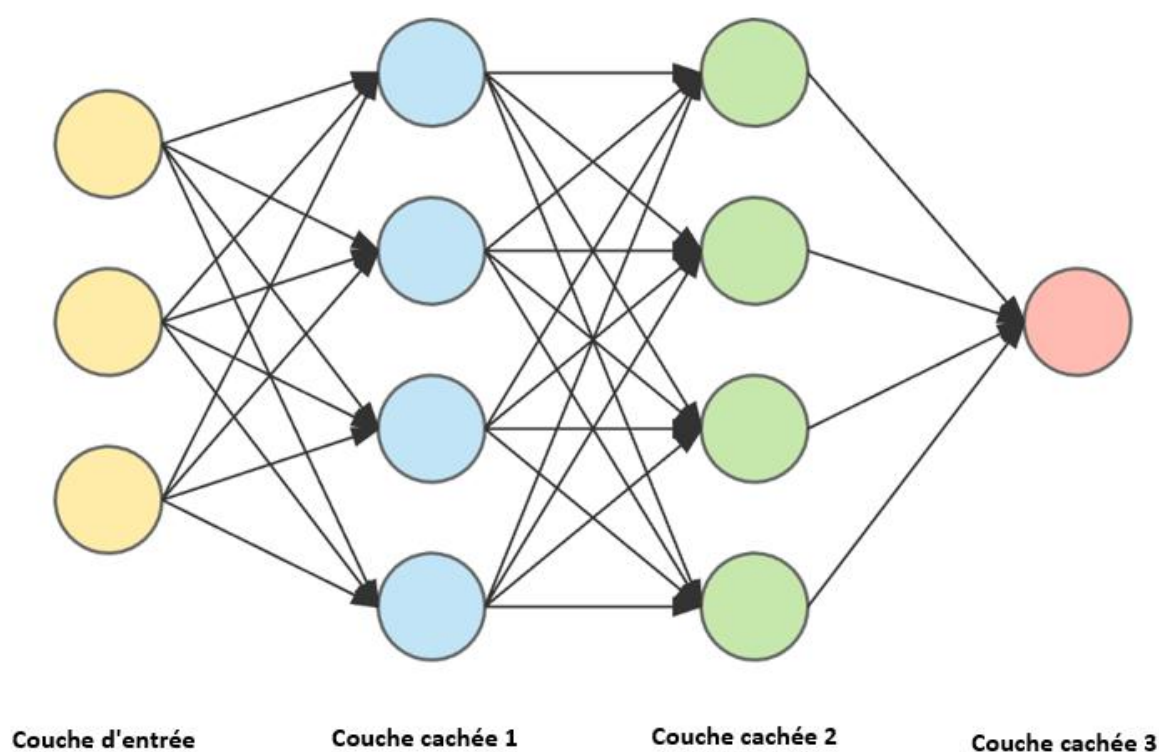


Figure 2.1 : Architecture générale d'un réseau de neurones profond avec plusieurs couches cachées.

2.5. 3 Différence entre l'Apprentissage profond et l'Apprentissage automatique

La distinction fondamentale entre ces deux approches réside dans la méthodologie d'extraction de l'information. L'étape d'extraction des caractéristiques pertinentes à partir des images est faite automatiquement, dans le processus de DL. On parle alors d'un apprentissage « de bout en bout » (End-to-End Learning) [25], à partir de données brutes, le réseau apprend de lui-même à identifier les descripteurs nécessaires pour accomplir la tâche assignée.

À l'inverse, dans un processus de Machine Learning classique, cette étape est manuelle. Elle repose sur l'expertise humaine pour sélectionner les caractéristiques significatives avant de les fournir à un modèle de classification.

Un autre avantage majeur des réseaux de Deep Learning réside dans leur scalabilité : leur capacité à améliorer continuellement leurs performances à mesure que le volume de données d'entraînement augmente [18]. Contrairement aux modèles de Machine Learning traditionnels qui tendent à saturer, le Deep Learning exploite pleinement la richesse des bases de données massives. [17] [23] (Figure II.2)

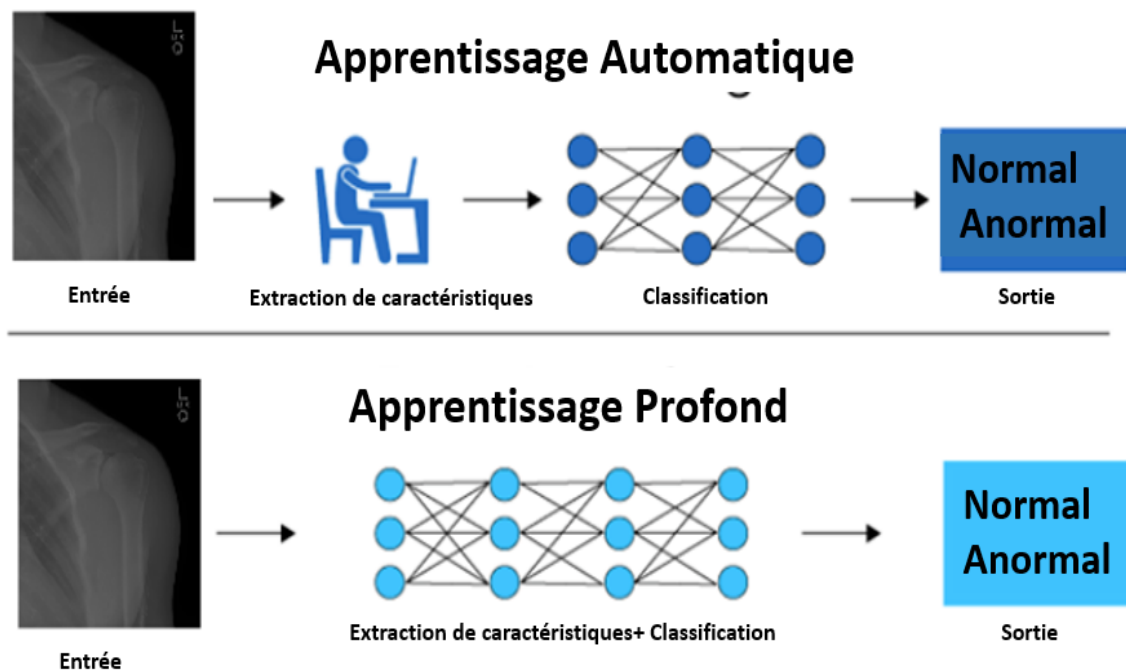


Figure 2.2 : Comparaison entre le Machine Learning et le Deep Learning pour le diagnostic médical.

2.5.4 Critères de choix entre les apprentissages automatique et profond

Le choix entre les approches d'apprentissage automatique (*Machine Learning*) et d'apprentissage profond (*Deep Learning*) dépend principalement des ressources disponibles, notamment la quantité de données et la capacité de calcul.

Les modèles de Deep Learning sont généralement très performants lorsqu'ils sont entraînés sur de vastes ensembles de données. Pour exploiter pleinement leur potentiel, il est souvent nécessaire de disposer de plusieurs milliers d'images ainsi que de ressources matérielles adaptées, telles que des processeurs graphiques (GPU), capables d'accélérer les calculs liés à l'entraînement des réseaux de neurones.

À l'inverse, lorsque le volume de données est limité ou que les ressources de calcul sont restreintes, les méthodes de Machine Learning constituent une alternative plus appropriée. Ces techniques nécessitent généralement moins de données pour l'apprentissage et présentent une complexité computationnelle plus faible. Elles peuvent ainsi fournir des performances satisfaisantes tout en réduisant les exigences matérielles et le temps de traitement.

Par conséquent, le recours au Deep Learning est particulièrement recommandé dans les situations où d'importantes bases de données et des équipements de calcul performants sont disponibles, tandis que le Machine Learning demeure une solution efficace pour les projets disposant de ressources plus modestes.

2. 5. 5 Création et entraînement des modèles de l'apprentissage profond

Dans le cadre de la classification d'images, trois méthodologies principales se distinguent. Le choix de l'une d'entre elles dépend directement de la disponibilité des données, de la puissance de calcul et de la spécificité de la tâche médicale visée.

2.5.5.1 l'entraînement à partir de zéro (*Training from Scratch*)

L'entraînement à partir de zéro consiste à développer un modèle de Deep Learning sans s'appuyer sur des connaissances préalablement acquises. Dans cette approche, l'architecture du réseau de neurones est conçue puis entraînée entièrement à partir d'un ensemble de données annotées. Le modèle apprend progressivement les représentations et les caractéristiques pertinentes directement à partir des exemples qui lui sont fournis.

Cette stratégie est particulièrement adaptée aux problématiques nouvelles pour lesquelles aucun modèle préexistant n'est disponible ou lorsque les données diffèrent fortement des bases de données couramment utilisées. Toutefois, elle requiert généralement un très grand nombre d'échantillons ainsi qu'une importante puissance de calcul. En raison de ces exigences, la phase d'apprentissage peut s'étendre sur plusieurs jours, voire plusieurs semaines, ce qui limite son utilisation dans de nombreuses applications pratiques [20], [23].

2.5.5.2 L'apprentissage par transfert (*Transfer Learning*)

L'apprentissage par transfert représente aujourd'hui l'une des approches les plus populaires en Deep Learning. Son principe repose sur la réutilisation d'un modèle déjà entraîné sur une vaste base de données, puis sur son adaptation à une nouvelle tâche spécifique.

Des architectures reconnues telles que ResNet, EfficientNet, MobileNet, Inception ou LeNet servent fréquemment de point de départ. Le modèle préentraîné est alors ajusté à l'aide d'un nouveau jeu de données contenant les classes d'intérêt de l'application étudiée. Cette méthode permet de bénéficier des connaissances déjà acquises par le réseau tout en réduisant considérablement les besoins en données et en ressources de calcul. Par conséquent, les temps d'entraînement sont souvent réduits à quelques heures, voire quelques minutes selon la taille du problème traité [24].

La figure 2.3 illustre le schéma du processus de Transfer Learning

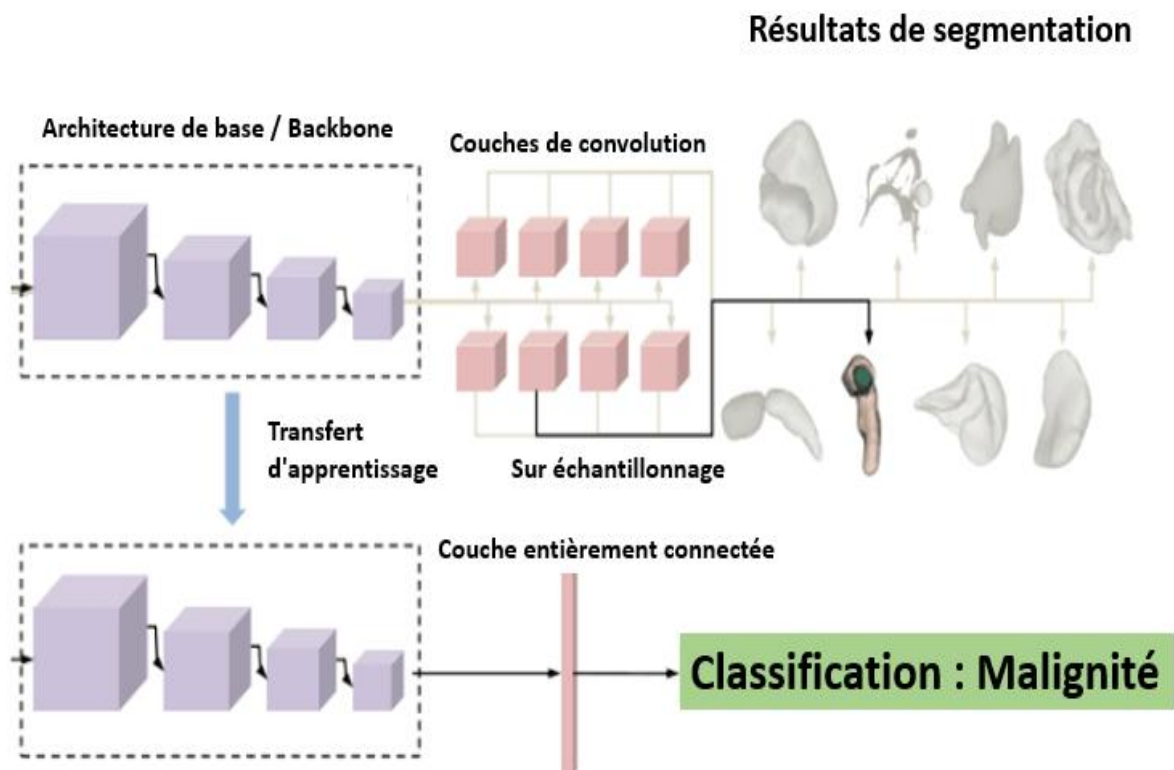


Figure 2.3 : Schéma du processus de Transfer Learning : passage d'une architecture pré-entraînée (Backbone) vers une tâche de classification médicale.

2.5.5.3 Extraction de caractéristiques (*Feature Extraction*)

L'extraction de caractéristiques constitue une autre stratégie d'exploitation des réseaux de neurones profonds. Dans ce cas, le réseau n'est pas utilisé directement pour effectuer la classification finale, mais comme un outil permettant de générer des représentations pertinentes des données d'entrée.

Les différentes couches du réseau apprennent progressivement des caractéristiques de plus en plus complexes à partir des images analysées. Les informations extraites peuvent ensuite être récupérées et utilisées comme variables d'entrée pour des algorithmes classiques d'apprentissage automatique, tels que les arbres de décision, les forêts aléatoires ou les machines à vecteurs de support (SVM).

Bien que moins répandue que l'apprentissage par transfert, cette approche demeure intéressante dans certains contextes spécifiques où l'on souhaite combiner la puissance de

représentation du Deep Learning avec la simplicité et l'interprétabilité de certains modèles de Machine Learning [20], [23].

2.6 Réseaux de neurones convolutifs (CNN)

Les réseaux de neurones convolutifs (Convolutional Neural Networks – CNN) constituent l'une des architectures les plus utilisées en vision par ordinateur et en analyse d'images médicales. Contrairement aux réseaux de neurones classiques, les CNN sont capables d'extraire automatiquement les caractéristiques importantes d'une image grâce à une succession de couches spécialisées. Cette capacité leur permet d'identifier des motifs complexes tels que les contours, les textures ou les structures anatomiques présentes dans les images médicales. Les CNN ont démontré d'excellentes performances dans les tâches de classification, de détection et de segmentation d'images. [17], [18]

2.6.1 Principe des CNN

Le principe fondamental des CNN repose sur l'apprentissage hiérarchique des caractéristiques visuelles. Les premières couches apprennent généralement des caractéristiques simples comme les bords et les formes élémentaires, tandis que les couches profondes extraient progressivement des représentations plus complexes et plus abstraites. Cette organisation hiérarchique permet aux CNN de construire automatiquement des descripteurs pertinents sans nécessiter une extraction manuelle des caractéristiques. [17], [18]

2.6.2 Couche de convolution

La couche de convolution représente l'élément principal d'un CNN. Elle applique des filtres appelés noyaux (kernels) sur l'image d'entrée afin d'extraire différentes caractéristiques visuelles. Chaque filtre est spécialisé dans la détection d'un motif particulier, comme des contours ou des textures. Le résultat de cette opération est une carte de caractéristiques (feature map) qui met en évidence les régions importantes de l'image. Les paramètres des filtres sont appris automatiquement durant l'entraînement du réseau. [18], [22]

2.6.3 Fonction ReLU

Après chaque opération de convolution, une fonction d'activation est appliquée afin d'introduire de la non-linéarité dans le réseau. La fonction la plus utilisée est la fonction ReLU (Rectified Linear Unit), définie par la sortie de la valeur d'entrée lorsqu'elle est positive et par zéro lorsqu'elle est négative. Cette fonction permet d'accélérer l'apprentissage, de réduire le problème du gradient qui disparaît et d'améliorer les performances globales du modèle. [18]

2.6.4 Pooling

Les couches de pooling sont utilisées pour réduire progressivement la dimension spatiale des cartes de caractéristiques produites par les couches de convolution. L'opération la plus courante est le Max Pooling, qui conserve uniquement la valeur maximale d'une région donnée. Cette réduction permet de diminuer le nombre de paramètres du réseau, de réduire le coût computationnel et d'améliorer la robustesse du modèle face aux petites variations présentes dans les images. [18], [22]

2.6.5 Fully Connected Layer

Les couches entièrement connectées (Fully Connected Layers) sont généralement placées à la fin du réseau. Elles reçoivent les caractéristiques extraites par les couches précédentes et réalisent la classification finale. Chaque neurone est connecté à tous les neurones de la couche précédente, permettant ainsi de combiner l'ensemble des informations apprises afin de prédire la classe de l'image analysée. Dans les applications médicales, cette couche produit généralement les probabilités associées aux différentes pathologies étudiées. [17], [18]

2.6.6 Avantages des CNN en imagerie médicale

Les CNN présentent plusieurs avantages pour l'analyse des images médicales. Ils permettent l'extraction automatique des caractéristiques pertinentes sans intervention manuelle, réduisent la dépendance à l'expertise humaine pour la conception des descripteurs et offrent des performances élevées dans les tâches de classification. Grâce à leur capacité à apprendre des représentations complexes, les CNN ont été largement utilisés pour la détection des tumeurs, des maladies pulmonaires, des pathologies cérébrales et des accidents vasculaires cérébraux (AVC). Leur efficacité explique leur adoption croissante dans les systèmes d'aide au diagnostic médical assistés par intelligence artificielle. [17][18][29][32]

2.7. Architectures de Classification

Dans cette section, nous détaillons les mécanismes internes des architectures sélectionnées. Ces modèles reposent sur des principes de conception distincts visant à optimiser l'extraction des caractéristiques et l'efficacité du traitement.

2.7. 1 MobileNet

L'architecture MobileNet se distingue par l'utilisation des Convolutions Séparables en Profondeur, une technique qui réduit massivement le nombre de paramètres et le coût de calcul par rapport aux CNN classiques [25].

Pour optimiser son déploiement sur des systèmes embarqués, des variantes comme Thin MobileNet ont été développées. Ce modèle réduit la taille du réseau (environ 9,9 Mo) en supprimant les couches redondantes et en remplaçant la fonction ReLU par la Drop Activation pour limiter le surapprentissage [26].

Par ailleurs, l'approche FD-MobileNet améliore la vitesse d'inférence en adoptant une stratégie de sous-échantillonnage rapide (Fast-Downsampling). En divisant par deux le nombre de couches tout en conservant une forte capacité d'information, elle permet une exécution fluide sur les processeurs mobiles ARM [27]. Enfin, l'intégration d'un multiplicateur de profondeur permet d'ajuster précisément le compromis entre latence et précision, rendant cette architecture particulièrement flexible pour le diagnostic médical en temps réel [28]. (Figure II.4)

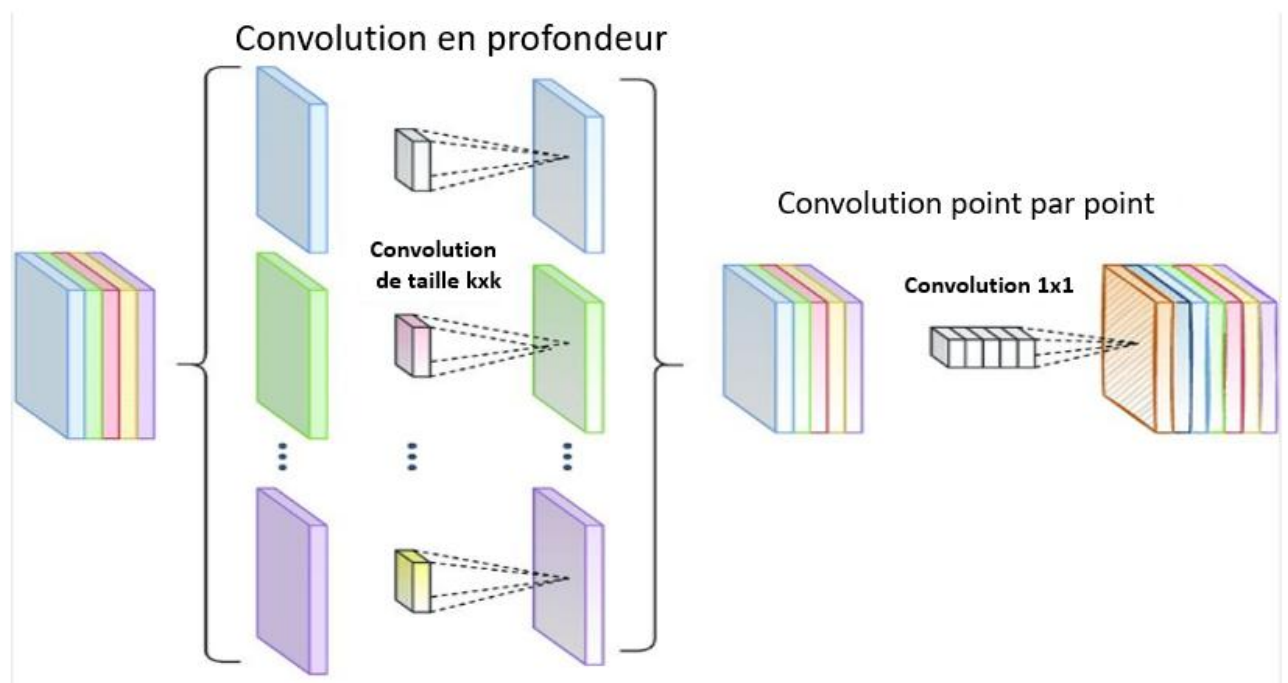


Figure 2.4 : Mécanisme des convolutions séparables en profondeur (Depthwise Separable Convolution) utilisé dans l'architecture MobileNet.

2.7. 2 EfficientNet

L'architecture EfficientNet s'est imposée comme une solution de pointe dans l'imagerie médicale automatisée. Sa force réside dans une méthode de mise à l'échelle systématique appelée Compound Scaling, qui équilibre de manière cohérente [31] la profondeur, la largeur et la résolution du réseau.

Contrairement à ses prédécesseurs, EfficientNet intègre des structures sophistiquées comme les blocs résiduels inversés de MobileNetV2, enrichis par des modules de compression et d'excitation (Squeeze-and-Excitation - SE). Ces modules permettent au réseau de recalibrer intelligemment les caractéristiques de l'image «en focalisant son "attention" sur les zones les plus informatives, ce qui améliore significativement la précision globale sans alourdir le coût de calcul [30].

L'efficacité de cette architecture a été largement démontrée dans le diagnostic de pathologies critiques. Par exemple, des travaux récents ont montré qu'un système basé sur EfficientNet peut atteindre une précision de 99,62 % dans la détection binaire de maladies respiratoires (comme le COVID-19) à partir d'images radiologiques. Cette capacité à fournir un diagnostic rapide et fiable fait d'EfficientNet un outil d'aide à la décision médicale indispensable pour les professionnels de santé dans des contextes de forte pression clinique [31]. (Figure II.5)

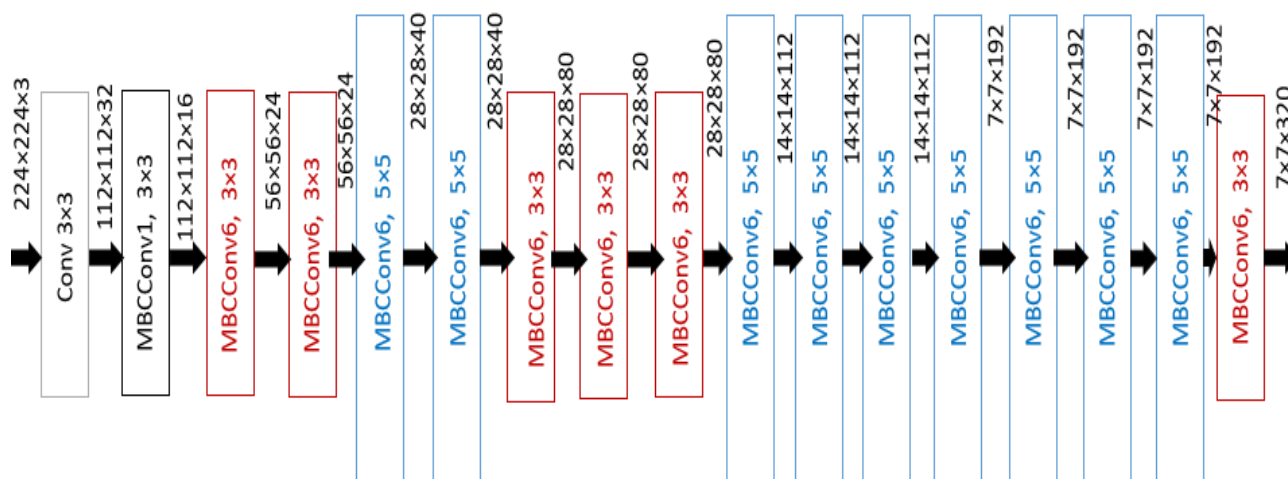


Figure 2.5 : Architecture détaillée d'EfficientNet.

2.7.3 La Révolution des Vision Transformers (ViT)

Les Vision Transformers (ViT) représentent une rupture majeure par rapport aux architectures convolutives traditionnelles. Initialement conçus pour le traitement du langage,

ils ont déclenché une véritable révolution en vision par ordinateur en traitant l'image comme une séquence globale de données [32].

Limites des approches locales et concept de Global Context : Contrairement aux CNN qui se limitent à une analyse locale (pixel par pixel), les ViT exploitent le mécanisme d'auto-attention pour capturer un contexte global dès les premières couches. Cette capacité est cruciale dans l'imagerie médicale (comme la tomodensitométrie) pour modéliser les dépendances spatiales complexes et identifier des anomalies structurelles précises [33], [34].

Mécanisme de Self-Attention et Patch Embedding : L'image est décomposée en patches (petits carrés) qui sont projetés dans un espace vectoriel. Le mécanisme d'attention permet ensuite au modèle de pondérer dynamiquement l'importance de chaque zone. Des études récentes ont démontré l'efficacité de cette approche «notamment le modèle ViT-b16 » dans la classification précise des accidents vasculaires cérébraux (AVC) et d'autres pathologies neurologiques [34], [35]. (Figure II.6)

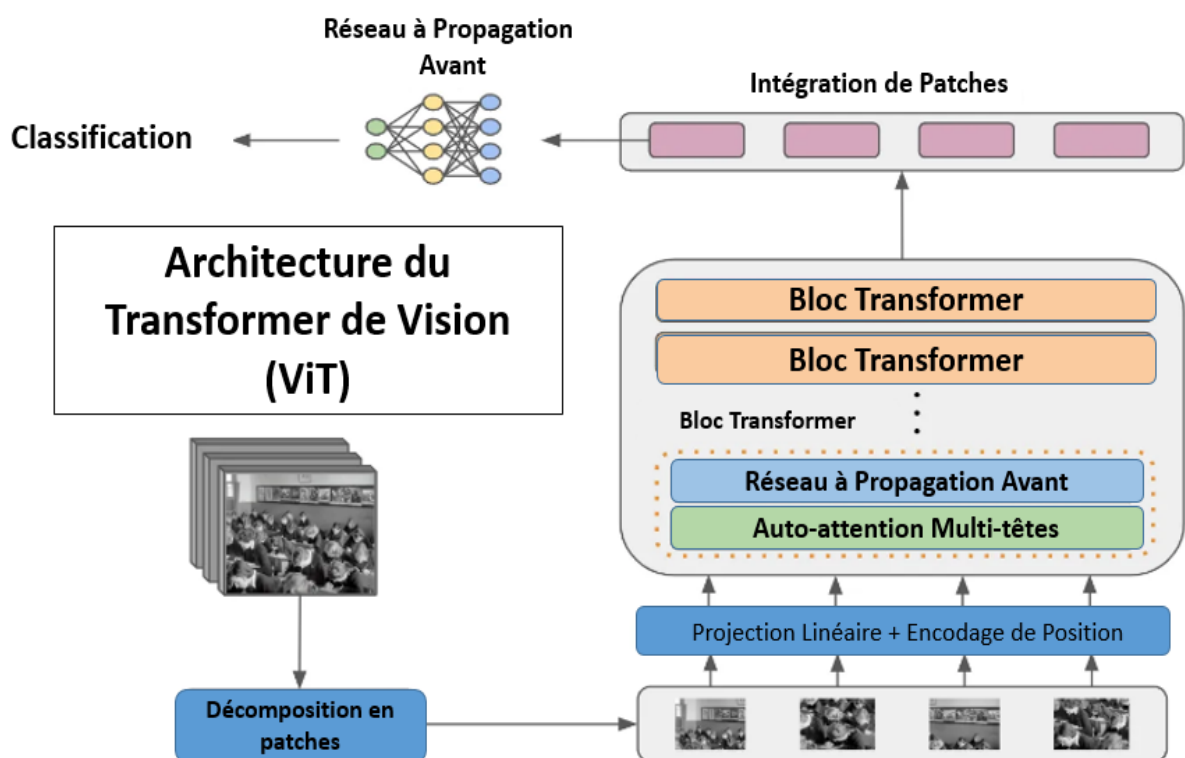


Figure 2.6 : Architecture détaillée d'un Vision Transformer (ViT).

2.7. 4 Swin Transformer

Le Swin Transformer est une architecture récente basée sur le mécanisme d'auto-attention, développée pour améliorer les performances dans les tâches de vision par ordinateur.

Contrairement aux modèles classiques, il traite l'image en plusieurs petites fenêtres locales appelées shifted windows, ce qui lui permet de capturer efficacement les informations locales et globales tout en réduisant le coût computationnel.

Cette architecture repose sur une structure hiérarchique où les caractéristiques extraites deviennent progressivement plus riches et plus abstraites à travers les différentes couches. Dans le domaine médical, le Swin Transformer est particulièrement adapté à l'analyse d'images complexes, car il peut mieux gérer le bruit, les faibles contrastes ainsi que les détails anatomiques difficiles à distinguer. Grâce à son mécanisme d'attention adaptatif, il se concentre sur les régions les plus importantes de l'image et améliore ainsi la qualité de classification et d'analyse des pathologies médicales [46]. (Figure II.7)

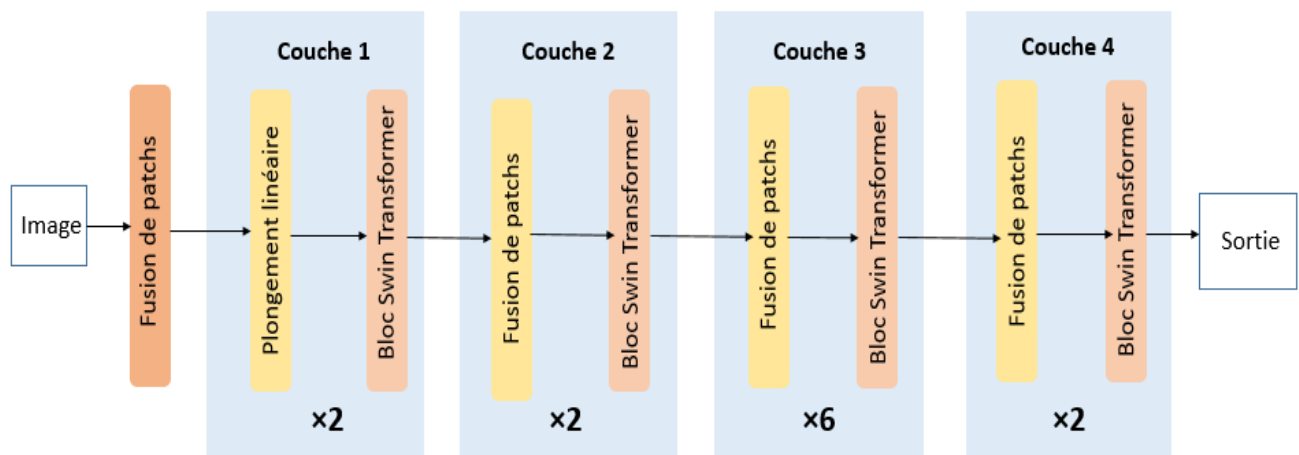


Figure 2.7 : L'architecture d'un transformateur Swin (Swin-T).

2.8 L'Intelligence Artificielle Explicable (XAI)

L'Intelligence Artificielle Explicable (Explainable Artificial Intelligence – XAI) regroupe un ensemble de méthodes permettant de comprendre, d'interpréter et de justifier les décisions prises par les modèles d'intelligence artificielle. Contrairement aux modèles traditionnels de Deep Learning souvent considérés comme des « boîtes noires », les approches XAI fournissent des explications sur les facteurs ayant influencé une prédiction donnée. L'objectif principal est d'améliorer la transparence, la confiance des utilisateurs et l'acceptation des systèmes intelligents dans les domaines critiques tels que la santé [37].

2.8.1 Importance de l'XAI en santé

Dans le domaine médical, la fiabilité et la compréhension des décisions algorithmiques sont essentielles. Les systèmes d'aide au diagnostic basés sur le Deep Learning peuvent atteindre des performances élevées, mais les professionnels de santé doivent comprendre les raisons ayant conduit à une décision avant de l'utiliser dans la pratique clinique. Les méthodes XAI permettent d'identifier les régions d'intérêt utilisées par le modèle et d'évaluer si celles-ci correspondent aux connaissances médicales. Elles contribuent ainsi à renforcer la confiance des médecins, à détecter d'éventuelles erreurs du modèle et à améliorer la sécurité des patients [34], [37].

2.8.2 Grad-CAM

Gradient-weighted Class Activation Mapping (Grad-CAM) est une méthode d'explicabilité largement utilisée pour les réseaux de neurones convolutifs. Elle génère une carte thermique (heatmap) mettant en évidence les zones de l'image ayant le plus contribué à la décision du modèle. Dans le cadre de la classification des AVC sur images CT, Grad-CAM permet de visualiser les régions cérébrales suspectes associées aux lésions hémorragiques ou ischémiques. Cette visualisation facilite l'interprétation des prédictions et la validation clinique des résultats obtenus [37].

2.8.3 Avantages et limites

Les méthodes XAI présentent plusieurs avantages, notamment l'amélioration de la transparence des modèles, le renforcement de la confiance des utilisateurs et la possibilité de détecter des biais ou des erreurs de classification. Elles constituent également un outil précieux pour la validation clinique des systèmes d'aide au diagnostic. Cependant, ces méthodes possèdent certaines limites. Les explications fournies peuvent parfois varier selon la méthode utilisée et ne représentent pas toujours parfaitement le raisonnement interne du modèle. De plus, l'interprétation des cartes explicatives nécessite souvent l'expertise d'un spécialiste du domaine médical [37].

2.9 État de l'Art

Plusieurs études récentes ont été réalisées sur la détection et la classification de l'Accident Vasculaire Cérébral (AVC) à partir d'images médicales. La majorité de ces approches repose sur l'utilisation de modèles de Deep Learning, qui ont montré des performances prometteuses dans l'analyse automatique et la classification des pathologies neurologiques.

- Les recherches récentes en classification automatique des accidents vasculaires cérébraux à partir d'images CT sans contraste (NCCT) confirment l'efficacité croissante des approches de Deep learning pour l'aide au diagnostic en situation d'urgence. Dans ce cadre Turyamuhaki et al. (2026) proposent un modèle d'ensemble profond innovant combinant des réseaux de neurones convolutionnels (CNN) et de Vision Transformers (ViT). L'étude s'appuie sur un jeu de données de 2 501 images (1 127 pathologiques et 1 374 normales), optimisé par des techniques d'augmentation de données et une réduction progressive de la fonction de Loss pour garantir une convergence stable. Les résultats expérimentaux démontrent une performance exceptionnelle avec une Accuracy (exactitude) de 99,77 %, tandis que l'analyse de la courbe ROC (Receiver Operating Characteristic) et de l'AUC confirme une capacité de distinction quasi parfaite entre les classes. Enfin, l'intégration de l'intelligence artificielle explicable (XAI), via LIME et la sensibilité à l'occlusion, permet de localiser précisément les zones lésionnelles, renforçant l'interprétabilité clinique du modèle pour une prise en charge rapide des patients. [38]

- Les recherches menées par Rajaramanan et al. (2024) propose une approche hybride robuste pour la classification tripartite des images scanner, distinguant les cas ischémiques, hémorragiques et normaux. Leur méthodologie repose sur l'utilisation du réseau EfficientNet-B0 en tant qu'extracteur de caractéristiques profonds, dont les vecteurs de sortie sont ensuite injectés dans des classifieurs statistiques classiques tels que les Machines à Vecteurs de Support (SVM) et le k-plus proches voisins (KNN). L'étude s'appuie sur un échantillon de 2 501 images CT, permettant au modèle de capturer des variations de densité subtiles souvent ignorées par les approches purement end-to-end. Les résultats démontrent l'efficacité de cette fusion avec une précision de 95,13 %, soulignant l'importance de l'optimisation des hyperparamètres pour la détection différentielle des lésions cérébrales aiguës. [39]

- Les travaux d'Okonkwo et al. (2025) introduisent une méthodologie d'analyse cognitive dédiée à la précision diagnostique des hémorragies intracrâniennes (ICH). L'architecture, basée sur un réseau ResNet-18, a été entraînée et validée sur un dataset comprenant 2 501 images, spécifiquement sélectionnées pour identifier la présence d'un saignement et classifier ses sous-types cliniques (IVH, SDH, SAH, EDH). L'originalité de leur approche réside dans l'équilibre trouvé entre une haute spécificité (97,1 %), visant à éviter les faux diagnostics, et une sensibilité de 86,2 %, adaptée aux exigences du triage d'urgence. En intégrant des mécanismes de compréhension des caractéristiques

morphologiques du sang, cette étude facilite une prise de décision neurochirurgicale plus ciblée. [40]

- Les recherches récentes mettent en avant les travaux de Tamilarasi et al. (2026) introduisent une architecture hybride originale nommée InHLeNet, qui fusionne InceptionV3Net pour l'extraction de caractéristiques multi-échelles et LeNet pour le traitement de données légères. Leur méthodologie intègre une couche d'analyse harmonique et l'utilisation de caractéristiques de motifs de gradient local (WSLGP) pour capturer les textures subtiles des tissus cérébraux. Fondée sur plus de 2 500 images CT, leur approche utilise un filtrage guidé et des techniques d'auto-augmentation pour améliorer la précision diagnostique. Les résultats affichent une Accuracy de 96,88 % et une sensibilité élevée (TPR de 97,98 %), surpassant les modèles standards. En outre, l'application de Grad-CAM fournit des cartes de chaleur permettant de visualiser les foyers pathologiques, assurant ainsi une transparence décisionnelle cruciale pour l'acceptation clinique du système. [41]

- L'étude de Santhanam et al. (2025) propose une architecture complexe combinant le traitement temporel et l'optimisation algorithmique sur une base de 2 501 images CT. Leur système diagnostique utilise un réseau BiLSTM (Bidirectional Long Short-Term Memory) pour analyser les corrélations spatiales entre les coupes de CT successives, simulant ainsi l'examen séquentiel réalisé par un radiologue. Pour maximiser la pertinence des données traitées, la méthodologie intègre un Algorithme Génétique chargé de la sélection optimale des descripteurs (features). Cette synergie permet une détection précoce des AVC tout en améliorant la robustesse du modèle face aux artefacts d'acquisition et au bruit d'image, garantissant ainsi une meilleure fiabilité diagnostique. [42]

- Les travaux de Gupta et al. (2025) se concentrent sur l'optimisation du temps de réponse médical via une détection multi-classes automatisée par CNN, utilisant également le jeu de données standard de 2 501 images. Leur recherche met en lumière l'importance d'un pipeline d'extraction automatique des zones pathologiques, couvrant à la fois l'ischémie et l'hémorragie, afin de réduire drastiquement le délai entre l'acquisition de l'image et l'interprétation. En s'appuyant sur des architectures convolutives optimisées pour l'inférence rapide, les auteurs démontrent que l'IA peut servir de "premier lecteur" efficace, capable de prioriser les cas critiques au sein des unités neuro-vasculaires. Cette approche vise directement à améliorer le pronostic fonctionnel des patients en accélérant l'accès aux thérapies de reperfusion. [43]

- Les recherches récentes explorent également le développement d'architectures personnalisées pour surpasser les modèles pré-entraînés standards. Dans ce contexte, Sabir

et al. (2025) proposent un nouveau modèle de réseau de neurones convolutionnels profonds (DCNN) spécifiquement optimisé pour la détection précoce des accidents vasculaires cérébraux. Leur méthodologie repose sur une structure tripartite comprenant un extracteur de caractéristiques à plusieurs couches, un module de fusion de caractéristiques pour créer une représentation plus informative de l'image, et un module de détection finale. L'étude s'appuie sur un jeu de données de 2 501 images CT, prétraitées pour améliorer le contraste et réduire le bruit, facilitant ainsi l'identification des signes subtils d'ischémie. Les résultats expérimentaux démontrent une performance robuste avec une Accuracy de 98,2 %, une précision de 98,3 % et une sensibilité de 98,2 %, surpassant plusieurs modèles de l'état de l'art. Enfin, les auteurs soulignent que l'efficacité du modèle réside dans sa capacité à fusionner des informations provenant de différentes profondeurs du réseau, ce qui améliore la transparence et la fiabilité du diagnostic automatisé pour une intervention médicale rapide. [44]

Les recherches récentes mettent en avant les travaux d'Inamdar et al. (2025), qui proposent une approche novatrice basée sur des mécanismes de double attention avec une intégration de la signification au niveau des patches pour la classification des AVC ischémiques. Leur méthodologie repose sur l'utilisation d'un extracteur de caractéristiques (backbone) VGG-16 couplé à des modules d'attention spatiale et de canal, permettant au modèle de se concentrer sélectivement sur les régions pathologiques les plus pertinentes. L'originalité de l'étude réside dans l'intégration d'un encodage de position (positional embedding) qui préserve les relations spatiales entre les différentes zones du cerveau. L'étude s'appuie sur un jeu de données de 2501 images CT, soumises à un prétraitement rigoureux incluant le filtrage et la normalisation. Les résultats expérimentaux sont très compétitifs, atteignant une Accuracy de 95,84 % et un score F1 de 95,82 %. Pour renforcer l'utilité clinique, les auteurs intègrent des techniques de XAI via des cartes d'activation, permettant de visualiser précisément les zones d'infarctus identifiées par le réseau. Cette approche améliore considérablement l'interprétabilité du diagnostic automatique, facilitant ainsi une intervention médicale rapide et précise. [45].

Le Tableau II.1 représente la synthèse des travaux récents sur la classification des AVC à partir d'images CT.

Référence	Année	Base de données	Méthode proposée	Objectif	Performances
Turyamuhaki et al. [40]	2026	2501 images CT	Modèle d'ensemble CNN + Vision	Classification des AVC	Accuracy = 99,77 % ; AUC élevée

			Transformer (ViT)		
Rajaramanan et al. [41]	2024	2501 images CT	EfficientNet-B0 + SVM/KNN	Classification ischémique, hémorragique et normale	Accuracy = 95,13 %
Okonkwo et al. [42]	2025	2501 images CT	ResNet-18	Détection et classification des hémorragies intracrâniennes	Sensibilité = 86,2 % ; Spécificité = 97,1 %
Tamilarasi et al. [43]	2026	Plus de 2500 images CT	InHLeNet (InceptionV3 + LeNet)	Classification des AVC	Accuracy = 96,88 % ; Sensibilité = 97,98 %
Santhanam et al. [44]	2025	2501 images CT	BiLSTM + Algorithme Génétique	Détection précoce des AVC	Amélioration de la robustesse diagnostique
Gupta et al. [45]	2025	2501 images CT	CNN optimisé	Détection multi-classes des AVC	Réduction du temps de diagnostic
Sabir et al. [46]	2025	2501 images CT	DCNN personnalisé	Détection précoce des AVC	Accuracy = 98,2 % ; Précision = 98,3 % ; Sensibilité = 98,2 %
Inamdar et al. [47]	2025	2501 images CT	VGG16 avec double mécanisme d'attention	Classification des AVC ischémiques	Accuracy = 95,84 % ; F1-score = 95,82 %

Tableau 2.1 : Synthèse des travaux récents sur la classification des AVC à partir d'images CT.

2.9.1 Analyse critique de l'état de l'art

L'analyse des travaux récents montre que les approches basées sur le Deep Learning ont considérablement amélioré les performances de détection et de classification des accidents vasculaires cérébraux à partir d'images CT. Les modèles convolutionnels tels que ResNet, EfficientNet et les architectures CNN personnalisées atteignent des taux de précision élevés, démontrant leur capacité à extraire automatiquement des caractéristiques pertinentes à partir des images médicales.

Cependant, plusieurs limites subsistent. Tout d'abord, les architectures CNN présentent une capacité limitée à capturer les dépendances globales entre les différentes régions d'une image. Leur fonctionnement repose essentiellement sur des opérations locales de convolution, ce qui peut entraîner une perte d'informations contextuelles importantes, notamment dans le cas de lésions diffuses ou de faible contraste. Cette limitation peut affecter la qualité de la classification dans certaines situations cliniques complexes.

Par ailleurs, la plupart des modèles de Deep Learning sont souvent considérés comme des « boîtes noires ». Bien qu'ils offrent d'excellentes performances, ils fournissent peu d'explications sur les mécanismes ayant conduit à une décision diagnostique donnée. Ce manque d'interprétabilité constitue un frein majeur à leur adoption en pratique clinique, où les médecins doivent pouvoir comprendre et justifier les résultats fournis par les systèmes d'aide au diagnostic.

Une autre limite concerne la capacité de généralisation des modèles. De nombreuses études sont développées et évaluées sur des bases de données spécifiques, parfois de taille limitée ou provenant d'un seul centre médical. Dans ces conditions, les performances obtenues peuvent diminuer lorsqu'un modèle est appliqué à des données provenant d'autres hôpitaux, scanners ou populations de patients. Cette problématique de généralisation demeure un défi majeur pour le déploiement clinique des systèmes d'intelligence artificielle.

Face à ces limitations, les recherches récentes s'orientent vers le développement de modèles hybrides combinant les avantages de plusieurs architectures. L'association des réseaux convolutionnels (CNN) avec les Vision Transformers (ViT) ou les mécanismes d'attention permet d'exploiter simultanément les caractéristiques locales et le contexte global des images médicales. Ces approches hybrides offrent généralement une meilleure robustesse, une meilleure capacité de généralisation et des performances supérieures à celles des architectures utilisées individuellement.

Enfin, l'intégration de techniques d'intelligence artificielle explicable (XAI), telles que Grad-CAM et LIME, apparaît aujourd'hui comme une nécessité pour renforcer la confiance des cliniciens envers les systèmes automatisés. Ces méthodes permettent de visualiser les régions d'intérêt ayant influencé la décision du modèle et contribuent ainsi à améliorer la transparence et l'acceptabilité clinique des solutions basées sur l'intelligence artificielle.

Au regard de ces constats, notre travail s'inscrit dans cette nouvelle génération d'approches en comparant plusieurs architectures avancées de Deep Learning, notamment MobileNet, EfficientNet, Vision Transformer (ViT) et Swin Transformer, afin d'identifier la solution la plus performante pour la classification automatique des AVC tout en garantissant la fiabilité et la robustesse des modèles.

2.10 Évaluation des Modèles de Classification

L'évaluation des performances est une étape décisive pour valider la fiabilité d'un système de diagnostic médical. Nous utilisons des métriques standardisées qui permettent de quantifier la précision et la capacité de généralisation de nos modèles.

- **Matrice de Confusion**

La matrice de confusion est un outil fondamental pour visualiser les performances du modèle. Elle permet de classer les prédictions en quatre catégories : les Vrais Positifs (VP), les Vrais Négatifs (VN), les Faux Positifs (FP) et les Faux Négatifs (FN). Dans le cadre clinique, l'objectif est de réduire au maximum les Faux Négatifs pour garantir qu'aucune pathologie ne soit ignorée [19]. (Figure II.7)

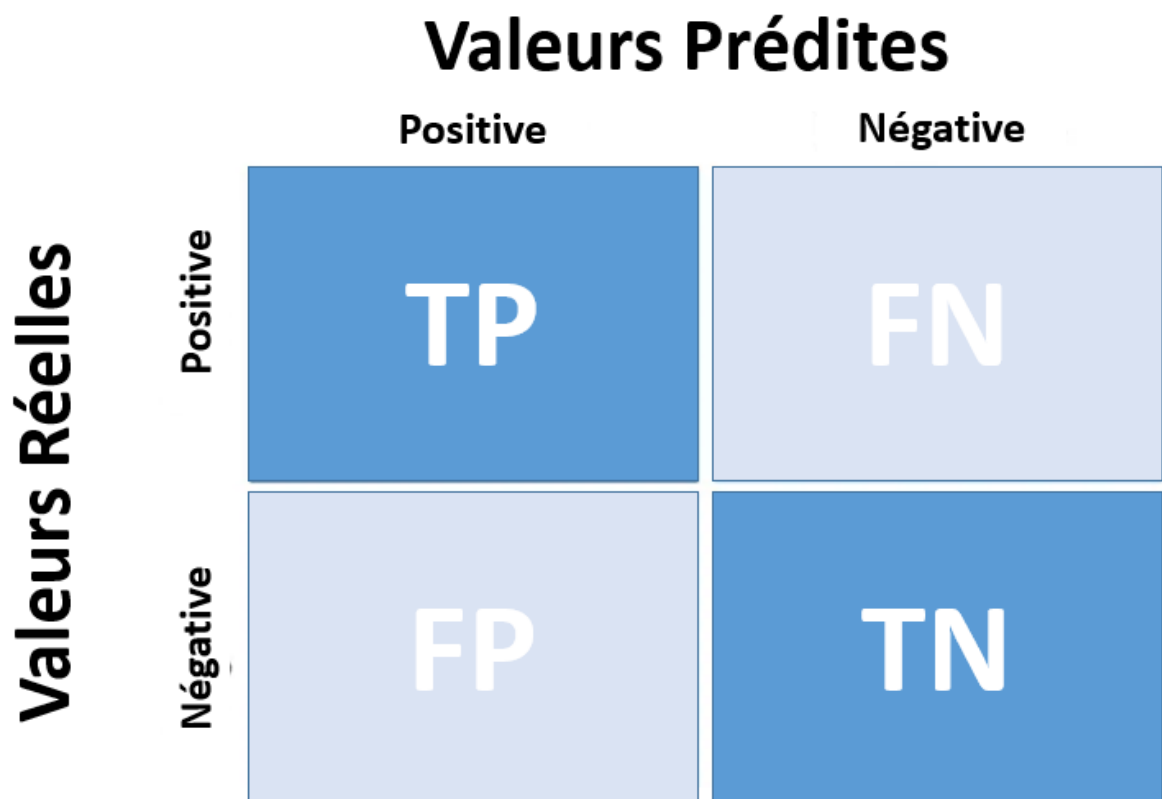


Figure 2.8 : Structure d'une matrice de confusion illustrant les quatre types de prédictions (VP, VN, FP, FN) pour l'évaluation du diagnostic médical.

- **Taux de Classification Correcte (Accuracy)**

Il s'agit du rapport entre le nombre de prédictions correctes (vrais positifs + vrais négatifs) et le nombre total d'instances évaluées [37].

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Où : TP (Vrai positif) ; TN (vrai négatif) ; FP (Faux positif) ; FN (Faux négatif).

- **Sensibilité (Recall) et Spécificité**

Ces deux paramètres sont essentiels pour juger de l'utilité pratique du modèle :

- **La Sensibilité** : Elle mesure l'aptitude du réseau à identifier correctement tous les patients réellement atteints d'une hémorragie [37].

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- **La Spécificité** : Elle évalue la capacité à exclure correctement les sujets sains afin d'éviter des interventions inutiles [37].

- **Précision**

La précision mesure la proportion de prédictions positives correctes par rapport au

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

nombre total de prédictions positives. Elle évalue la qualité des prédictions positives du modèle [37].

- **F1-Score**

Le F1-score mesure la moyenne harmonique de la précision et du rappel, fournissant une évaluation équilibrée des performances du modèle [37].

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Courbe ROC et Aire sous la courbe (AUC)**

Parmi les indicateurs couramment utilisés pour évaluer les performances d'un modèle de classification, la courbe ROC (*Receiver Operating Characteristic*) et l'aire sous cette courbe,

appelée AUC (*Area Under the Curve*), occupent une place importante. Ces outils permettent d'analyser la capacité du modèle à distinguer correctement les différentes classes.

La courbe ROC est construite en représentant le taux de vrais positifs (*True Positive Rate* ou sensibilité) en fonction du taux de faux positifs (*False Positive Rate*), calculé pour plusieurs seuils de décision. L'étude de cette courbe permet d'apprécier le compromis entre la détection des cas positifs et la limitation des erreurs de classification. Plus la courbe se rapproche de l'angle supérieur gauche du graphique, plus le modèle présente une bonne capacité discriminante.

L'AUC correspond à la surface délimitée sous la courbe ROC et fournit un indicateur synthétique des performances globales du classifieur. Sa valeur est comprise entre 0,5 et 1. Une valeur proche de 0,5 traduit un comportement comparable à un choix aléatoire, tandis qu'une valeur proche de 1 indique une excellente aptitude du modèle à différencier les classes étudiées. Contrairement à certaines métriques sensibles à la répartition des données, l'AUC conserve généralement sa pertinence même en présence d'un déséquilibre important entre les classes, ce qui en fait un critère d'évaluation particulièrement robuste [36].

L'utilisation conjointe de la courbe ROC, de l'AUC et des autres mesures de performance permet d'obtenir une analyse plus complète du comportement du modèle. Ces indicateurs ne se limitent pas à mesurer le taux global de bonnes classifications, mais évaluent également l'efficacité du système à détecter correctement les cas positifs tout en réduisant les erreurs liées aux faux positifs et aux faux négatifs. Ils constituent ainsi des outils essentiels pour vérifier la fiabilité, la robustesse et la pertinence des prédictions produites par le modèle.

2.11 Conclusion

Au terme de ce chapitre, les fondements théoriques nécessaires à la compréhension des méthodes modernes de diagnostic assisté par intelligence artificielle ont été présentés. Après avoir introduit les notions d'intelligence artificielle, d'apprentissage automatique et d'apprentissage profond, nous avons étudié les principes de fonctionnement des réseaux de neurones profonds ainsi que les principales stratégies d'entraînement, notamment l'apprentissage par transfert, largement utilisé en imagerie médicale.

Les architectures retenues dans cette étude, à savoir MobileNet, EfficientNet, Vision Transformer (ViT) et Swin Transformer, ont également été détaillées afin de mettre en évidence leurs spécificités, leurs avantages et leur potentiel pour l'analyse automatique des images

médicales. Par ailleurs, les différentes métriques d'évaluation utilisées pour mesurer les performances des modèles de classification ont été présentées, permettant ainsi d'établir une comparaison objective entre les approches étudiées.

L'analyse critique de l'état de l'art a montré que, malgré les excellentes performances obtenues par les modèles de Deep Learning dans la détection et la classification des accidents vasculaires cérébraux, plusieurs défis demeurent, notamment en matière d'interprétabilité, de généralisation et de robustesse clinique. Ces limitations ont conduit à l'émergence de nouvelles approches hybrides intégrant des mécanismes d'attention et des techniques d'intelligence artificielle explicable afin d'améliorer à la fois les performances et la confiance des utilisateurs.

L'ensemble de ces éléments constitue le cadre théorique de notre travail et justifie le choix des architectures étudiées. Le chapitre suivant sera consacré à la méthodologie proposée, en présentant la base de données utilisée, les étapes de prétraitement des images, les modèles implémentés ainsi que le protocole expérimental adopté pour la classification automatique des accidents vasculaires cérébraux à partir d'images de tomodensitométrie cérébrale.

Chapitre 3

Méthodologie Expérimentale

3.1 Introduction

La classification automatisée des accidents vasculaires cérébraux (AVC) et des hémorragies intracrâniennes à partir d'images de tomodensitométrie (scanner CT) constitue un défi majeur pour l'aide au diagnostic médical d'urgence. La réussite d'un tel système repose non seulement sur la puissance des algorithmes d'apprentissage profond, mais également sur la rigueur de la chaîne de traitement des données en amont et sur la mise en place d'un protocole expérimental strict.

Ce chapitre a pour objectif de détailler la méthodologie globale et l'environnement technique mis en œuvre dans le cadre de cette étude afin de garantir une classification fiable, robuste et cliniquement pertinente des images tomographiques.

Pour ce faire, notre démarche est structurée en trois phases principales et complémentaires. Dans un premier temps, nous présenterons l'environnement de développement, les outils logiciels ainsi que les bases de données médicales exploitées. Une attention particulière sera accordée à la phase critique de prétraitement et d'augmentation de données, indispensable pour pallier les artefacts radiologiques et le déséquilibre inhérent aux classes cliniques. Dans un deuxième temps, nous exposerons les fondements théoriques et les configurations des quatre architectures de réseaux de neurones profonds sélectionnées pour leur efficacité (EfficientNet-B0, MobileNet-V2, ViT-B/16 et Swin-T). Enfin, nous détaillerons le protocole expérimental d'entraînement et de validation, avant de définir l'arsenal de métriques statistiques et cliniques retenu pour évaluer les performances de nos modèles.

La figure 3.1 suivante représente le pipeline général du notre travail.

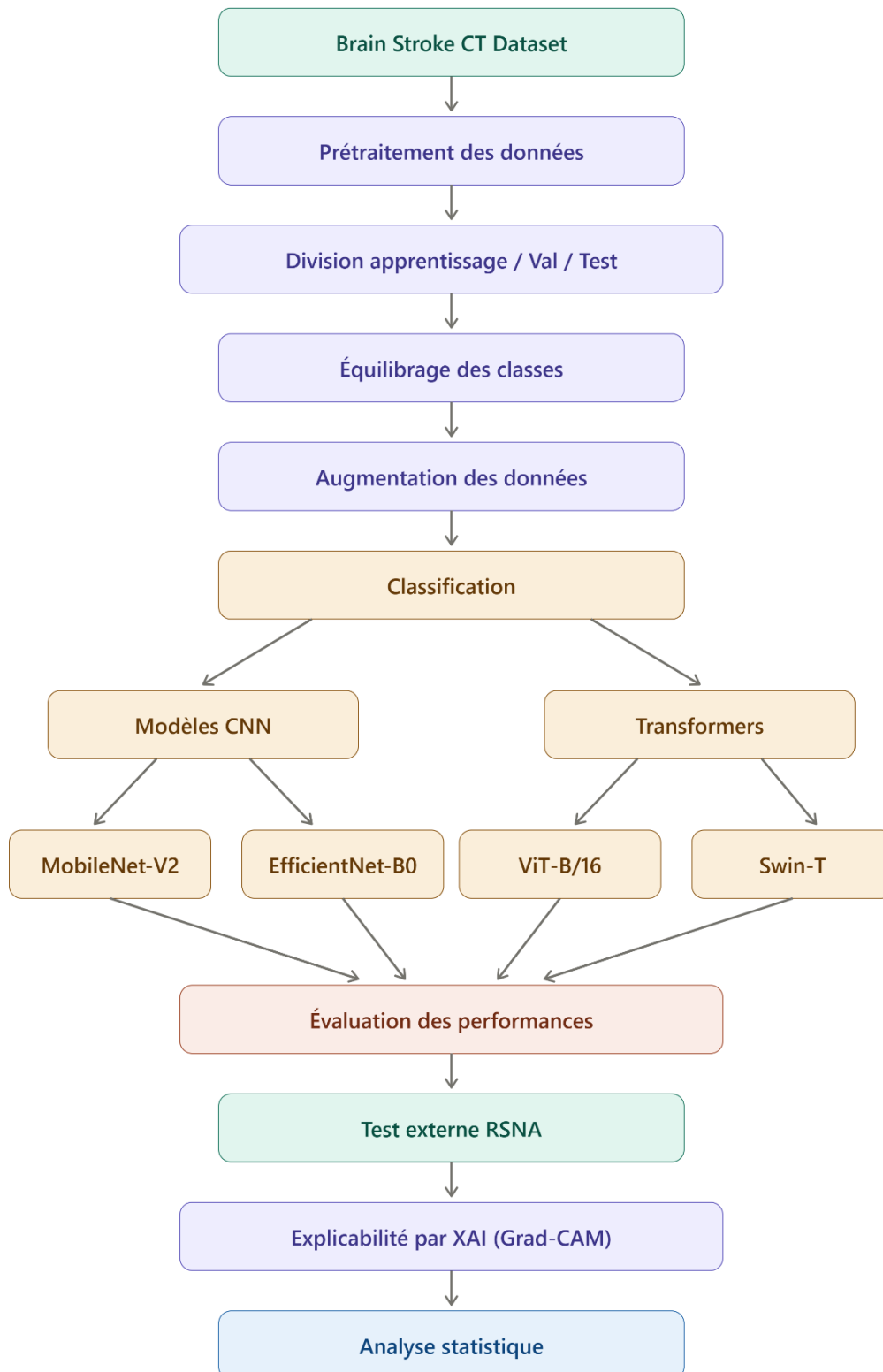


Figure 3.1 : Pipeline général de notre méthodologie.

3.2 Description des bases de données

Dans le cadre de ce travail, deux bases de données publiques de tomodensitométrie (TDM) cérébrale ont été utilisées : le **Brain Stroke CT Dataset**, exploité pour l'entraînement et la validation des modèles de classification des AVC, et le **RSNA Intracranial Hemorrhage Detection Dataset**, réservé exclusivement à l'évaluation externe indépendante afin de mesurer la capacité de généralisation des modèles développés.

3.2.1 Brain Stroke CT Dataset

Afin d'entraîner et d'évaluer les performances des modèles de classification automatique des Accidents Vasculaires Cérébraux (AVC), nous avons exploité le jeu de données public intitulé : **Brain Stroke CT Dataset** [<https://www.kaggle.com/datasets/ozguraslank/brain-stroke-ct-dataset>].

La base de données **Brain Stroke CT Dataset** contient un total de 8873 images de tomodensitométrie (TDM) cérébrale, conçues pour faciliter la détection et la classification automatique des types d'AVC à l'aide de techniques d'intelligence artificielle. Chaque image est annotée par des experts cliniques afin de distinguer les tissus sains des différentes pathologies cérébrales. Les images sont réparties en trois classes distinctes : les cas normaux, les cas d'ischémie et les cas d'hémorragie (Bleeding). [49]

Le tableau 3.1 suivant présente un résumé des trois classes et la répartition des images au sein de la base de données.

Classe	Etat Clinique	Nombre d'images
Normal	Cerveau sain (Absence de lésion)	4427
Ischémie	AVC Ischémique	2260
Hémorragie	AVC Hémorragique (Bleeding)	2186
Total	8873	

Tableau 3.1 : Classification de la base de données Brain Stroke CT Dataset

La figure 3.2 Suivante présente des exemples d'images d'Avc de cette base de données à partir l'image tomодensitométrique.

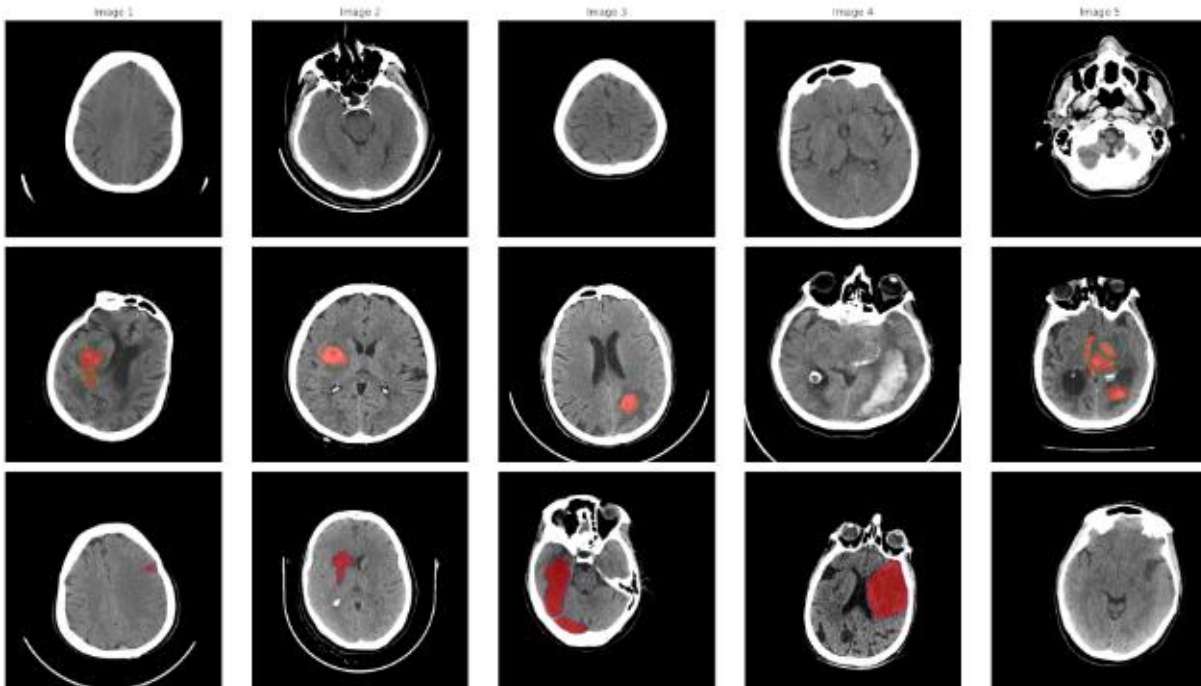


Figure 3.2 : Exemples d'images CT représentant les trois classes de la base Brain Stroke CT Dataset.

3.2.2 RSNA Intracranial Hemorrhage Detection Dataset

Afin d'évaluer la capacité de généralisation des modèles entraînés sur le **Brain Stroke CT Dataset**, nous avons conduit une phase de **test externe indépendant** en exploitant le jeu de données public : **RSNA Intracranial Hemorrhage Detection Dataset** [<https://www.kaggle.com/competitions/rsna-intracranial-hemorrhage-detection>].

La base de données RSNA Intracranial Hemorrhage Detection Dataset contient un total de **752 803 images** de tomодensitométrie (TDM) cérébrale au format DICOM, constituant l'un des plus grands corpus publics disponibles pour la détection automatique des hémorragies intracrâniennes. Chaque image a été annotée par des radiologues experts afin de distinguer les coupes cérébrales normales des différentes formes d'hémorragie intracrânienne. [50]

Le tableau 3.3 suivant présente un résumé des classes et la répartition des images au sein de la base de données.

Classe	Etat Clinique	Nombre d'images
Normal	Cerveau sain (Absence de lésion)	644 874
Hémorragie	Présence d'au moins un type d'hémorragie	107 933
Hémorragie Épidurale	Collection sanguine entre la dure-mère et la boîte crânienne	3 800
Hémorragie Sous-Durale	Collection sanguine entre la dure-mère et l'arachnoïde	22 400
Hémorragie Sous-Arachnoïdienne	Saignement dans l'espace sous-arachnoïdien	35 200
Hémorragie Intraparenchymateuses	Saignement à l'intérieur du tissu cérébral	37 600
Hémorragie Intraventriculaire	Saignement dans les ventricules cérébraux	16 100
Total	867 907	

Tableau 3.2 : Répartition des classes du dataset RSNA Intracranial Hemorrhage Detection utilisé pour le test externe.

La figure 3.3 suivante montre des exemples de coupes tomодensitométriques cérébrales du dataset RSNA.

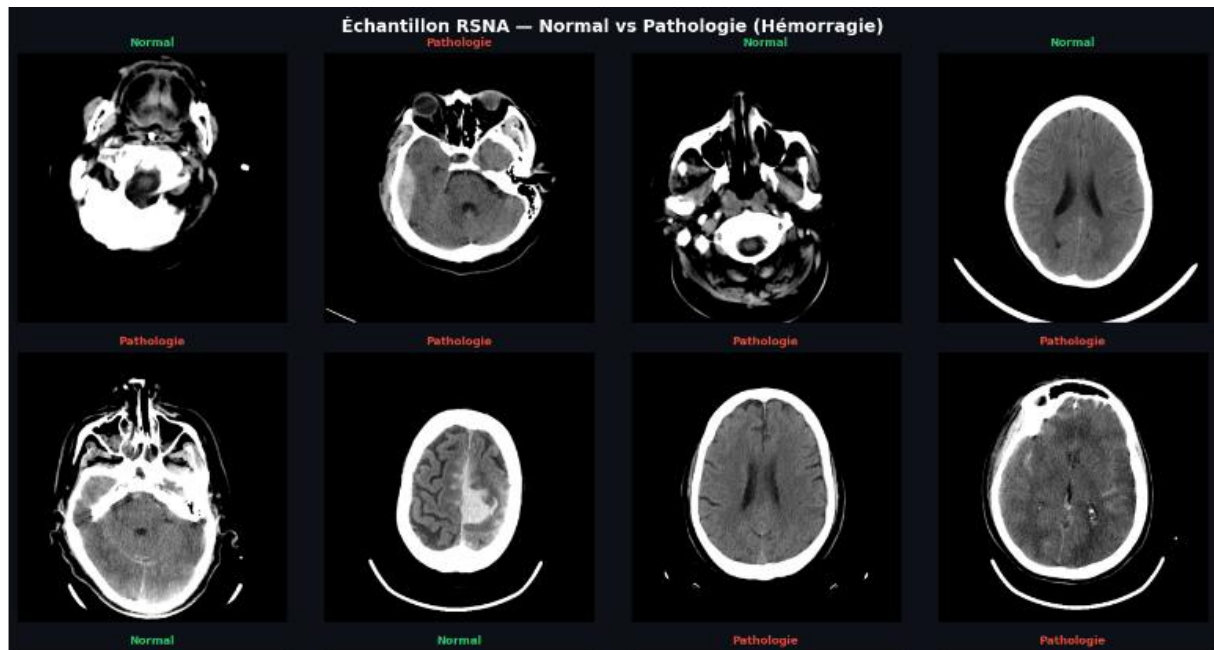


Figure 3.3 : Exemples d’images CT de la base RSNA Intracranial Hemorrhage Detection.

3.3 Méthodes proposées

Cette partie présente la méthodologie de prétraitement adoptée dans le cadre de cette étude afin d’améliorer la qualité des images médicales et d’optimiser les performances des modèles de Deep learning. L’approche proposée repose sur plusieurs étapes complémentaires incluant le nettoyage morphologique des images CT, le redimensionnement, la normalisation des données ainsi que la préparation des images pour l’apprentissage automatique. Chaque étape joue un rôle essentiel dans la réduction du bruit, l’amélioration de l’homogénéité des données et l’extraction efficace des caractéristiques anatomiques pertinentes nécessaires à une classification précise et fiable.

- L’organigramme présenté dans la figure 3.4 illustre les différentes étapes mises en œuvre dans le cadre de notre projet.



Figure 3. 4 : Organigramme des différentes phases de notre approche.

3.3.1 Prétraitement des images CT

La qualité des données constitue un facteur essentiel influençant directement les performances des modèles de Deep learning. Pour cette raison, une phase de prétraitement rigoureuse a été appliquée aux images CT cérébrales afin d'améliorer leur qualité, de réduire le bruit et de faciliter l'extraction des caractéristiques anatomiques pertinentes. Ce processus comprend plusieurs étapes successives allant du nettoyage morphologique des images jusqu'à leur normalisation, dans le but d'obtenir un jeu de données homogène et optimisé pour l'apprentissage automatique.

Dans un premier temps, chaque image CT subit une étape de nettoyage morphologique avancé afin de supprimer les informations non pertinentes telles que les zones noires périphériques, les artefacts d'acquisition et les structures situées en dehors du crâne. Cette étape permet aux modèles de se concentrer uniquement sur les régions cérébrales importantes pour la classification.

Le processus débute par la conversion des images en niveaux de gris afin de simplifier les traitements morphologiques et de faciliter l'analyse des intensités anatomiques. Ensuite, un seuillage binaire avec un seuil fixé à 70 est appliqué pour séparer automatiquement la région cérébrale du fond noir et des zones parasites. Afin d'améliorer la qualité du masque obtenu, une opération d'érosion morphologique utilisant un noyau elliptique de taille 5×5 est réalisée pour éliminer les petits artefacts et les connexions indésirables autour du cerveau.

Par la suite, tous les contours présents dans l'image sont détectés, puis seul le contour possédant la plus grande surface est conservé puisqu'il correspond généralement à la boîte crânienne. À partir de ce contour principal, un masque binaire représentant la région cérébrale est construit puis légèrement dilaté afin de préserver correctement les contours anatomiques importants et éviter la perte d'informations utiles.

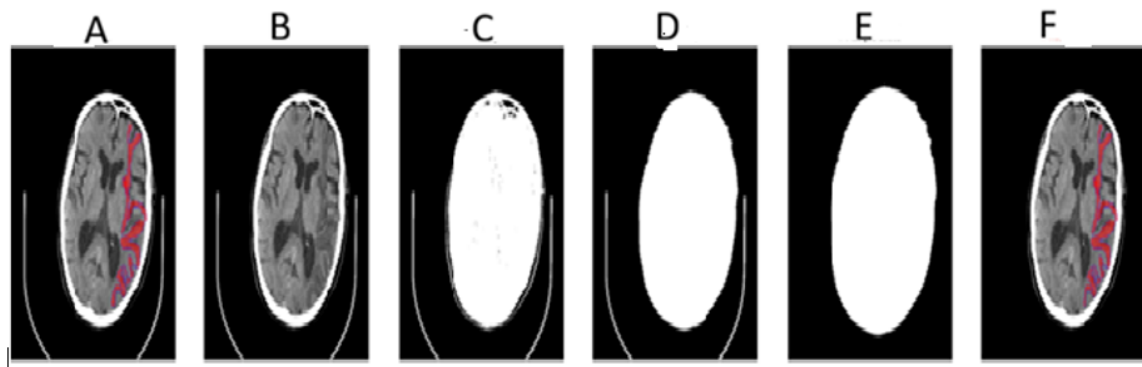
Enfin, ce masque est appliqué à l'image originale afin d'effectuer une opération de skull stripping, permettant de supprimer complètement le fond noir ainsi que les structures externes non pertinentes tout en conservant uniquement la région cérébrale utile à la classification. Cette procédure améliore l'homogénéité des images et réduit l'influence des informations inutiles sur l'apprentissage des modèles.

Après le nettoyage morphologique, toutes les images sont redimensionnées à une résolution uniforme de 224×224 pixels, correspondant à la taille d'entrée requise par les

architectures de Deep learning utilisées dans cette étude. Cette uniformisation garantit une cohérence dans le traitement des données et facilite l'apprentissage des réseaux neuronaux.

Une étape de normalisation est ensuite appliquée afin d'harmoniser les valeurs numériques des pixels. Les intensités initialement comprises entre 0 et 255 sont converties vers l'intervalle $[0,1]$, ce qui améliore la stabilité du calcul des gradients et favorise une convergence plus rapide pendant l'entraînement. Une standardisation supplémentaire basée sur les statistiques du jeu de données ImageNet est ensuite réalisée à l'aide d'une moyenne de $[0.485, 0.456, 0.406]$ et d'un écart-type de $[0.229, 0.224, 0.225]$ pour les trois canaux RGB. Cette étape permet d'aligner la distribution des données médicales avec celle utilisée lors du pré-entraînement des modèles, améliorant ainsi l'efficacité du transfert d'apprentissage et les performances globales de classification.

La Figure 3.5 ci-dessous montre les étapes de prétraitement des données pour les classes cliniques Bleeding, Ischemia et Normal.



A : image original ; **B** : Niveaux de gris ; **C** : Seuillage ; **D** : Morphologie ; **E** : Masque dilaté ; **F** : image final

Figure 3. 5 : Pipeline de prétraitement appliqué aux images CT.

Cette figure illustre les différentes étapes du pipeline de prétraitement appliqué aux images CT, incluant la conversion en niveaux de gris, le seuillage, les opérations morphologiques, la création du masque cérébral et le nettoyage de l'image finale.

La figure 3.6 présente une comparaison avant et après le prétraitement.

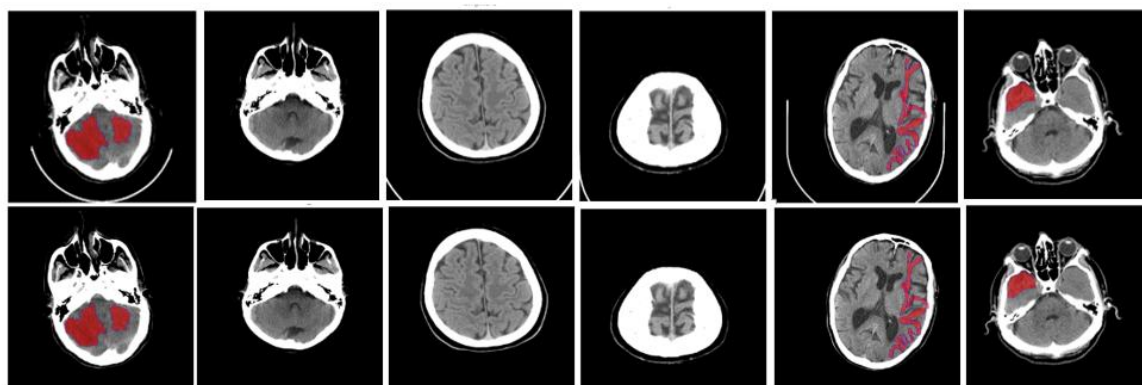


Figure 3. 6 : Comparaison avant/après prétraitement.

Cette figure compare les images CT avant et après prétraitement afin de visualiser l'impact des étapes appliquées sur la qualité et la focalisation des régions importantes.

3.3.2 Division des données

Après la préparation du dataset, les images ont été réparties en trois ensembles selon une division de type 80 % train / 10 % validation / 10 % test. Cette étape permet d'entraîner le modèle puis d'évaluer ses performances de manière plus fiable.

L'ensemble Train est utilisé pour l'apprentissage du modèle, l'ensemble Validation sert à suivre les performances pendant l'entraînement et à ajuster certains paramètres, tandis que l'ensemble Test est réservé à l'évaluation finale. Afin de garder une répartition cohérente des classes dans chaque ensemble, une division stratifiée (stratified split) a été utilisée. Le même protocole a été appliqué aux cas multiclasse et binaires. Les deux figures (Figure 3. 7)

(Figure 3. 8) représentent les trois ensembles de division.

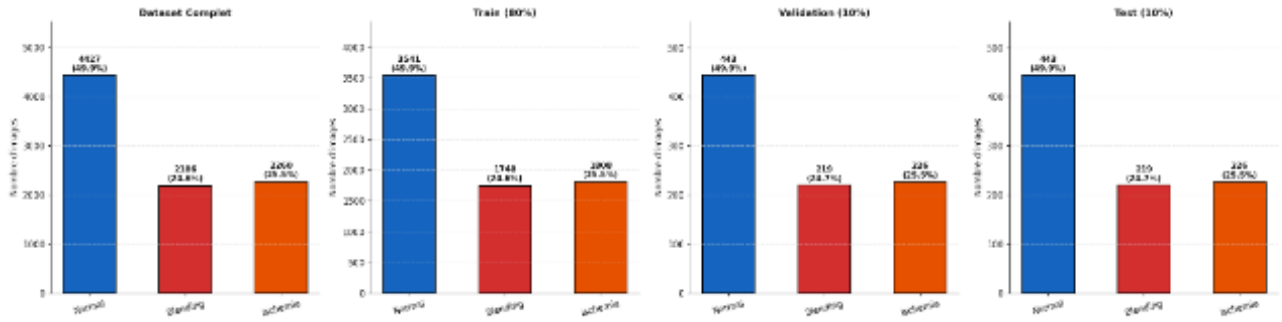


Figure 3. 7 : Répartition des données pour la classification multiclasse dans les ensembles d'entraînement, de validation et de test.

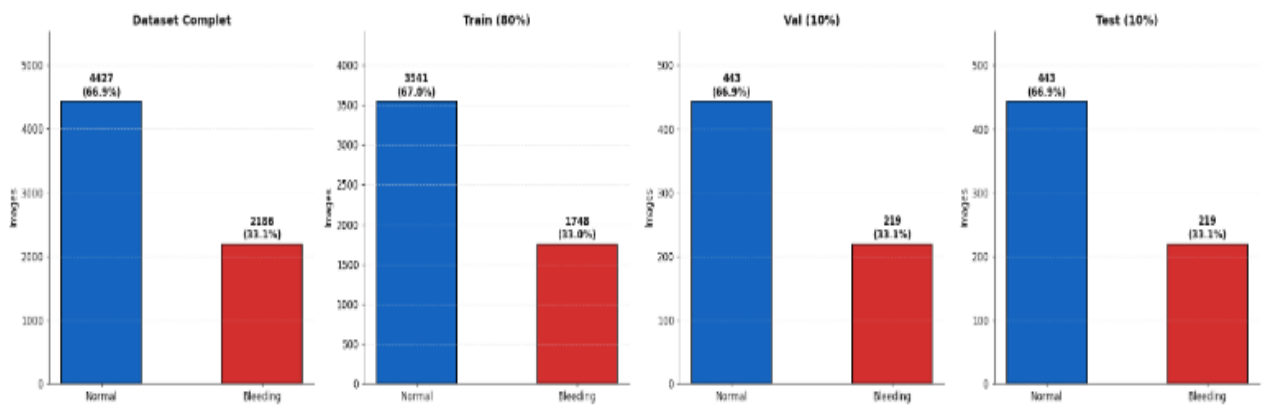


Figure 3. 8 : Répartition des données pour la classification binaire dans les ensembles d'entraînement, de validation et de test.

3.3.3 Équilibrage des données

Après la division du jeu de données en ensembles d'entraînement, de validation et de test, une étape d'équilibrage a été appliquée à l'ensemble d'entraînement afin de réduire l'effet du déséquilibre entre les classes. Cette étape a été réalisée aussi bien dans le cas de la classification multiclasse (Normal, Bleeding et Ischemia) que dans le cas de la classification binaire (Normal et Pathologie), (Normal et Ischémie), (Normal et Bleeding).

Dans cette étude, la méthode `WeightedRandomSampler` a été utilisée afin d'assurer une meilleure représentation des classes pendant l'apprentissage. Cette méthode attribue un poids plus important aux classes les moins représentées et un poids plus faible aux classes majoritaires, ce qui permet au modèle d'apprendre de manière plus équilibrée.

L'équilibrage a été appliqué uniquement à l'ensemble d'entraînement, tandis que les ensembles de validation et de test ont été conservés sans modification afin de garantir une évaluation plus réaliste et plus fiable des performances du modèle.

Les deux figures suivantes (Figure 3. 9) (Figure 3. 10) représentent des exemples de distribution des classes (multiclasse, binaire) avant et après l'équilibrage.

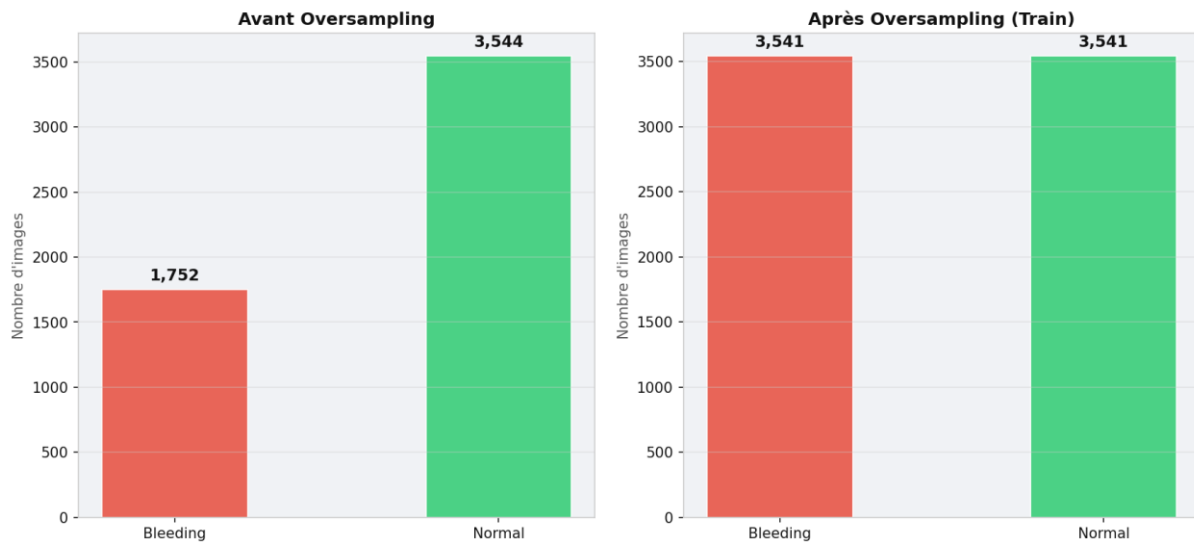


Figure 3.9 : distribution des trois classes avant et après l'équilibrage.

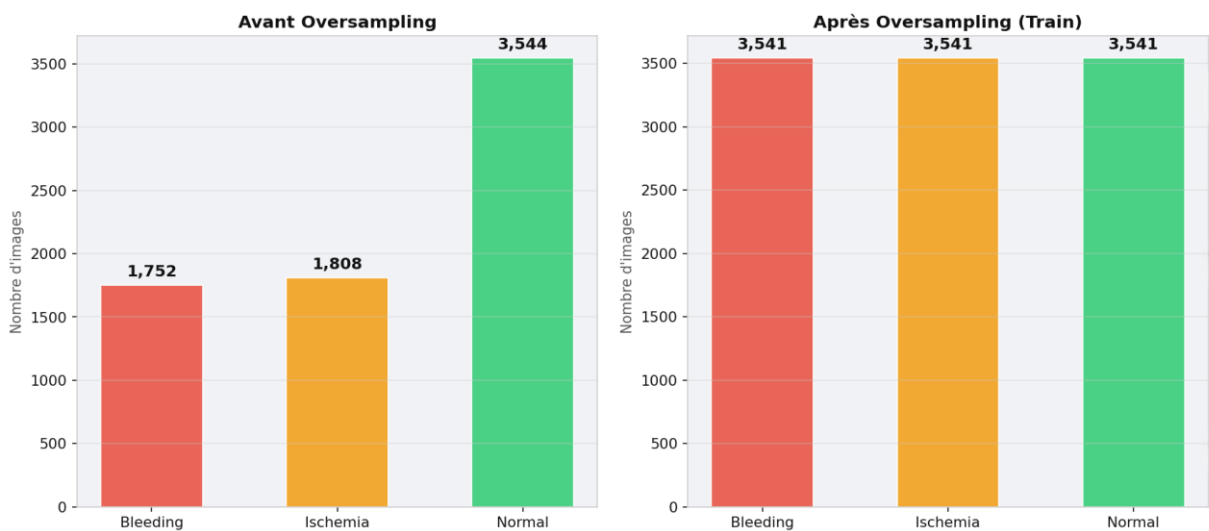


Figure 3.10 : distribution des deux classes avant et après l'équilibrage.

3.3.4 Augmentation des données

Afin d'améliorer la capacité de généralisation des modèles et de limiter le risque de surapprentissage (overfitting), une technique d'augmentation de données a été appliquée exclusivement à l'ensemble d'entraînement. Cette étape consiste à générer dynamiquement, durant le processus d'apprentissage, des variantes artificielles des images originales tout en préservant leurs caractéristiques diagnostiques essentielles. Comme illustré sur la Figure 3.10, nous avons mis en œuvre une série de transformations aléatoires, combinant des modifications géométriques et photométriques.

Ces transformations incluent des retournements horizontaux et verticaux pour varier les orientations anatomiques, des rotations contrôlées (jusqu'à 20°) pour simuler les variations de positionnement du patient, ainsi que des transformations affines (translations et zooms) pour renforcer la résilience spatiale. Par ailleurs, des ajustements de contraste, de luminosité et, occasionnellement, des conversions en niveaux de gris ont été intégrés afin d'accroître la robustesse des modèles face à l'hétérogénéité des équipements d'imagerie et aux variations de calibration. En exposant le modèle à ces versions modifiées lors de chaque cycle d'entraînement, nous favorisons une meilleure compréhension des structures cérébrales, garantissant ainsi une performance plus stable et fiable face à de nouvelles données cliniques.

Toutes les transformations ont été appliquées de manière contrôlée afin de préserver la cohérence anatomique des structures cérébrales.

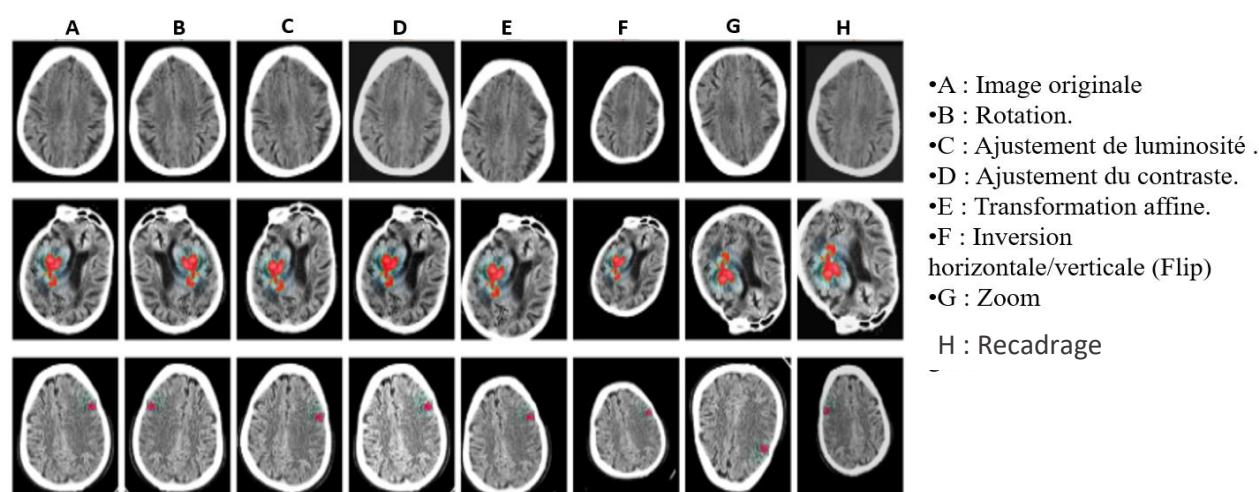


Figure 3. 11 : Exemples d'images après l'augmentation des données.

3.4 Classification

Pour répondre aux objectifs de notre recherche, nous avons segmenté notre corpus de données en deux types d'expérimentations distinctes : une classification multiclasse et une classification binaire.

a) Classification multiclasse

Dans un premier temps, une classification multiclasse a été réalisée afin de distinguer trois catégories d'images cérébrales CT : Normal, Bleeding et Ischemia. Le but de cette étape est d'apprendre au modèle à reconnaître la différence entre un cerveau normal, un cas d'hémorragie cérébrale et un cas d'ischémie. Chaque image est associée à une seule classe selon

son diagnostic. Le tableau 3. 3 suivant présente les différentes classes de BDD en classification multiclasse.

Cas	Classes
Multiclasse	Normal
	Bleeding
	Ischemia

Tableau 3. 3 : les différentes classes de BDD en classification multiclasse.

b) Classification binaire

Après la classification multiclasse, plusieurs expériences de classification binaire ont été effectuées afin d'évaluer le comportement du modèle dans des situations plus spécifiques. Dans ce cas, la classification est réalisée entre seulement deux classes. Le tableau 3. 4 suivant présente les deux classes de BDD en classification binaire.

Cas	Classes
Binaire 1	Normal
	Pathologie
Binaire 2	Normal
	Ischemia
Binaire 3	Normal
	Bleeding

Tableau 3. 4 : les deux classes de BDD en classification binaire.

Dans le premier cas binaire (Normal / Pathologie), les classes Bleeding et Ischemia sont regroupées dans une seule catégorie appelée Pathologie. Cela permet d'évaluer si le modèle peut distinguer un cerveau normal d'un cerveau présentant une anomalie.

3.5 Modèles d'apprentissage profond utilisés

Cette section se concentre sur l'évaluation et la description des quatre architectures de réseaux de neurones profonds que nous avons développées et configurées pour la classification automatique des accidents vasculaires cérébraux. Notre approche repose sur le paradigme de

l'apprentissage par transfert (*Transfer Learning*) appliqué à quatre modèles de l'état de l'art, reconnus pour leur efficacité dans les tâches de vision par ordinateur : EfficientNet-B0, MobileNet-V2, ViT-B/16 et Swin-T. Structurellement, notre démarche s'articule autour d'un extracteur de caractéristiques pré-entraîné et maintenu initialement figé, auquel nous avons greffé une tête de classification personnalisée, redimensionnée et optimisée par nos soins. Cette configuration sur mesure a pour objectif de discriminer efficacement nos trois catégories cliniques cibles à partir d'images de scanner CT cérébral : les accidents hémorragiques (Bleeding), ischémiques (Ischemia) ainsi que les cas sains (Normal).

Le Tableau 3. 5 ci-dessous, montre pourquoi nous avons choisi précisément ces modèles de classification.

Modèle	Type	Motivation
MobileNetV2	CNN léger	Rapidité
EfficientNet-B0	CNN optimisé	Excellent compromis précision/coût
ViT-B16	Transformer	Vision globale
Swin-T	Transformer hiérarchique	Analyse multi-échelle

Tableau 3.5 : Motivation du choix des modèles de DL.

3.5.1 Architecture EfficientNet-B0

Pour l'analyse automatique des scanners cérébraux, nous avons mis en œuvre une architecture EfficientNet-B0 spécialisée, dont nous avons optimisé la structure pour répondre aux exigences complexes de la détection des hémorragies. Ce framework repose sur un extracteur de caractéristiques (backbone) hybride : les couches initiales sont gelées pour préserver les détecteurs de textures fondamentaux appris sur des bases de données massives, tandis que les couches supérieures sont soumises à un fine-tuning ciblé pour s'adapter aux signatures radiologiques spécifiques des images CT. Par rapport à l'architecture originale, nous avons substitué la tête de classification standard par une structure dense à trois niveaux (512 et 256 neurones) et avons introduit une régularisation par couches de Dropout progressives (0,4, 0,3 et 0,2). Cette modification majeure permet non seulement d'accroître la profondeur d'abstraction des caractéristiques, mais également de stabiliser l'apprentissage et de prévenir le surapprentissage, un risque critique dans le domaine médical. Cette architecture modulaire, associée à notre prétraitement par fenêtrage multicanal (Brain, Subdural, Bone), permet de

transformer efficacement les données brutes DICOM en un outil diagnostique robuste, capable de générer des scores de probabilité fiables tout en garantissant une cohérence clinique rigoureuse dans l'analyse des lésions ciblées. La (Figure 3.11) ci-dessous expose l'agencement de notre modèle Efficientnet-B0.

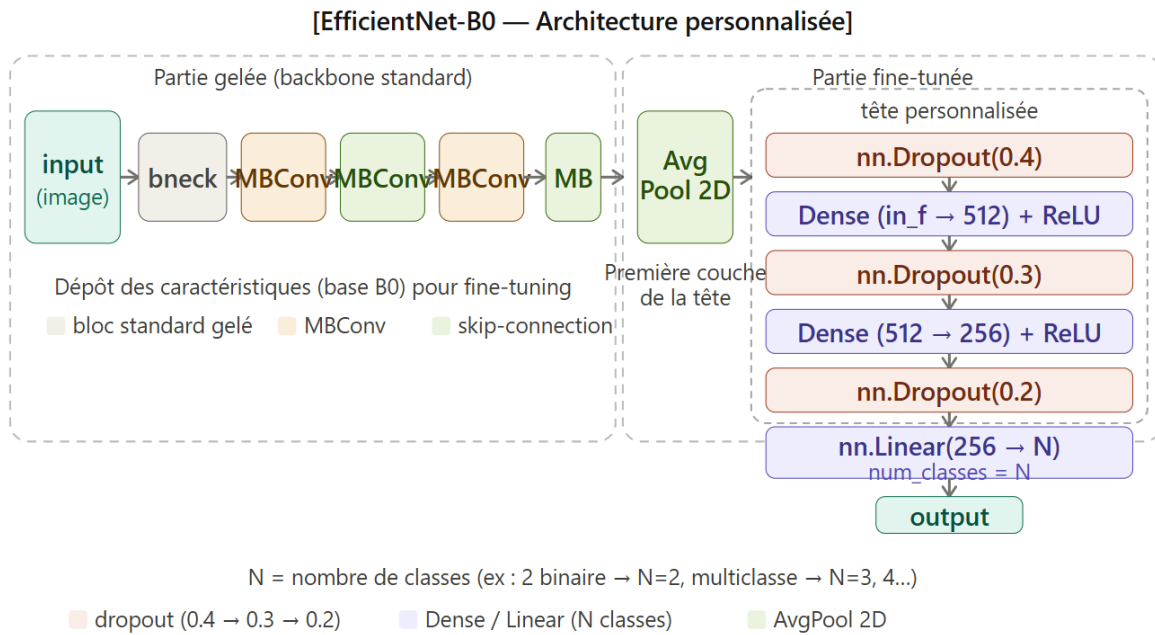


Figure 3.12 : Notre architecture pour EfficientNet-B0.

3.5.2 Architecture MobileNet-V2

Pour l'analyse des scanners cérébraux, nous avons exploité l'architecture MobileNet-V2, reconnue pour sa légèreté et sa vitesse d'inférence élevée. Ce modèle repose sur une conception optimisée pour les environnements à ressources limitées, utilisant des blocs Inverted Residuals et des Linear Bottlenecks. Chaque bloc intègre une convolution d'expansion 1×1 , une convolution séparable en profondeur (Depthwise) 3×3 , et une projection linéaire 1×1 , permettant de réduire considérablement la complexité computationnelle tout en préservant l'intégrité morphologique des descripteurs grâce aux connexions résiduelles.

Dans notre implémentation, nous avons conservé le réseau de base (backbone) figé afin d'exploiter ses capacités d'extraction, tout en apportant une refonte majeure de la structure aval pour l'adapter à nos contraintes cliniques. Nous avons éliminé la couche de classification native pour y greffer une tête décisionnelle sur mesure. Après une agrégation globale (AvgPool2D), le flux transite par un bloc résiduel inversé additionnel, spécifiquement configuré pour raffiner les canaux latents (compression de 11 vers 9 dimensions). Cette compression

sémantique, couplée à une régularisation par Dropout (0,2), assure une robustesse accrue avant que la couche terminale ne projette les données vers l'output.

Cette architecture personnalisée, totalisant environ 3,4 millions de paramètres, offre un compromis optimal entre rapidité et performance. Elle est conçue pour être polyvalente : sans altérer l'extracteur de caractéristiques, le modèle s'adapte automatiquement aux tâches de classification binaire ou multiclasse en ajustant le nombre de neurones de la couche de sortie, garantissant ainsi une classification performante des pathologies hémorragiques, comme illustré dans (la Figure 3. 12).

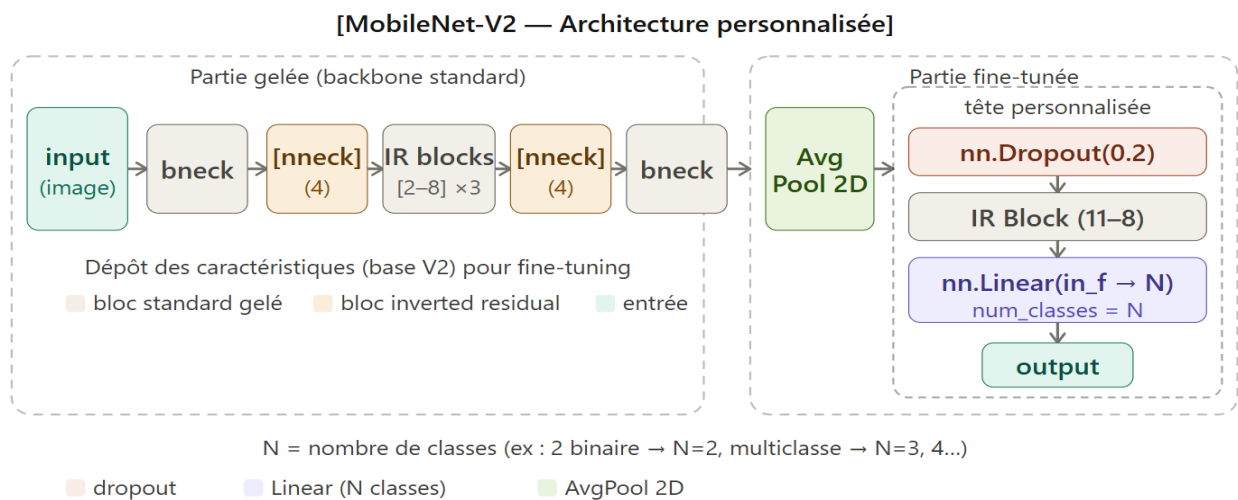


Figure 3.13 : Notre architecture pour MobileNet-V2.

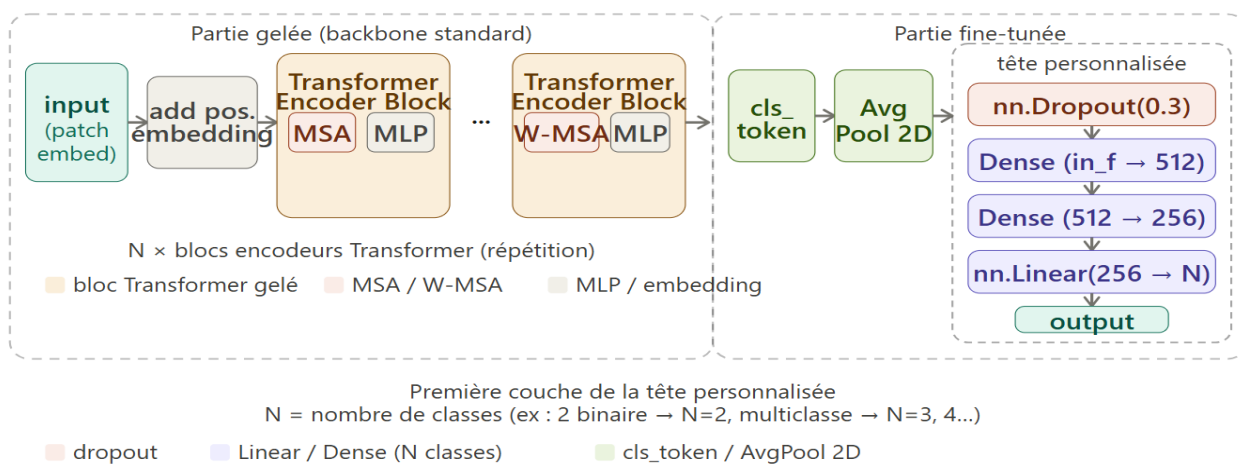
3.5.3 Architecture ViT-B/16 (Vision Transformer)

Pour l'analyse des images médicales complexes, nous avons intégré l'architecture ViT-B/16 (Vision Transformer), qui traite l'image comme une séquence de segments visuels (patches de 16×16 pixels). Contrairement aux réseaux de neurones convolutifs (CNN), ce modèle remplace les convolutions traditionnelles par des mécanismes d'auto-attention globale (MSA), permettant de modéliser les interactions anatomiques à longue distance dès les premières couches. Le flux de traitement segmente l'image en une grille de patches, enrichis par un encodage de position absolue, avant de les soumettre aux 12 blocs du Transformer Encoder. Par rapport à l'architecture originale, qui se termine par une simple couche linéaire, nous avons intégralement refondu la tête de classification pour imposer une logique décisionnelle spécialisée.

Nous avons isolé le jeton de classification (cls_token), qui synthétise le contexte global de la scène médicale, et nous l'avons enrichi par une couche d'agrégation globale (AvgPool2D) agissant comme un filtre de lissage pour stabiliser le signal. Ce flux est ensuite transmis à une

tête personnalisée, structurée en deux niveaux de couches denses (512 et 256 neurones) et sécurisée par une couche de Dropout (0,3) pour prévenir le surapprentissage. Cette architecture modifiée conserve sa polyvalence : elle s'adapte dynamiquement à une classification binaire ou multiclasse en ajustant le nombre de neurones de la couche de sortie, tout en laissant l'extracteur de contexte (les blocs Transformer) figé. Grâce à sa capacité à modéliser les dépendances spatiales globales, ce ViT-B/16 personnalisé constitue un outil puissant pour l'analyse des images médicales où la compréhension étendue du contexte visuel est déterminante pour la précision diagnostique. Comme détaillé dans la (Figure 3.13).

Figure 3.14 : Notre architecture pour ViT-B/16
[ViT-B/16 — Architecture personnalisée]



3.5.4 Architecture Swin-T (Swin Transformer)

Pour l'analyse des scanners cérébraux, nous avons intégré le Swin Transformer (Swin-T), une architecture hiérarchique qui optimise les Vision Transformers. Contrairement à une attention globale, le Swin-T utilise des mécanismes d'auto-attention calculés au sein de fenêtres locales décalées (W-MSA), permettant de capturer des relations contextuelles multi-échelles avec une complexité computationnelle linéaire. Le modèle segmente l'image en patches qui sont progressivement fusionnés à travers trois stades hiérarchiques, combinant ainsi la précision des détails locaux et la compréhension globale du contexte visuel.

Dans notre pipeline, après le partitionnement initial, le flux transite par les stades hiérarchiques où l'action combinée des fenêtres d'attention et des perceptrons (MLP) affine la représentation. Par rapport à l'architecture Swin-T originale, qui se termine par une tête de classification standard très superficielle, nous avons opéré une refonte majeure de la structure aval pour imposer une logique décisionnelle spécialisée. Après une agrégation globale (AvgPool/GAP), nous avons inséré une couche de Dropout (0,3) pour prévenir le

surapprentissage, suivie d'une structure dense à deux niveaux (512 et 256 neurones) avant la couche finale.

Le choix de cette architecture dense à deux niveaux ($512 \rightarrow 256$) est stratégique : elle permet une compression sémantique progressive, où le réseau "résume" les informations complexes extraites par le backbone pour les filtrer vers une décision finale précise. Cette profondeur supplémentaire permet d'apprendre des relations non-linéaires plus fines entre les descripteurs extraits et le diagnostic, tout en offrant les points d'ancrage nécessaires pour nos couches de régularisation. Cette personnalisation permet de projeter plus finement l'espace latent enrichi vers l'output, assurant une corrélation directe entre les structures anatomiques globales et les foyers pathologiques locaux. À l'instar des autres architectures, ce modèle demeure polyvalent : il s'adapte automatiquement aux tâches de classification binaire ou multiclasse en ajustant le nombre de neurones de la couche terminale (N), sans altérer l'extracteur de caractéristiques. Grâce à sa capacité unique à gérer les détails multi-échelles, le Swin-T ainsi modifié constitue un outil performant pour l'identification précise des classes cliniques, comme le détaille la (Figure 3. 14).

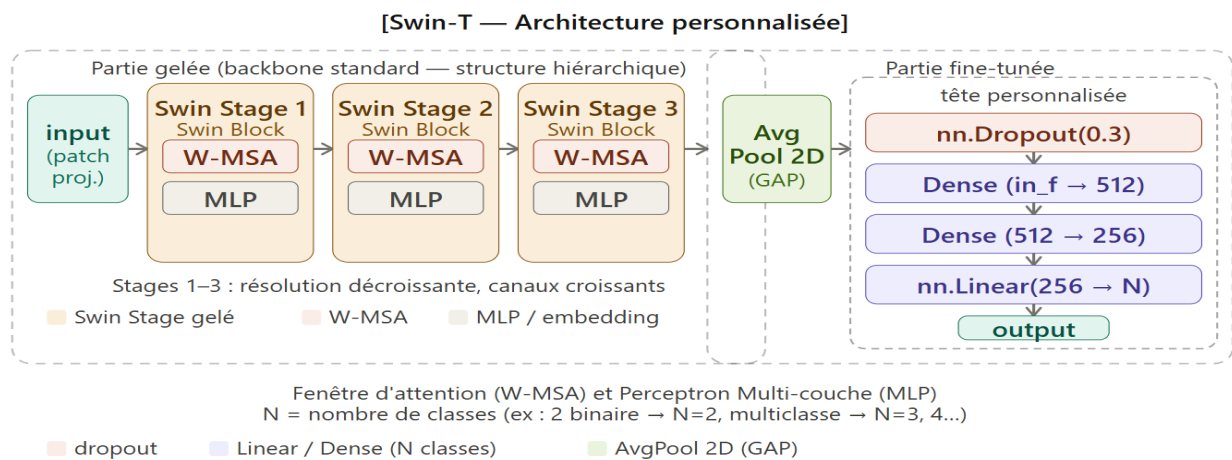


Figure 3.15 : Notre architecture pour Swin-T.

3.6 Hyperparamètres d'Optimisation

Afin de stabiliser la convergence des réseaux de neurones, d'optimiser l'extraction des caractéristiques cliniques et de garantir une meilleure généralisation face à la variabilité des coupes scénographiques, les hyperparamètres ont été fixés de manière rigoureuse au sein de notre environnement de développement.

Le tableau 4.6 récapitule les variables et hyperparamètres optimaux retenus pour L'entraînement de nos quatre architectures personnalisées.

Hyperparamètres	Configuration logicielle / Valeur
Taille du lot	16
Résolution des entrées	224 × 224 pixels
Taux d'apprentissage (Learning Rate)	1×10^{-4} (1e-4)
Nombre maximum d'époques	30
Algorithme d'optimisation	Adam
Fonction de perte	Entropie Croisée
Graine aléatoire	42

Tableau 3.6 : Configuration synthétique des hyperparamètres d'entraînement.

La reproductibilité et l'optimisation de la phase d'apprentissage reposent sur une sélection rigoureuse et standardisée des hyperparamètres d'entraînement, synthétisés dans le tableau précédent. Afin de garantir l'équilibre optimal entre l'efficacité computationnelle et la finesse des mises à jour des poids du réseau, l'apprentissage est orchestré avec une taille de lot (batch size) fixée à 16. L'optimisation des modèles s'appuie sur l'algorithme Adam, largement plébiscité pour ses capacités de convergence adaptative en imagerie médicale. Le taux d'apprentissage initial est configuré de manière conservatrice à 10^{-4} (1e-4) sur un horizon maximal de 30 époques, un réglage stratégique visant à opérer un fine-tuning progressif sans altérer les descripteurs visuels préalablement acquis par les architectures pré-entraînées. La fonction de perte de l'Entropie Croisée est exploitée comme indicateur de minimisation de l'erreur pour la classification. En amont, toutes les données tomographiques subissent un redimensionnement uniforme à une résolution d'entrée de 224×224 pixels pour s'aligner nativement sur les exigences structurelles des réseaux testés. Enfin, la fixation d'une graine aléatoire à la valeur 42 verrouille l'initialisation et le partitionnement des données, scellant ainsi la stricte reproductibilité scientifique de notre protocole expérimental.

Le tableau 3.7 suivant présente les hyperparamètres pour chaque modèle étudié.

Modèle	Backbone gelé	Fine-Tuning	Dropout
MobileNet	Oui	Partiel	0.4
EfficientNet	Oui	Partiel	0.4
ViT	Oui	Partiel	0.4
Swin-T	Oui	Partiel	0.4

Tableau 3.7 : Hyperparamètres par modèle.

3.7 Métriques d'évaluation des modèles

Pour évaluer les performances des quatre architectures de DL (EfficientNet-B0, MobileNet-V2, ViT-B/16 et Swin-T), plusieurs métriques complémentaires ont été employées. Au cours de l'entraînement, la perte (Loss) par cross-entropie et l'Accuracy sont suivies à chaque époque sur les ensembles d'entraînement et de validation, permettant de détecter le surapprentissage et de conserver le meilleur modèle. Sur l'ensemble de test, la matrice de confusion constitue le point de départ de l'évaluation : il s'agit d'un tableau carré de dimension 3×3 (une ligne et une colonne par classe : Bleeding, Ischemia, Normal) dont chaque cellule (i, j) indique le nombre d'images appartenant réellement à la classe i et prédites comme appartenant à la classe j ; la diagonale principale regroupe ainsi les prédictions correctes, tandis que les cases hors diagonale représentent les erreurs de classification. À partir de cette matrice sont dérivées, pour chaque classe, la sensibilité (taux de vrais positifs), la spécificité (taux de vrais négatifs), la précision, le F1-score (moyenne harmonique de la précision et du rappel), ainsi que le nombre de faux négatifs (FN) et de faux positifs (FP). Ces métriques par classe sont agrégées en une moyenne macro pour comparer globalement les modèles. Enfin, les courbes ROC et l'AUC (aire sous la courbe) sont calculés selon une stratégie One-vs-Rest pour chacune des trois classes, ainsi qu'une AUC macro moyenne, offrant une évaluation de la capacité discriminante des modèles indépendamment du seuil de décision.

3.8 Explicabilité des modèles

Malgré leurs excellentes performances en classification d'images médicales, les modèles d'apprentissage profond demeurent souvent difficiles à interpréter en raison de leur structure complexe. Cette absence de transparence leur vaut fréquemment l'appellation de « boîtes noires », car il est difficile de comprendre précisément quels éléments ont conduit le modèle à prendre une décision donnée. Dans le domaine médical, où les décisions diagnostiques

doivent être fiables, explicables et validées par les professionnels de santé, cette problématique constitue un enjeu majeur.

Afin de réduire cette opacité, le domaine de l'Intelligence Artificielle Explicable (IAX) propose plusieurs méthodes permettant de rendre les prédictions des modèles plus compréhensibles. Ces approches offrent des explications visuelles ou analytiques qui mettent en évidence les régions importantes ayant influencé la décision du réseau neuronal, améliorant ainsi la confiance des médecins dans les systèmes d'aide au diagnostic basés sur l'intelligence artificielle.

Dans ce travail, nous avons utilisé la méthode Grad-CAM (Gradient-weighted Class Activation Mapping), particulièrement adaptée aux modèles de vision par ordinateur. Cette technique génère des cartes de chaleur (heatmaps) à partir des gradients calculés dans les dernières couches convolutionnelles du réseau. Les zones les plus activées correspondent aux régions de l'image ayant le plus contribué à la prédiction finale du modèle.

L'application de Grad-CAM sur les scanners cérébraux permet ainsi de visualiser les régions anatomiques suspectes associées aux hémorragies intracrâniennes ou aux lésions ischémiques détectées par le modèle. Cette interprétation visuelle facilite la validation clinique des résultats et améliore la compréhension du comportement du système de classification automatique. La méthode Grad-CAM est aujourd'hui largement utilisée dans les applications médicales récentes grâce à sa simplicité, sa robustesse et sa capacité à fournir des explications visuelles intuitives directement exploitables par les radiologues [51].

3.9 Conclusion

Ce chapitre a détaillé la démarche méthodologique adoptée pour la conception et l'implémentation de notre système de classification automatique des accidents vasculaires cérébraux (AVC) à partir d'images de tomodensitométrie. Après avoir présenté l'environnement technique et la structure de notre base de données, nous avons exposé le pipeline rigoureux de préparation des données, incluant les étapes essentielles de prétraitement, d'équilibrage et d'augmentation des données, visant à optimiser la qualité des entrées et la robustesse de nos modèles.

Par la suite, nous avons décrit les quatre architectures de réseaux de neurones profonds sélectionnées — EfficientNet-B0, MobileNet-V2, ViT-B/16 et Swin-T — en mettant en avant

les modifications structurelles et les adaptations que nous avons apportées pour spécialiser ces modèles aux contraintes cliniques de notre étude. Enfin, l'arsenal des métriques d'évaluation retenu nous permettra de mesurer avec précision les performances diagnostiques de ces approches. Ce cadre expérimental robuste constitue désormais le fondement nécessaire pour l'analyse des résultats obtenus, qui sera détaillée dans le chapitre suivant.

Chapitre 4

Résultats et discussion

4.1 Introduction

Ce chapitre est dédié à l'exposition, l'analyse approfondie et la discussion critique des résultats expérimentaux obtenus, visant à valider l'efficacité de notre chaîne de traitement amont pour le diagnostic automatisé des accidents vasculaires cérébraux (AVC) par scanner CT cérébral. Face aux défis majeurs du déséquilibre des classes, de la ressemblance sémantique inter-pathologique et du glissement de domaine (Domain Shift), nos expérimentations ont été structurées autour de quatre scénarios décisionnels complémentaires : une classification multi-classes simultanée (Sain, Ischémie, Hémorragie), un dépistage global des anomalies (Normal vs Pathologique), et deux tâches binaires ciblées respectivement sur l'urgence hémorragique (Normal vs Bleeding) et ischémique (Normal vs Ischemia).

Après une évaluation interne rigoureuse sur le *Brain stroke CT Dataset*, ce travail met un accent crucial sur la robustesse et la généralisation de nos deux meilleurs modèles, EfficientNet-B0 et Swin-T, testés sur la base de données internationale indépendante RSNA. Les performances de ces architectures ont été soumises à une validation statistique approfondie (tests de McNemar, Wilcoxon et analyse Bootstrap) qui prouve la supériorité et la résilience des mécanismes d'attention visuelle face aux variations inter-hospitalières. Enfin, une confrontation directe avec les méthodes récentes de l'état de l'art (2024–2026) vient démontrer la valeur ajoutée et la pertinence clinique de notre contribution avant de conclure ce chapitre.

4.2 Classification automatique des accidents vasculaires cérébraux sur scanner CT cérébral

Pour analyser finement le comportement de ces modèles face aux contraintes cliniques, nous avons structuré nos expérimentations autour de trois scénarios décisionnels :

1. Une approche multi-classes simultanée à trois catégories (Bleeding, Ischemia, Normal).
2. une tâche de dépistage global des anomalies (Normal vs Pathologie)
3. Une tâche binaire ciblée sur l'urgence hémorragique (Normal vs Bleeding).
4. Une tâche binaire dédiée à la détection des zones ischémiques (Normal vs Ischemia).

Les performances ont été rigoureusement quantifiées et comparées via un ensemble de métriques cliniquement représentatives : l'exactitude (Accuracy), la précision, le rappel (sensibilité), le F1-score macro, la spécificité et l'aire sous la courbe ROC (AUC). Cette double

approche, multi-classes et binaire, permet non seulement d'évaluer la robustesse de chaque architecture face au déséquilibre initial du jeu de données, mais aussi de mesurer précisément le gain diagnostique obtenu par la simplification de l'espace décisionnel en contexte médical d'urgence.

4.2. 1 Évaluation comparative de la classification multi-classes

Cette section est dédiée à l'analyse approfondie et à la confrontation des performances des quatre architectures de réseaux de neurones profonds au sein d'un espace décisionnel multiclassés. L'objectif fondamental de ce premier scénario est d'évaluer l'aptitude de chaque modèle à discriminer simultanément les trois catégories cliniques cibles à partir de coupes scanographiques cérébrales. Cette configuration permet d'éprouver la robustesse des extracteurs de caractéristiques face aux défis conjoints du déséquilibre natif du jeu de données et de la ressemblance sémantique inter-pathologique, tout en mesurant l'apport direct de notre chaîne de traitement amont.

Le Tableau 4. 1 suivante présente les résultats comparatifs des quatre modèles en classification multi-classes (Bleeding / Ischemia / Normal).

Modèle	Précision	Rappel	F1-Score	Sensibilité	Spécificité	AUC
EfficientNet-B0	97 %	97.03%	97.24 %	97.07 %	98.47 %	99.35 %
MobileNet-V2	96.33 %	95.33%	95.94 %	95.60 %	97.92 %	97.91 %
ViT-B/16	97.33 %	96.76%	97.08 %	96.85 %	98.40 %	99.12 %
Swin-T	98.33 %	97.69%	97.94%	97.69 %	98.80 %	99.80 %

Tableau 4. 1 : Résultats comparatifs des architectures d'apprentissage profond pour la classification multiclassés des AVC.

L'examen des performances reportées dans le Tableau 4. 1 met en lumière la nette supériorité des architectures basées sur les mécanismes d'attention visuelle, tout en confirmant la haute compétitivité des modèles convolutifs. Parmi les modèles testés, le Swin-T se distingue par des performances remarquables et s'impose comme l'architecture leader sur l'ensemble des métriques. Il atteint une Accuracy de 97,97 % et un F1-Score de 97,94 %, soutenus par la meilleure capacité de discrimination avec une AUC de 99,80 %. L'analyse détaillée de ses capacités prédictives montre une excellente précision (98,33 %) et un rappel de 97,69 %. Sur le plan clinique, sa sensibilité élevée (97,69 %) combinée à une spécificité robuste (98,80 %) garantit une identification extrêmement fiable des

pathologies tout en minimisant les faux positifs. EfficientNet-B0 occupe la deuxième position en affichant une excellente complémentarité, avec une Accuracy de 97,30 %, un F1-Score de 97,24 % et une AUC de 99,35 %. Ce modèle convolutifs fait preuve d'un équilibre remarquable entre sa sensibilité (97,07 %) et sa spécificité (98,47 %). Il est suivi de près par le ViT-B/16, qui enregistre également des résultats solides avec une Accuracy de 97,18 %, une précision de 97,33 % et une AUC de 99,12%. Enfin, bien que MobileNet-V2 se positionne en retrait par rapport aux autres architectures notamment avec le rappel (95,33 %) et l'au-delà de la courbe (AUC de 97,91 %) les plus bas, ses scores globaux restent particulièrement élevés et compétitifs pour un modèle léger.

En résumé, le Swin-T s'impose comme l'architecture la mieux adaptée à cette tâche complexe de classification. Ces résultats d'un niveau globalement très élevé soulignent l'importance cruciale de notre pipeline de prétraitement incluant le nettoyage morphologique, la normalisation et l'équilibrage des classes par sur-échantillonnage pour optimiser l'efficacité et la convergence des modèles sur les images de scanner CT cérébral.

La figure 4. 1 suivante illustre les matrices de confusion des quatre modèles d'apprentissage profond pour la classification multiclasse (Bleeding, Ischemia, Normal).

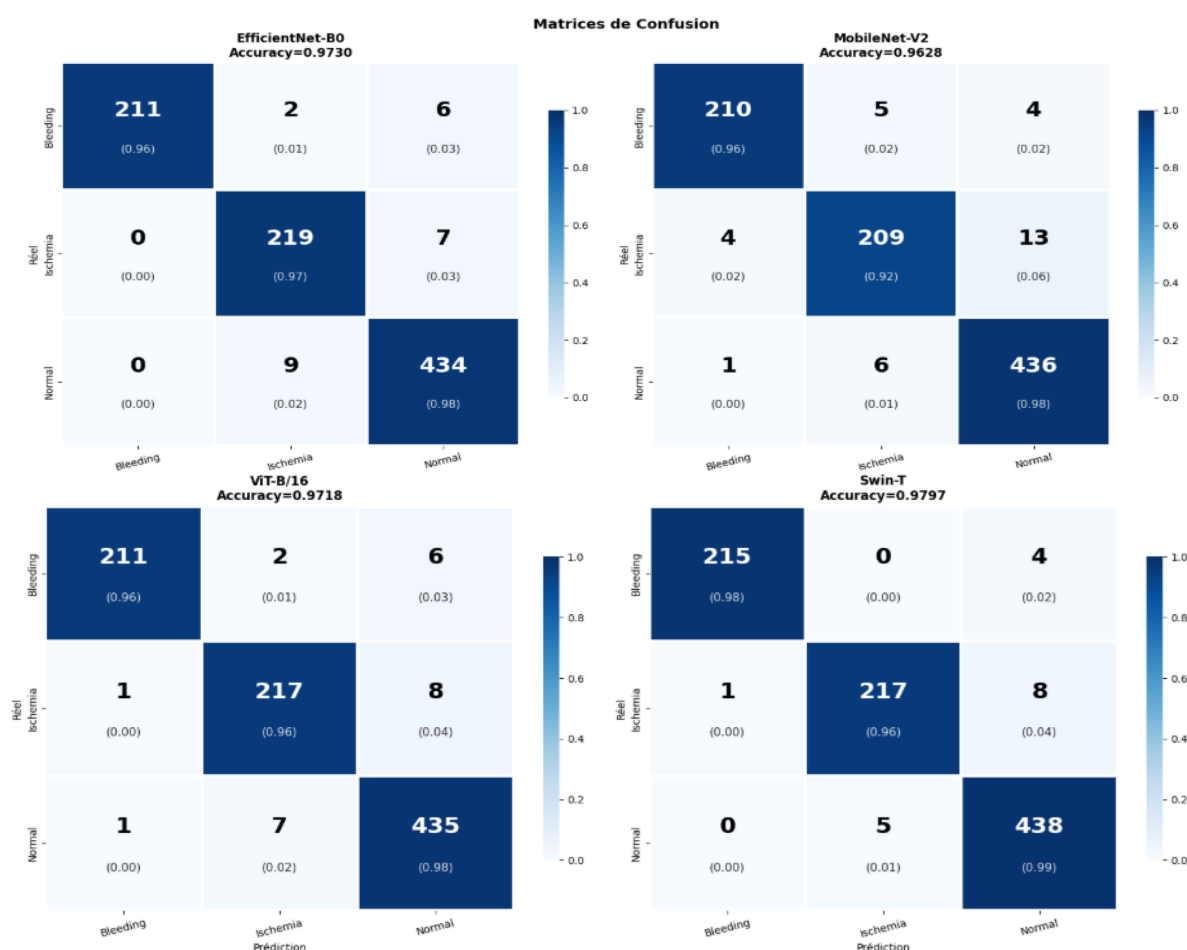


Figure 4. 1 : Matrices de confusion des quatre modèles d'apprentissage profond pour la classification multiclasse (Bleeding, Ischemia, Normal).

L'analyse des matrices de confusion montre que les quatre architectures étudiées sont capables de distinguer efficacement les trois classes cliniques considérées (Normal, Bleeding et Ischemia). Le modèle Swin-T a obtenu les meilleures performances globales avec une Accuracy de 97,97 %, suivi de près par EfficientNet-B0 et ViT-B/16, qui présentent également des taux de classification très élevés. MobileNet-V2 affiche des résultats légèrement inférieurs, tout en conservant une bonne capacité de discrimination entre les différentes classes.

Les principales erreurs observées concernent la confusion entre certaines images ischémiques et normales, ce qui peut être expliqué par la présence de lésions discrètes présentant des caractéristiques visuelles proches du tissu cérébral sain sur les images tomодensitométriques. En revanche, les cas d'hémorragie sont généralement mieux identifiés grâce à leur contraste plus marqué.

4. 2. 2 Évaluation comparative de la classification binaire (normal/ pathologie)

Cette section est dédiée à l'analyse approfondie et à la confrontation des performances des quatre architectures de réseaux de neurones profonds au sein d'un espace décisionnel binaire visant le dépistage global des anomalies. L'objectif fondamental de ce second scénario est d'évaluer l'aptitude de chaque modèle à distinguer les coupes cérébrales saines de celles présentant une pathologie, en regroupant les cas d'hémorragie et d'ischémie sous une classe unique « Pathologie ». Cette configuration permet d'éprouver la robustesse des extracteurs de caractéristiques dans une tâche de triage clinique simplifié, tout en mesurant l'impact de la simplification de l'espace de sortie sur la réduction des confusions inter-pathologiques et l'amélioration de la sensibilité du système.

Le Tableau 4. 2 suivant présente les résultats comparatifs des quatre modèles en classification binaire (Normal vs Pathologie).

Modèle	Précision	Rappel	F1-Score	Sensibilité	Spécificité	AUC
EfficientNet-B0	97.41%	97.41%	97.41%	97.41%	97.41%	99.5%
MobileNet-V2	97.73%	97.75%	96.73%	96.74%	96.74%	99.4%
ViT-B/16	95.72%	95.72%	95.72%	95.72%	95.72%	99.3%
Swin-T	97.18%	97.18%	97.17%	97.17%	97.17%	99.7%

Tableau 4. 2 : Résultats comparatifs des architectures d'apprentissage profond pour la classification binaire (normal pathologie) des AVC.

L'examen des performances reportées dans le Tableau 4.2 met en lumière l'excellente capacité de généralisation des quatre architectures testées pour cette tâche de classification. Parmi les modèles évalués, EfficientNet-B0 s'impose comme l'architecture la plus performante, affichant une performance homogène et optimale avec une Accuracy, une Précision, un Rappel, un F1-Score, une Sensibilité ainsi qu'une Spécificité uniformes de 97,41 %, complétés par une solide capacité de discrimination (AUC de 99,5 %). Le Swin-T se distingue également par la meilleure capacité de discrimination globale du groupe, enregistrant la valeur d'AUC la plus élevée avec 99,7 %, tout en maintenant des scores de performance très robustes (Accuracy de 97,18 %). Le MobileNet-V2 démontre une grande efficacité, particulièrement avec des taux de Précision (97,73 %) et de Rappel (97,75 %) très compétitifs. Enfin, le ViT-B/16 complète ce panorama avec des résultats solides (Accuracy de 95,72 % et AUC de 99,3 %), confirmant la pertinence de l'ensemble de nos modèles pour cette application.

En résumé, ces résultats d'un niveau globalement très élevé soulignent l'importance cruciale de notre pipeline de prétraitement et la pertinence de nos choix architecturaux. La complémentarité entre la précision des modèles convolutifs et la puissance discriminante des architectures basées sur l'attention confirme la robustesse de notre approche pour l'analyse automatisée des images de scanner CT cérébral.

La figure 4.2 suivante illustre les matrices de confusion des quatre modèles d'apprentissage profond pour la classification binaire (Normal, Bleeding)

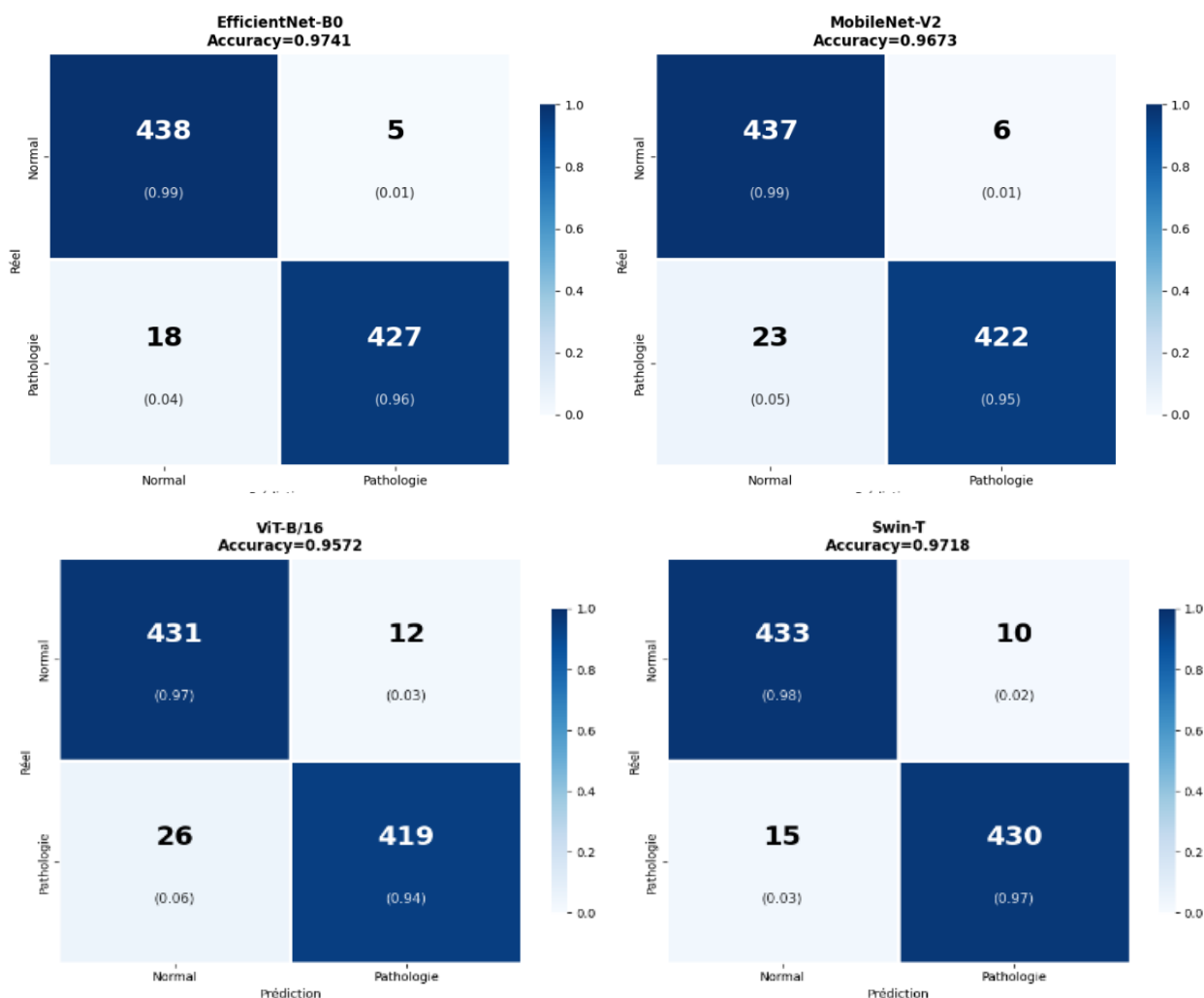


Figure 4. 2 : Matrices de confusion des quatre modèles d'apprentissage profond pour la classification binaire (Pathologie, Normal).

L'analyse des matrices de confusion présentées montre que les quatre architectures étudiées sont capables de distinguer efficacement les deux classes cliniques considérées (Normal et Pathologie). Le modèle EfficientNet-B0 a obtenu les meilleures performances globales avec une Accuracy de 97,41 %, suivi de près par Swin-T (97,18 %) et MobileNet-V2 (96,73 %), tandis que ViT-B/16 présente également un taux de classification très élevé (95,72 %).

Les principales erreurs observées concernent la confusion entre certaines images pathologiques et normales, ce qui reste marginal compte tenu de la précision globale des modèles. La simplification de l'espace décisionnel vers une configuration binaire a permis de réduire significativement les ambiguïtés, facilitant ainsi une identification robuste des anomalies cérébrales par rapport au tissu cérébral sain.

4. 2. 3 Évaluation comparative de la classification binaire (normal/ bleeding)

Cette section est dédiée à l'analyse approfondie et à la confrontation des performances des quatre architectures de réseaux de neurones profonds au sein d'un espace décisionnel binaire spécifiquement orienté vers l'urgence hémorragique. L'objectif fondamental de ce troisième scénario est d'évaluer la capacité de chaque modèle à isoler avec une haute précision les cas d'hémorragie intracrânienne par rapport aux examens normaux. Cette configuration, cruciale pour les applications cliniques d'urgence, permet de mesurer l'efficacité de nos modèles dans la détection d'une pathologie critique dont le diagnostic rapide conditionne le pronostic vital du patient. Elle permet également d'isoler la performance du système face au défi spécifique que représente la détection des foyers hémorragiques dans des volumes CT complexes.

Le Tableau 4. 3 suivant présente les résultats comparatifs des quatre modèles en classification binaire (Normal vs Bleeding).

Modèle	Précision	Rappel	F1-Score	Sensibilité	Spécificité	AUC
EfficientNet-B0	99 %	99.50 %	99.15 %	99.09 %	99.09 %	99.70 %
MobileNet-V2	98.50 %	99 %	98.64 %	98.75 %	98.75 %	99.90 %
ViT-B/16	98.50 %	98.50 %	98.46 %	98.18 %	98.18 %	99.70 %
Swin-T	99.50 %	99.50 %	99.32 %	99.20%	99.20%	99.91%

Tableau 4. 3 : Résultats comparatifs des architectures d'apprentissage profond pour la classification binaire des AVC.

Parmi l'ensemble des configurations de tableau 4. 3, le modèle Swin-T se distingue par des performances remarquables et s'impose comme l'architecture leader sur la quasi-totalité des métriques. Il atteint une exactitude (*Accuracy*) de 99,40 % et un *F1-Score* de 99,32 %, soutenus

par une capacité de discrimination parfaite matérialisée par une AUC de 99.91 %. L'analyse détaillée de ses capacités prédictives montre un équilibre idéal avec une précision et un rappel s'élevant tous deux à 99,50 %. Sur le plan clinique, sa sensibilité élevée (99,20 %) combinée à une spécificité tout aussi robuste (99,20 %) garantit une identification extrêmement fiable des pathologies tout en minimisant drastiquement les risques de faux positifs et de faux négatifs. EfficientNet-B0 occupent la deuxième position en affichant une excellente complémentarité. Ce modèle convolutif fait preuve d'une haute compétitivité avec une *F1-Score* de 99,15 % et une AUC solide de 99,70 %. Il démontre également un équilibre notable entre sa sensibilité et sa spécificité, toutes deux mesurées à 99,09 %. Le modèle léger MobileNet-V2 surprend positivement en se positionnant au troisième rang global. Bien que ses scores de précision (98,50 %) et de *F1-Score* (98,64 %) soient légèrement en retrait, il surclasse EfficientNet-B0 et ViT-B/16 en termes de pouvoir discriminatoire avec une AUC remarquable de 99,90 %, prouvant son efficacité malgré une architecture optimisée pour la légèreté. Enfin, le ViT-B/16 ferme la marche tout en maintenant des standards de performance très élevés. Il enregistre un *F1-Score* de 98,46 % ainsi qu'une sensibilité et spécificité de 98,18 %.

En résumé, l'architecture à attention fenêtrée Swin-T s'impose comme la solution la mieux adaptée à cette tâche complexe de classification. Ces résultats, d'un niveau globalement exceptionnel, valident et soulignent l'importance cruciale de notre pipeline de prétraitement incluant le nettoyage morphologique, la normalisation et l'équilibrage des classes par sur-échantillonnage pour optimiser la convergence et l'efficacité des modèles sur les images de scanner CT cérébral.

La figure 4. 3 suivante illustre les matrices de confusion des quatre modèles d'apprentissage profond pour la classification binaire (Normal, Bleeding)

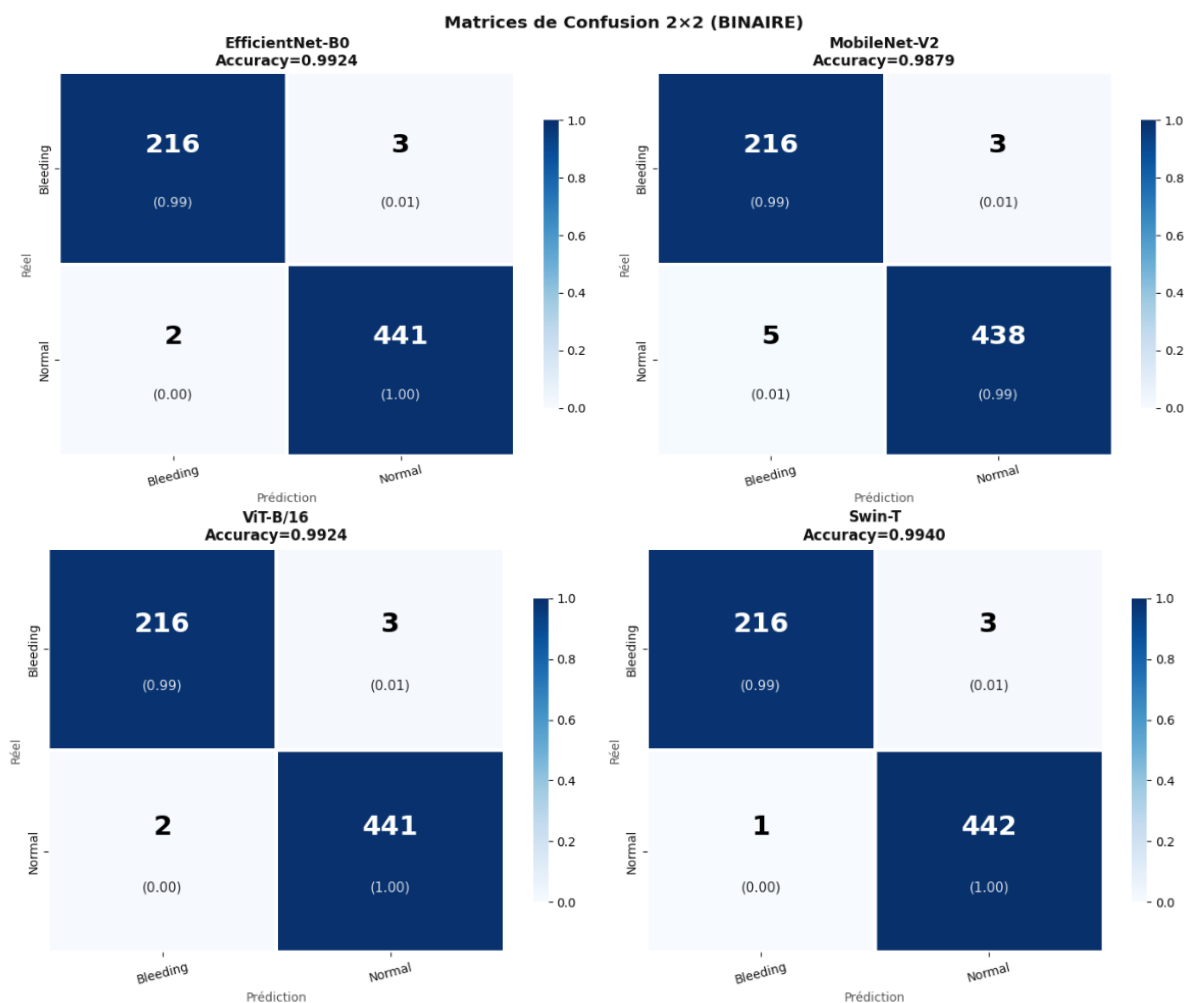


Figure 4. 3 : Matrices de confusion des quatre modèles d'apprentissage profond pour la classification binaire (Bleeding, Normal).

L'analyse des matrices de confusion montre que les quatre architectures étudiées sont capables de distinguer très efficacement les deux classes cliniques considérées (Normal et Bleeding). Le modèle Swin-T a obtenu les meilleures performances globales avec une Accuracy de 99,40 %, suivi de très près par EfficientNet-B0 (99,24 %), MobileNet-V2 (98,79 %) et ViT-B/16 (98,64 %), qui présentent tous des taux de classification extrêmement élevés. L'examen des matrices de confusion binaire montre que les modèles parviennent à une séparation quasi parfaite entre les tissus sains et les signes d'hémorragie. Les rares erreurs observées se traduisent par de très faibles taux de faux positifs et de faux négatifs, ce qui s'avère crucial pour la sécurité diagnostique. Cette précision accrue s'explique par la simplification de l'espace décisionnel, permettant une distinction nette entre le tissu cérébral sain et les foyers hémorragiques grâce à leur contraste radiologique marqué sur les images tomodensitométriques.

4.2. 4 Évaluation comparative de la classification binaire (normal/ Ischemia)

Cette section est dédiée à l'évaluation des performances des quatre architectures sur une seconde tâche de classification binaire, restreinte aux classes *Normal* et *Ischemia*. Cette configuration cible spécifiquement la détection des accidents vasculaires cérébraux ischémiques. Les anomalies associées à l'ischémie présentent des caractéristiques visuelles particulièrement subtiles sur les images de tomodensitométrie (scanner CT), ce qui rend leur identification nettement plus complexe que celle des hémorragies. L'enjeu est ici d'évaluer la sensibilité des modèles face à ces signatures radiologiques discrètes, un défi majeur pour l'aide au diagnostic clinique d'urgence.

Le Tableau 4. 4 regroupe l'ensemble des métriques d'évaluation, offrant ainsi une perspective comparative directe de l'efficacité de chaque modèle dans cette configuration binaire.

Modèle	Précision	Rappel	F1-Score	Sensibilité	Spécificité	AUC
EfficientNet-B0	97.50 %	98 %	97.83 %	97.77 %	97.77 %	99.80 %
MobileNet-V2	97 %	96.5 %	96.80 %	96.45 %	96.45 %	99.50 %
ViT-B/16	96 %	96 %	95.97 %	95.77 %	95.77 %	99 %
Swin-T	97 %	97.5 %	97.17 %	97.32 %	97.32 %	99.60 %

Tableau 4. 4 : Synthèse comparative des métriques de performance des quatre architectures d'apprentissage profond pour la classification binaire (Normal vs Ischemia)

L'examen des performances reportées dans le Tableau 4.4 met en lumière la nette supériorité des architectures basées sur les mécanismes d'attention visuelle, tout en confirmant la haute compétitivité des modèles convolutifs. Parmi les modèles testés, l'architecture EfficientNet-B0 a obtenu les meilleures performances globales avec une Accuracy de 98,06 %,

une Précision de 97,50 %, un Rappel de 98 % et un F1-Score de 97,83 %. Ce modèle se distingue également par une Sensibilité et une Spécificité de 97,77 %, ainsi qu'une aire sous la courbe maximale avec une AUC de 99,80 %, ce qui en fait l'architecture la plus robuste de cette étude pour la détection de l'ischémie. Le modèle Swin-T se positionne comme la deuxième approche la plus performante, affichant une Accuracy de 97,46 %, un F1-Score de 97,17 % et une AUC de 99,60 %. L'architecture MobileNet-V2 suit de près avec une Accuracy de 97,16 %, une Précision de 97 % et un F1-Score de 96,80 %, démontrant une grande efficacité. Enfin, le modèle ViT-B/16 affiche les résultats les plus faibles de l'étude avec une Accuracy de 96,41 % et un F1-Score de 95,97 %, tout en conservant une excellente capacité de discrimination globale comme le montre son AUC de 99,00 %.

En résumé, EfficientNet-B0 s'impose ici comme l'architecture la mieux adaptée à cette tâche de classification.

La figure 4. 4 suivante illustre les matrices de confusion des quatre modèles d'apprentissage profond pour la classification binaire (Normal, Ischemia)

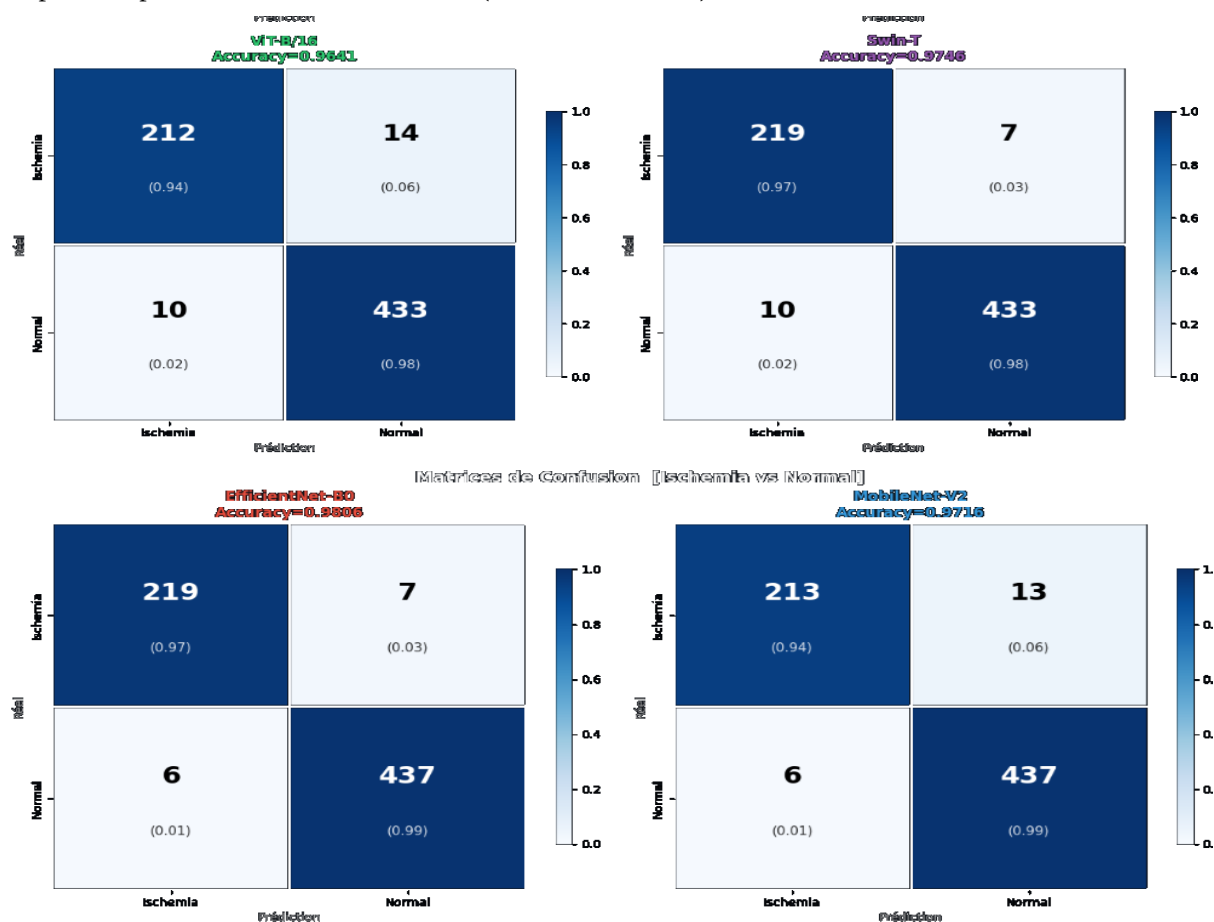


Figure 4. 4 : Matrices de confusion des quatre modèles d'apprentissage profond pour la classification binaire (Ischemia, Normal).

L'analyse des matrices de confusion montre que les quatre architectures étudiées sont capables de distinguer très efficacement les deux classes cliniques considérées (Normal et Ischemia). Parmi les modèles testés, EfficientNet-B0 s'impose comme l'architecture la plus performante avec une Accuracy globale de 98,06 %. Il est suivi par Swin-T (97,46 %), MobileNet-V2 (97,16 %) et ViT-B/16 (96,41 %). L'examen de la matrice de confusion d'EfficientNet-B0 confirme une séparation robuste entre les tissus sains et les signes d'ischémie, avec un taux de faux négatifs et de faux positifs extrêmement réduit, garantissant ainsi une fiabilité diagnostique optimale. Cette précision élevée s'explique par la capacité du modèle à isoler les lésions ischémiques, même lorsqu'elles présentent des caractéristiques densitométriques subtiles sur les images tomodensitométriques (NCCT). Cette minimisation critique des faux négatifs atteste de la sécurité diagnostique du modèle, renforçant sa pertinence pour une aide à la décision clinique en temps réel.

4.3 La généralisation

Après avoir validé la performance de nos modèles sur nos données internes, il est impératif d'évaluer leur comportement en conditions réelles. Un système d'aide au diagnostic robuste ne doit pas seulement être performant en environnement contrôlé ; il doit avant tout démontrer une réelle capacité de généralisation. Cela implique une fiabilité constante, même lors de l'analyse d'images issues de protocoles d'acquisition différents ou de sources cliniques variées.

C'est pourquoi, en complément de nos tests internes, nous avons soumis nos modèles à une évaluation externe rigoureuse en utilisant la base de données RSNA, telle que décrite dans le chapitre précédent. Cette base, composée d'images que le système n'a jamais rencontrées durant sa phase d'apprentissage, constitue un véritable banc d'essai pour mesurer sa robustesse.

Pour valider cette généralisation, nous avons réalisé des analyses statistiques approfondies sur la capacité de discrimination binaire des modèles, en se concentrant sur la distinction entre les classes « Normal » et « Bleeding ». Cette classification binaire est cruciale pour le triage clinique, car elle permet d'isoler immédiatement les cas d'hémorragies intracrâniennes nécessitant une intervention urgente, tout en confirmant l'absence de pathologie chez les patients sains. La confrontation des résultats obtenus sur ces deux classes, entre les tests internes et l'évaluation externe, nous permet de quantifier précisément l'impact du changement de distribution des données (dataset shift) sur la fiabilité diagnostique.

Pour atténuer ce glissement de domaine et sécuriser le transfert des modèles vers ce nouvel environnement clinique, notre chaîne d'adaptation met en œuvre quatre mécanismes complémentaires :

- **Stochastic Weight Averaging (SWA)** : Cette technique calcule la moyenne des poids du modèle au fil des époques afin de stabiliser ses paramètres. Elle permet d'obtenir des configurations plus équilibrées, aidant ainsi le modèle à mieux réagir face à des images provenant d'autres constructeurs hospitaliers.
- **Pseudo-Labeling itératif** : Ce mécanisme permet au modèle de prédire lui-même des étiquettes sur les images non annotées du domaine cible. Les prédictions les plus sûres sont ensuite réinjectées petit à petit dans l'entraînement, ce qui améliore la précision globale sans nécessiter d'effort d'annotation humain.
- **Test-Time Augmentation (TTA)** : Lors de la phase de test, chaque image subit plusieurs modifications géométriques et d'intensité, telles que des rotations, des zooms ou des variations de luminosité. Le modèle analyse l'ensemble de ces versions et combine leurs résultats, ce qui réduit considérablement les erreurs d'interprétation visuelle.

4.4. Comparaison entre test interne et test externe

L'évaluation de nos modèles sur un jeu de données externe, confrontée aux performances obtenues lors du test interne, constitue un indicateur déterminant de leur fiabilité et de leur capacité à maintenir une précision diagnostique sur des données cliniques variées. Cette démarche nous a permis de mesurer la robustesse réelle des modèles face à des conditions d'acquisition différentes de celles rencontrées durant l'entraînement. Dans ce cadre, nous avons réalisé une comparaison approfondie des résultats obtenus pour la classification binaire des classes Normal et Bleeding.

Cette analyse comparative est cruciale, car elle met en évidence la capacité de nos architectures à isoler les lésions hémorragiques indépendamment de la source des images.

Le tableau 4.5 suivant représente la comparaison des performances des modèles (EfficientNet-B0 et Swin-T) entre le test interne et l'évaluation externe (RSNA).

	Modèle	Accuracy	F1-Score	Sensibilité	Spécificité	AUC
Test interne (Brain stroke CT Dataset)	EfficientNet-B0	99.24%	99.15 %	99.09 %	99.09 %	99.70 %
	Swin-T	99.40 %	99.32 %	99.20%	99.20%	99.91 %
Test externe(RSNA)	EfficientNet-B0	84.67 %	85.21 %	85.21 %	81.00 %	92.68 %
	Swin-T	86.83 %	86.40 %	83.67 %	90 %	94.38 %

Tableau 4.5 : Comparaison des performances des modèles (EfficientNet-B0 et Swin-T) entre le test interne et l'évaluation externe (RSNA).

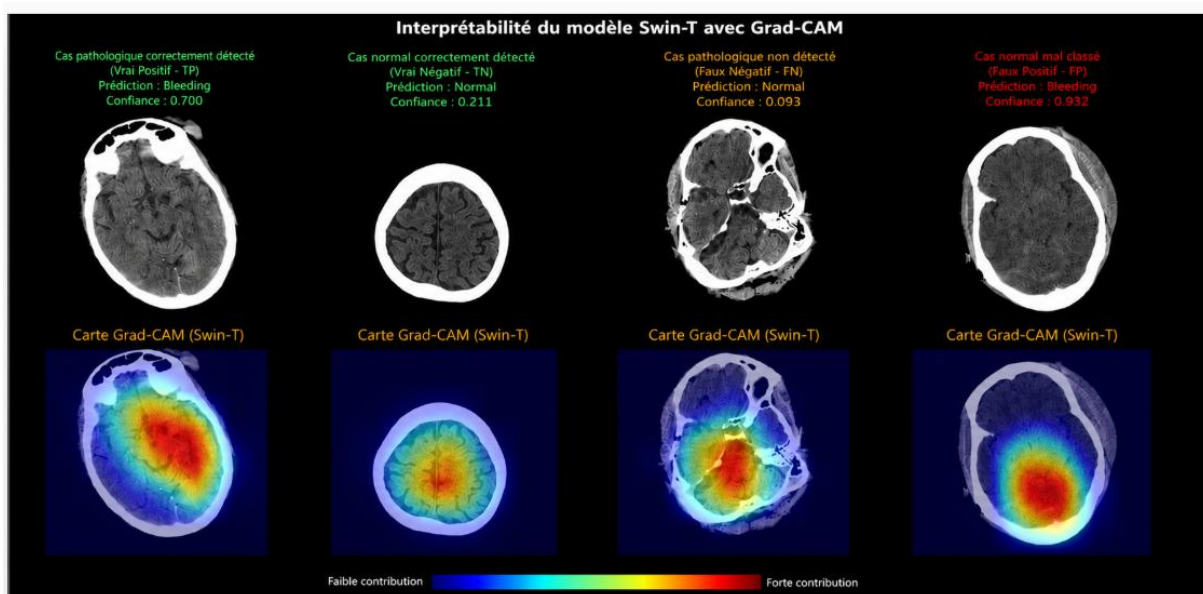
La réussite de notre évaluation externe constitue l'un des résultats les plus significatifs de ce travail. En soumettant nos modèles à une base de données totalement nouvelle (RSNA), jamais rencontrée lors de la phase d'entraînement, nous avons démontré leur réelle capacité à généraliser leurs acquis à des conditions cliniques hétérogènes. L'analyse comparative entre les tests internes et externes illustre parfaitement cette montée en compétence : alors que nos modèles atteignent des métriques quasi-optimales en environnement contrôlé (Accuracy > 99 %), le passage à la base RSNA constitue un véritable test de robustesse. Malgré une baisse de performance inhérente au dataset shift entre les protocoles d'acquisition, nos modèles conservent une efficacité diagnostique remarquable avec une Accuracy dépassant les 84 %. Cette confrontation est révélatrice, car elle permet de distinguer une simple mémorisation de données d'une réelle compréhension des signatures pathologiques. Dans ce contexte, le modèle Swin-T se distingue nettement par sa stabilité supérieure, affichant une spécificité de 90 % contre 81 % pour l'EfficientNet-B0. Cette supériorité face à des données inédites confirme que le mécanisme d'auto-attention hiérarchique du Swin-T excelle dans l'extraction de caractéristiques morphologiques universelles, moins dépendantes des spécificités de notre base initiale. Avec une AUC de 94,38 % en contexte externe, le Swin-T prouve sa capacité à filtrer

le "bruit" des données externes pour se focaliser sur les lésions réelles. Ces résultats valident la pertinence de nos têtes décisionnelles personnalisées et confirment que, par leur résilience diagnostique, nos architectures sont prêtes à être déployées au-delà de notre jeu de données original, garantissant ainsi une fiabilité clinique constante face à l'hétérogénéité des scanners hospitaliers.

4.5. Explicabilité des modèles

Dans cette partie, nous nous intéressons à l'interprétabilité du modèle Swin-T personnalisé à l'aide de la méthode Grad-CAM (Gradient-weighted Class Activation Mapping). Malgré leurs excellentes performances en classification d'images médicales, les modèles d'apprentissage profond restent souvent difficiles à interpréter en raison de leur structure complexe. Dans le domaine médical, il est essentiel de comprendre les raisons ayant conduit le modèle à prendre une décision donnée afin de garantir la fiabilité des résultats diagnostiques.

La méthode Grad-CAM permet de générer des cartes thermiques mettant en évidence les régions de l'image ayant le plus influencé la prédiction finale du modèle. Les figures (4.5) présentent plusieurs exemples de scanners cérébraux accompagnés de leurs cartes d'activation correspondantes. Les zones les plus activées apparaissent sous forme de régions colorées intenses et correspondent principalement aux foyers hémorragiques présents dans les images CT.



La figure 4.5 : Visualisation de l'interprétabilité des prédictions du modèle Swin-T à l'aide de la méthode Grad-CAM pour la classe Bleeding.

Les résultats obtenus montrent que le modèle Swin-T concentre efficacement son attention sur les régions anatomiques pertinentes associées aux hémorragies intracrâniennes. Cette analyse confirme ainsi la cohérence médicale des prédictions réalisées et démontre que le modèle se base sur des caractéristiques cliniquement significatives plutôt que sur des artefacts ou des régions non importantes de l'image. Grâce à cette étape d'explicabilité, la confiance dans le système de classification proposé est considérablement renforcée.

L'image présente une analyse d'interprétabilité du modèle **Swin-T** à l'aide de la méthode **Grad-CAM** appliquée à des scanners cérébraux CT. Quatre situations de classification sont illustrées : **VP**, **VN**, **FN** et **FP**. Pour chaque cas, l'image originale est affichée en haut, tandis que la carte thermique Grad-CAM correspondante apparaît en bas afin de visualiser les régions ayant le plus influencé la décision du modèle.

Dans le cas **VP**, le modèle détecte correctement la présence d'une hémorragie intracrânienne. La carte Grad-CAM montre une forte activation rouge et jaune localisée précisément sur la région pathologique, indiquant que le modèle concentre son attention sur la zone hémorragique pertinente.

Dans le cas **VN**, l'image normale est correctement classifiée. Les activations restent faibles et diffuses, traduisant l'absence de régions suspectes importantes dans le scanner cérébral.

Le cas **FN correspond** à une hémorragie non détectée par le modèle. Bien que certaines zones internes soient activées, l'attention du réseau reste insuffisante ou mal localisée, ce qui explique l'erreur de classification vers la classe normale.

Enfin, dans le cas **FP** une image normale est incorrectement prédite comme pathologique. La carte Grad-CAM révèle une forte activation dans une région anatomique non hémorragique, montrant que le modèle a été influencé par des structures ou variations d'intensité ressemblant à une anomalie.

Globalement, cette analyse démontre que le modèle Swin-T parvient majoritairement à focaliser son attention sur des régions médicalement pertinentes. Les cartes d'activation confirment ainsi la cohérence clinique des prédictions et améliorent la transparence du système de classification automatique des hémorragies intracrâniennes.

4. 6 Analyse statistique

L'analyse statistique a été réalisée afin d'évaluer les performances des modèles de classification sur l'ensemble de test. Les résultats ont été mesurés à l'aide de plusieurs métriques telles que l'Accuracy, la Sensibilité, la Spécificité et le F1-score. Une attention particulière a été portée à la réduction des faux négatifs afin d'améliorer la fiabilité du diagnostic des hémorragies intracrâniennes.

La figure 4.6 suivante représente l'analyse statistique comparative entre EfficientNet-B0 et Swin-T sur les datasets interne et RSNA externe.

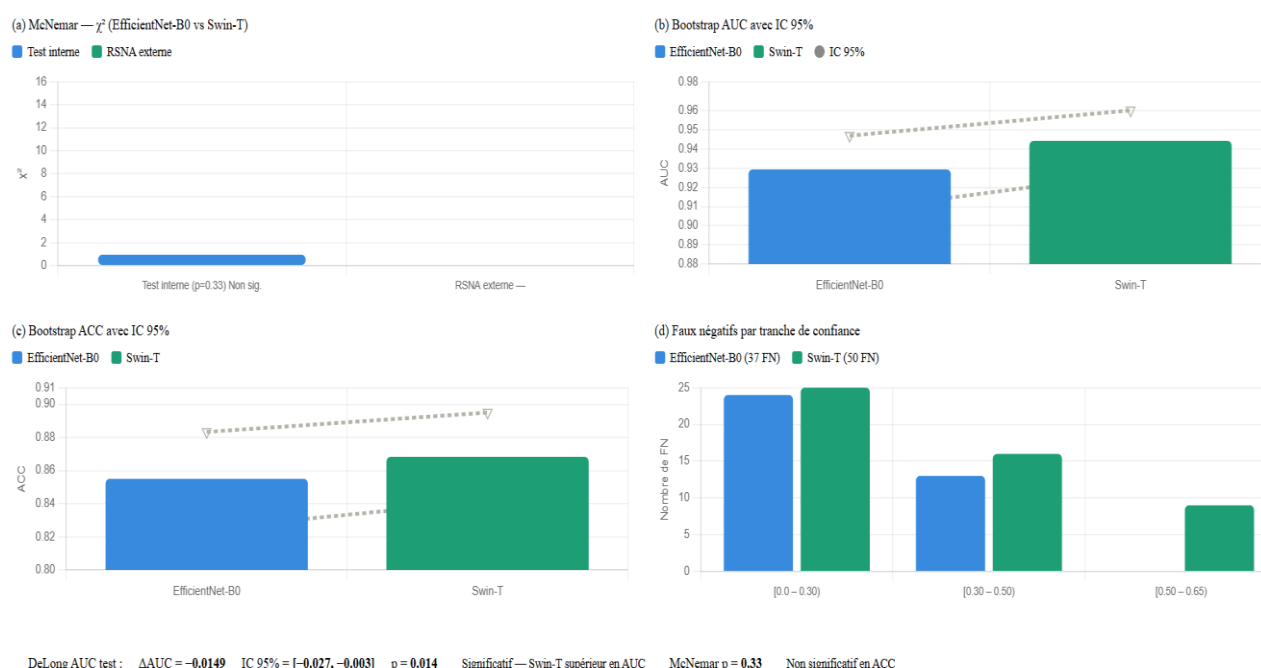


Figure 4.6 : Analyse statistique comparative entre EfficientNet-B0 et Swin-T sur les BDD interne et RSNA externe.

Les résultats statistiques présentés dans la Figure 4.6 révèlent une comparaison nuancée entre EfficientNet-B0 et Swin-T. Le test de McNemar indique que les deux modèles ne diffèrent pas significativement en termes d'Accuracy sur le dataset interne, ce qui est corroboré par les intervalles de confiance Bootstrap qui se chevauchent partiellement. En revanche, le test de DeLong confirme une supériorité statistiquement significative du Swin-T en AUC ($\Delta AUC = -0,0149$, IC 95% = $[-0,027 ; -0,003]$, $p = 0,014$), traduisant une meilleure capacité de discrimination globale. Toutefois, l'analyse des faux négatifs par tranche de

confiance introduit une réserve clinique importante : EfficientNet-B0 génère 37 faux négatifs contre 50 pour Swin-T, soit 26 % de moins, avec un avantage maintenu sur toutes les tranches de confiance. Dans un contexte médical où un faux négatif correspond à un AVC non détecté, ce résultat nuance la supériorité du Swin-T et souligne qu'une meilleure AUC ne se traduit pas nécessairement par une meilleure sécurité diagnostique.

4.7 Comparaison avec l'état de l'art

Cette section confronte les performances de nos architectures (EfficientNet-B0 et Swin-T) avec les travaux de référence publiés sur les jeux de données Brain Stroke CT Dataset et RSNA. Cette démarche permet de situer la pertinence de nos modèles face aux standards actuels de la classification d'AVC par imagerie.

4.7. 1 Brain Stroke CT Dataset

Le Tableau 4.6 synthétise une analyse comparative des taux de classification (TC) obtenus par nos modèles par rapport aux approches de l'état de l'art, évaluées sur le même jeu de données. Cette mise en perspective souligne la supériorité de notre méthodologie, confirmant que nos architectures permettent une meilleure discrimination des caractéristiques pathologiques que les méthodes existantes.

Auteurs	Modèles	TC
Rajaramanan et al. (2024)	EfficientNet-B0	95.13%
d'Inamdar et al. (2025),	VGG-16	95,84 %
Tamilarasi et al. (2026)	InHLeNet	98,2 %,
Sabir et al. (2025)	DCNN	96,88 %
nos modèles	EfficientNet-B0 personnalisé	99.24 %
	MobileNet-V2 personnalisé	98.79 %
	ViT-B/16 personnalisé	98.64 %
	Swin-T personnalisé	99.49 %

Tableau 4.6 : Comparaison des performances des modèles de classification des AVC ischémiques et hémorragiques basés sur le Deep Learning

Le Tableau 4.6 met en évidence la supériorité de nos modèles par rapport aux travaux récents de la littérature. Alors que les modèles de référence atteignent des taux de classification compris entre 95,13 % (Rajaramanan et al. 2024) et 98,2 % (Tamilarasi et al. 2026), nos quatre modèles personnalisés les surpassent tous sans exception. Le Swin-T obtient le meilleur résultat global avec 99,49 %, suivi d'EfficientNet-B0 à 99,24 %, ViT-B/16 à 98,64 % et MobileNet-V2 à 98,79 %. Il est particulièrement notable qu'EfficientNet-B0 dans sa version personnalisée dépasse de près de 4 points sa version standard utilisée par Rajaramanan et al. (2024), confirmant l'apport significatif de notre tête de classification adaptée. Ces résultats valident la qualité de notre pipeline d'entraînement et démontrent que l'intégration de mécanismes d'attention, notamment dans le Swin-T, offre une meilleure capacité de discrimination des structures cérébrales pathologiques.

4.8 Conclusion

Cette étude démontre l'intérêt des techniques d'apprentissage profond pour la classification automatique des accidents vasculaires cérébraux (AVC) à partir d'images de scanner cérébral CT. Grâce à une stratégie de prétraitement efficace et à l'optimisation des architectures utilisées, les modèles développés ont atteint des performances remarquables dans les différentes tâches de classification.

Les expérimentations menées sur le Brain Stroke CT Dataset ont permis d'évaluer plusieurs architectures de deep learning, à savoir EfficientNet-B0, MobileNet-V2, ViT-B/16 et Swin-T. Les résultats obtenus montrent que les modèles basés sur les mécanismes d'attention, notamment Swin-T, offrent les meilleures performances en classification multiclasse et dans la détection des hémorragies intracrâniennes. De son côté, EfficientNet-B0 s'est distingué par son excellente capacité à identifier les cas d'ischémie ainsi qu'à détecter globalement les anomalies cérébrales.

L'évaluation réalisée sur la base de données externe RSNA a permis de vérifier la robustesse et la capacité de généralisation des modèles face à des données provenant de contextes cliniques différents. Malgré l'existence d'un décalage entre les bases de données, les modèles ont conservé des performances satisfaisantes, confirmant ainsi leur potentiel pour une utilisation dans des environnements réels.

Par ailleurs, l'analyse d'interprétabilité basée sur la méthode Grad-CAM a permis de mettre en évidence les régions d'intérêt ayant influencé les décisions du modèle. Les cartes d'activation obtenues montrent que les prédictions reposent principalement sur des zones anatomiques pertinentes, ce qui renforce la fiabilité et la transparence du système proposé.

En conclusion, les résultats obtenus confirment l'efficacité des architectures d'apprentissage profond pour l'aide au diagnostic des AVC à partir d'images CT cérébrales. Les performances élevées obtenues, associées à une bonne capacité de généralisation et à une interprétabilité satisfaisante, soulignent le potentiel de ces approches pour assister les professionnels de santé dans la détection précoce et fiable des accidents vasculaires cérébraux.

Conclusion Générale

L'Accident Vasculaire Cérébral (AVC) demeure l'une des principales causes de mortalité et de handicap dans le monde, ce qui rend indispensable le développement d'outils capables d'assister les cliniciens dans un diagnostic rapide et fiable. Dans ce contexte, ce travail a porté sur la conception d'un système intelligent de classification des AVC à partir d'images de scanner CT cérébral en exploitant les capacités de l'apprentissage profond.

Quatre architectures pré-entraînées ont été étudiées et comparées dans différents scénarios de classification binaire et multiclassées. Les résultats obtenus sur la base Brain Stroke CT Dataset ont démontré d'excellentes performances, avec une supériorité globale du modèle Swin Transformer, tandis qu'EfficientNet-B0 s'est distingué par sa robustesse et sa capacité à réduire les faux négatifs. La comparaison avec les travaux récents de la littérature a confirmé la compétitivité de l'approche proposée.

L'évaluation externe réalisée sur la base RSNA a permis de vérifier la capacité de généralisation des modèles dans un contexte de changement de domaine. Malgré les variations entre les jeux de données, les performances sont restées élevées grâce aux mécanismes d'adaptation mis en œuvre, confirmant ainsi la robustesse du système développé.

Par ailleurs, l'intégration de la méthode Grad-CAM a permis d'améliorer l'interprétabilité des prédictions en mettant en évidence les régions anatomiques ayant influencé les décisions du modèle. L'analyse statistique basée sur les tests de McNemar, Bootstrap et DeLong a également apporté une validation rigoureuse des résultats obtenus.

En conclusion, les résultats de cette étude démontrent le fort potentiel des architectures d'apprentissage profond pour l'aide au diagnostic automatisé des AVC sur des images CT cérébrales. Ils soulignent également l'importance de la généralisation, de l'explicabilité et de la validation statistique comme éléments essentiels pour favoriser une future intégration de ces systèmes dans la pratique clinique.

Références Bibliographiques

- [01] : Johns Hopkins Medicine. "Anatomy of the Brain". [En ligne]. Disponible sur : [hopkinsmedicine.org]. (Consulté le 11 Février 2026).
- [02] : Cleveland Clinic. " Brain : Structure and Function". [En ligne]. Disponible sur : [clevelandclinic.org]. (Consulté le 11 Février 2026).
- [03] : Rouvière, H., & Delmas, A. (2005). Anatomie humaine : Tome 1, Tête et cou. 15ème édition, Éditions Masson, Paris.
- [04] : Osborn, A. G. (2017). Osborn's Brain : Imaging, Pathology, and Anatomy. 2nd Edition, Elsevier Health Sciences.
- [05] : Molloy, P. T., Bilaniuk, L. T., et al. (1995). "Brainstem tumors in patients with Neurofibromatosis type 1". Neurology, 45(10).
- [06] : Shibata, S. (1989). "Ultrastructure of capillary walls in human brain tumors". Acta Neuropathologica, 78(5).
- [07] : Wakasugi, N., & Hanakawa, T. (2021). "Overlapping Molecular and Circuit Pathophysiology in Alzheimer's". Frontiers in Systems Neuroscience.
- [08] : Chen, Z., & Zhong, C. (2013). "Decoding Alzheimer's disease from perturbed cerebral glucose metabolism". Progress in Neurobiology.
- [09] : Martirosyan, A., Ansari, R., et al. (2024). "Unravelling cell type-specific responses to Parkinson's Disease". Molecular Neurodegeneration.
- [10] : Boursin, P., Paternotte, S., et al. (2018). "Sémantique, épidémiologie et sémiologie des AVC". Soins, 63(828).
- [11] : Majumder, D. (2024). "Ischemic Stroke: Pathophysiology and Evolving Treatment Approaches". Neuroscience Insights.
- [12] : Kuriakose, D., & Xiao, Z. C. (2020). "Pathophysiology and Treatment of Stroke". International Journal of Molecular Sciences.
- [13] : Vaquier, C., Nedellec, M., et al. (2019). "Les AVC hémorragiques, état des connaissances". Soins Gériatrie.

- [14] : Qureshi, A. I., et al. (2001). "Intracerebral hemorrhage". *New England Journal of Medicine*, 344(19).
- [15] : LAMRI, M. A. (2020). Détection et classification des hémorragies intracrâniennes par Deep Learning. Mémoire de Master, Université de Boumerdes.
- [16] : Chilamkurthy, S., et al. (2018). "Deep learning for head CT scans". *The Lancet*, 392(10160).
- [17] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), 436–444.
- [18] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [19] Shinde, P. P., & Shah, S. (2018). A Review of Machine Learning and Deep Learning Applications. *IEEE*.
- [20] Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- [21] Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- [22] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- [23] Chen, B., & Chen, S. (2026). *Deep Learning*. Springer.
- [24] Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.
- [25] Howard, A. G., et al. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv:1704.04861*.
- [26] Sinha, D., & El-Sharkawy, M. (2019). Thin MobileNet: An Enhanced MobileNet Architecture. *IEEE UEMCON*.
- [27] Qin, Z., et al. (2018). FD-MobileNet: Improved MobileNet with a Fast Downsampling Strategy. *IEEE ICIP*.

- [28] Chen, H. Y., & Su, C. Y. (2018). An Improved Hybrid MobileNet. International Conference on Systems and Signals.
- [29] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ICML.
- [30] Hoang, V. T., & Jo, K. H. (2021). Practical Analysis of EfficientNet Architecture.
- [31] Marques, G., Agarwal, D., & de la Torre Díez, I. (2020). Automated Medical Diagnosis of COVID-19 through EfficientNet Convolutional Neural Network. Applied Soft Computing.
- [32] Moutik, O., et al. (2023). Réseaux neuronaux convolutifs ou transformateurs de vision : qui remportera la course ?. Sensors.
- [33] Haroun, K. (2025). Exploration de la compression dynamique dans les Vision Transformers pour la classification d'images. Thèse de doctorat.
- [34] Abbaoui, W. (2024). Impact de l'IA sur la Santé Digitale : Cas des AVC et de la médecine de précision. Thèse.
- [35] Khenouche, C. (2024). Exploration de l'efficacité des modèles de transformateurs dans les tâches de classification du cancer.
- [36] Fawcett, T. (2006). An Introduction to ROC Analysis. Pattern Recognition Letters, 27(8), 861–874.
- [37] Elaouaber, Z. A. (2023). Système d'aide au diagnostic pour la détection automatique du glaucome et de la rétinopathie diabétique. Thèse de doctorat, Université de Tlemcen.
- [38] Turyamuhaki, N., Sameena, A., & Ahishakiye, E. (2026). Deep Ensemble Learning with Convolutional and Transformer Models for Brain Stroke Classification. Discover Artificial Intelligence.
- [39] Rajaramanan, G., Sathyavani, S. S., & Kumar, R. (2024). Classification of Brain Ischemia and Hemorrhagic Stroke Using a Hybrid Method. IEEE.
- [40] Okonkwo, C., Anyanele, A. N., & Emagbetere, J. O. (2025). Innovative Cognitive Stroke Detection Analysis Using Deep Learning Methodology. IEEE NIGERCON.
- [41] Tamilarasi, S., et al. (2026). Stroke Disease Classification from Computed Tomography Images Using Inception Harmonic LeNet and Wavelet-Symmetrically Weighted Local Gradient Pattern Features. Computational Biology and Chemistry.
- [42] Gupta, S., Sharma, P., & Verma, A. (2025). Brain Stroke Detection Using CT Images. IEEE International Conference on Computing and Communication.

- [43] Saleem, M., Ahmad, S., & Mahmud, T. (2026). StrokeFuse-AttnNet: A Hybrid Architecture for Robust Stroke Detection Using Global-Local Feature Fusion and Self-Attention. IEEE Access.
- [44] Sabir, M. A., & Ashraf, F. (2025). Development of a Novel Deep Convolutional Neural Network Model for Early Detection of Brain Stroke Using CT Scan Images. Multimedia Tools and Applications.
- [45] Inamdhar, M. A., et al. (2025). Dual Attention Mechanisms with Patch-Level Significance Embedding for Ischemic Stroke Classification in Brain CT Images. Informatics in Medicine Unlocked.
- [46] Kumar, A., Yadav, S. P., & Kumar, A. (2025). An Improved Feature Extraction Algorithm for Robust Swin Transformer Model in High-Dimensional Medical Image Analysis. Computers in Biology and Medicine.
- [47] Goodfellow, I., Bengio, Y., & Courville, A. Deep Learning. MIT Press, 2016.
- [48] Nielsen, M. Neural Networks and Deep Learning. Determination Press, 2015.
- [49] O. Aslan, "Brain Stroke CT Dataset," Kaggle, 2023. [Online]. Available: [Kaggle Dataset](#)
- [50] Flavio Pesenti, MD et al. RSNA Intracranial Hemorrhage Detection. Kaggle, 2019. <https://kaggle.com/competitions/rsna-intracranial-hemorrhage-detection>
- [51] Gupta, V., Sharma, A., & Kumar, R. (2024). "Explainable Deep Learning Framework for Brain CT Scan Classification Using Grad-CAM Visualization." Journal of Medical Imaging and Artificial Intelligence, Vol. 7, No. 2, pp. 115–126.