



UNIVERSITE ECHAHID HAMMA LAKHDAR - EL OUED
FACULTÉ DES SCIENCES EXACTES
Département D'Informatique



Mémoire de Fin D'étude
Présenté pour l'obtention du Diplôme de
MASTER ACADEMIQUE

Domaine : **Mathématique et Informatique**

Filière : **Informatique**

Spécialité : **Systèmes Distribués et Intelligence Artificielle**

Présenté par :

- Chekima Nour EL Islem
- Mansouri Aridj

Thème

**Comparaison entre les performances de CNN
et SwinUnetR Transformer pour le
développement d'un système de segmentation
3D et classification automatique des tumeurs
cérébrales.**

Soutenue le 02-06- 2024 Devant le jury :

M. **Boucherit** Ammar MCA Président

M. **Retima** Farida MAA Rapporteur

M. **Hamoud** Meriem MCB Encadreur

Année Universitaire : 2023/2024

شكر وتقدير

الحمد لله الذي أعاننا على إتمام هذا العمل. الحمد لله، الحمد لله الذي بنعمته تتم الصالحات. كما نتقدم بخالص الشكر والامتنان لمشرفتنا الدكتورة مريم حمود على دعمها وتوجيهاتها التي لا تقدر بثمن طوال فترة إعداد هذه المذكرة.

ولوالدينا العزيزين لا توجد كلمات تعبر عن امتناننا. والشكر الجزيل لجميع أعضاء لجنة التحكيم الذين وافقوا على تحكيم عملنا المتواضع. نود أيضاً أن نشكر جميع أعضاء هيئة التدريس في قسم علوم الحاسب الآلي، خاصة أولئك الذين قاموا بتدريسنا طوال مسيرتنا الجامعية.

وأخيراً، إلى جميع زملائنا وأصدقائنا الذين ساعدونا، شكراً لكم على دعمكم لنا. مع خالص شكرنا وتقديرنا.



نور الإسلام:

لم تكن الرحلة قصيرة أبداً ولا ينبغي لها أن تكون. لم يكن الحلم قريباً ولا الطريق سهلاً، لكني فعلتها ونلتها. وكل ذلك بفضل الله أولاً. الحمد لله حباً وشكراً وامتناناً، الذي بفضلها ها أنا اليوم أنظر إلى حلم طال انتظاره وقد أصبح واقعاً أفخر به.

إلى سندي الدائم، وقوتي بعد الله، داعمتي الأولى والأبدية، "أمي"، أهديك هذا الإنجاز الذي لولا تضحياتك لما كان له وجود. ممتنة لأن الله قد اصطفاك لي من بين البشر أماً، يا خير سند وعوض. فأنت التي أنرت لي دربي. إلى من دعمني بلا حدود وأعطاني بلا مقابل، أبي.

إلى من قيل فيهم: "سَنَشُدُّ عُضْدَكَ بِأَخِيكَ"

إلى من مد يده دون كلل ولا ملل وقت ضعفي، "أخي"، أدامك الله ضلعاً ثابتاً لي.

إلى أخي الأكبر، رغم رغبتني في حضورك إلا أن المسافات كانت العائق. وفقك الله في طريقك ويسره لك يا رب.

إلى أخوتي الصغار اللذين كانا يكسران الجو العكر ويرسمان البسمة على وجهي. شكراً لكم جميعاً.

إلى الذكورة المشرفة مريم حمود،

أود أن أتقدم لك بأسمى عبارات الشكر والامتنان على دعمك اللامحدود وإرشاداتك القيمة. لقد كنت دائماً نعم الموجه والمرشد، وكان لاهتمامك وتفانيك أثر كبير في تطوير معرفتي وصل مهاراتي. شكراً لك على كل ما قدمته من وقت وجهد، وعلى صبرك وتفهمك خلال جميع المراحل. بفضلك، أصبحت قادراً على تحقيق هذا الإنجاز بكل ثقة وفخر.

إلى صديقتي العزيزة أريج، شكراً لك. تعاونك وروحك المرحية جعلتنا من رحلتنا الأكاديمية تجربة ممتعة وملينة بالتحديات المثمرة. إلى صديقتي وزميلاتي، أرايتم؟ كانت رحلة قصيرة وها قد انتهت.

أريج:

الحمد لله شكراً وامتناناً على البدء والختام.

وآخر دعوانا أن الحمد لله رب العالمين.

بعد عناء السنين والسهرة، والفشل الذي جعلت منه سلماً يوصلني إلى طريق الناجحين، ها أنا قد وصلت.

في اللحظة الأكثر فخرًا أهدي عملي هذا إلى من أحمل اسمه بكل افتخار، إلى من كَلَّمَهُ اللهُ بالهبة والوقار، الذي حصد أشواك عن دربي ليجهد لي طريق العلم. طاب بك العمر يا سيد الرجال، طبت لي عمراً يا "أبي وقرة عيني".

إلى قدوتي الأولى ومعنى الحب والتفاني، إلى من جعلت الجنة تحت قدميها، وسهلت لي الشدائد بدعائها، إلى التي احتضنتني قلبها قبل يدها، إلى القلب الدافئ والشمعة التي كانت لي في الليالي المظلمة، سر قوتي ونجاحي، التي طالما حلمت أن تقر عينها برؤيتي في يوم كهذا، "أمي".

إلى ظلي الثابت وأمان أيامي، إلى ملهبي ونجاحي من شددت عضدي بهم، "أخوتي وأخواتي".

إلى صديقة المواقف لا السنين، شريكة الدرب والطموح البعيد، إلى من كانت دوما موضع ابتلاء عثرات حياتي، صديقتي "نور".

ولا أنسى مؤطرتي في هذا العمل، الطاقة الإيجابية، الأخت، الصديقة، المرشدة، المساعدة.... لا تكفي الكلمات لوصفها، "حمود مريم".

وفي الختام، أتمنى أن نخطى بأيام تليق بنا، وبضحكات تنعش قلوبنا، وأن نمضي في دروب نجحنا.

Résumé

La détection des tumeurs cérébrales à l'aide de l'apprentissage profond, de Transformer et de l'apprentissage automatique est un domaine essentiel et prometteur, car les gliomes sont classés comme des tumeurs primaires parce qu'ils proviennent des cellules du cerveau. En exploitant les progrès continus de l'intelligence artificielle, il est désormais possible d'améliorer considérablement le diagnostic précoce des tumeurs cérébrales en analysant avec précision les images médicales et en identifiant les changements associés à la maladie. Transformer permet d'analyser les données des patients et de prédire les résultats potentiels. Cette capacité permet aux cliniciens de prendre des décisions thérapeutiques plus précises, tout en améliorant la rapidité et la précision du diagnostic et de la classification. Par conséquent, l'utilisation de Transformer et de l'apprentissage profond pour détecter le cancer du cerveau réduit le temps et les coûts associés à la maladie. En outre, cette approche améliore la qualité des soins aux patients et contribue à optimiser les résultats des traitements grâce à un diagnostic précis et rapide et à un traitement personnalisé. Les développements technologiques en cours conduisent à de nouvelles améliorations et applications pratiques des modèles Transformer et de l'apprentissage profond dans le domaine de la détection du cancer du cerveau. L'objectif principal de notre travail était d'appliquer ces modèles à l'analyse d'images médicales 3D d'imagerie par résonance magnétique (IRM) afin de classer et de détecter avec précision les cancers du cerveau. Pour ce faire, nous avons utilisé plusieurs techniques, notamment les réseaux neuronaux convolutifs (CNN), une méthode largement utilisée dans le domaine de l'apprentissage profond basé sur les réseaux neuronaux. Les réseaux neuronaux convolutifs (CNN) ont la capacité d'extraire automatiquement des caractéristiques de haut niveau à partir des données. Les régions d'intérêt qui en résultent sont recadrées, puis introduites dans un SVM pour la classification afin de distinguer les gliomes de bas grade (LGG) des gliomes de haut grade (HGG) sur la base de la quantification. Le modèle que nous proposons a donné des résultats très prometteurs, qui permettront aux cliniciens de prendre des décisions diagnostiques plus précises en fournissant des solutions d'imagerie médicale à haute résolution.

Mots clés

Tumeur au Cerveau, Glioma, Transformer, Classification, Machine learning, deep learning , SVM.

Abstract

Detecting brain tumors using deep learning, Transformer and machine learning is an essential and promising field, as gliomas are classified as primary tumors because they originate from brain cells. By exploiting ongoing advances in artificial intelligence, it is now possible to significantly improve the early diagnosis of brain tumors by accurately analyzing medical images and identifying changes associated with the disease. Transformer makes it possible to analyze patient data and predict potential outcomes. This capability enables clinicians to make more precise therapeutic decisions, while improving the speed and accuracy of diagnosis and classification. As a result, using Transformer and deep learning to detect brain cancer reduces the time and costs associated with the disease. What's more, this approach improves the quality of patient care and helps optimize treatment outcomes through accurate, rapid diagnosis and personalized treatment. Ongoing technological developments are leading to further improvements and practical applications of Transformer models and deep learning in the field of brain cancer detection. The main goal of our work was to apply these models to the analysis of 3D medical MRI images in order to accurately classify and detect brain cancer. To achieve this, we used several techniques, including convolutional neural networks (CNNs), which is a widely used method in the field of neural network-based deep learning. Convolutional neural networks (CNNs) have the ability to automatically extract high-level features from data. The resulting regions of interest are cropped and then fed to SVM for classification to distinguish low-grade gliomas (LGG) from high-grade gliomas (HGG) based on quantification. Our proposed model showed very promising results, which will enable clinicians to make more accurate diagnostic decisions by providing high-resolution medical imaging solutions.

Key words

Brain cancer, Glioma, Transformer, Classification, Machine learning, deep learning , SVM.

الملخص

يُعد التصوير بالرنين المغناطيسي التقنية الأكثر شيوعاً للتشخيص غير الجراحي لسرطان الدماغ، حيث يوفر صوراً دماغية فائقة الجودة. لتوفير تشخيص دقيق للورم الدبقي وتحديد درجة خبثه، تُعد عمليتا التجزئة والتصنيف أمراً ضرورياً لتحقيق الكشف التلقائي عن درجة الورم دون تدخل بشري.

تُعتبر تجزئة وتصنيف الورم الدماغي في الصور الطبية مهمةً معقدةً بسبب التغيرات المورفولوجية والموقع العشوائي للآفات الدماغية. في هذا السياق، أحرز التعلم العميق، ولا سيما تقنية Transformers، تقدماً كبيراً في مجالات الرؤية الحاسوبية وتحليل الصور الطبية.

تستند استراتيجيتنا إلى كتل تحويل التجزئة ثلاثية الأبعاد (D3) الفعالة المعروفة باسم SwinUnetR، بالإضافة إلى استخدام أربعة أنواع من التصوير بالرنين المغناطيسي لتحديد مواقع الأورام الدماغية وتجزئتها. يتم استخراج أهم السمات من الصور المنجزة ثم نقلها إلى مصنفي SVM و ANN لتحديد نوع الورم إلى: أورام دبقيّة منخفضة الدرجة (LGG) أو أورام دبقيّة عالية الدرجة (HGG).

أظهرت الاختبارات الصارمة كفاءة النظام المقترح، حيث حقق معدل دقة يزيد عن 94%. وتكمن ميزة نهجنا في قدرته على مساعدة أخصائيو الأشعة في تحديد وتصنيف الأورام الدبقيّة بدقة وفعالية.

الكلمات المفتاحية

التعلم العميق ، سرطان الدماغ ، الورم الدبقي ، التحويل ، التصنيف ، التعلم الآلي

SVM

TABLE DES MATIÈRES

Table des matières	i
Table des figures	v
Liste des tableaux	viii
Introduction générale	1
I Les paradigmes de l'intelligence artificielle : l'apprentissage automatique, l'apprentissage profond et les Transformers	5
I.1 Introduction	6
I.2 Définition de la vision par ordinateur	6
I.3 Définition de l'analyse d'image et ses divers niveaux	7
I.3.1 Analyse de bas niveau d'image	7
I.3.2 Analyse de haut niveau d'image	8
I.4 La texture d'image et son analyse	8
I.5 Les atouts de l'intelligence artificielle : l'apprentissage automatique et l'apprentissage profond	9
I.5.1 L'apprentissage automatique	9
I.5.2 L'apprentissage profond	14
I.6 Les différents modèles de réseaux de neurone convolutifs	21
I.6.1 U-Net	21
I.6.2 SegNet, VggNet et AlexNet	22
I.6.3 ResNet	24

I.7	Transformer	24
I.7.1	Swin Transformer	26
I.7.2	Swin UNET	27
I.7.3	Swin UNETR	27
I.7.4	TransUNet	28
I.7.5	U-Netmer : U-Net meets Transformer	28
I.7.6	MMCFormer : Missing Modality Compensation Transformer	29
I.7.7	CoTr	30
I.7.8	HiFormer : Hierarchical Multi-scale Representations Using Transformers for Medical Image Segmentation	31
I.7.9	UNetFormer	32
I.8	Conclusion	33
II Exploration des paradigmes de l'intelligence artificielle pour le diagnostic automatique du cancer du cerveau : problématique et état de l'art		34
II.1	Introduction	35
II.2	Le cancer du cerveau	35
II.2.1	Définition du cancer du cerveau	35
II.2.2	Types des tumeurs cérébrales	36
II.2.3	Dépistage du cancer du cerveau basé imagerie par résonance magnétique (IRM)	38
II.3	Diagnostic automatique des tumeurs cérébrales	40
II.4	Problématique tirée dans le pronostic automatique du cancer du cerveau	43
II.5	Etat de l'art sur l'aide au diagnostic médical du cancer du cerveau	44
II.6	Conclusion	47
III Approche proposée pour la segmentation 3D du gliome et le pronostic auto- matique en LGG ou HGG		48
III.1	Introduction :	49
III.2	Plan d'action	49
III.3	Méthodologie proposée pour le diagnostic automatique du cancer du cerveau	53
III.4	Segmentation des tumeurs cérébrales dans les IRM	55
III.4.1	Base de données utilisée	55
III.4.2	Prétraitement des IRM cérébrales	57

III.4.3 Segmentation des tumeurs cérébrales basée Swin UnetR	58
III.4.4 Segmentation des tumeurs cérébrales basée CNN (SegResNet)	60
III.5 Classification des gliomes en LGG ou HGG	63
III.5.1 Base de données utilisée pour la classification	63
III.5.2 Cropping des gliomes	64
III.5.3 Extraction et sélection des descripteurs	65
III.5.4 Classification des gliomes à bas grade et à haut grade par les machines à vecteurs de support (SVM)	73
III.6 Conclusion	74
IV Evaluation des performances de l'approche proposée	75
IV.1 Introduction :	76
IV.2 Environnement De Développement Et Outils Utilisés	76
IV.2.1 Langage Python	76
IV.2.2 Matériel de configuration	78
IV.3 Mesures de performance utilisées	79
IV.3.1 Matrice de Confusion	79
IV.3.2 Précision de la classification (Accuracy)	80
IV.4 Les codes sources utilisés	81
IV.5 Evaluation des performances de l'approche de segmentation et de classification du gliome basée SwinUnetR	84
IV.5.1 Evaluation des performances de la segmentation	84
IV.5.2 Evaluation des performances de la classification	91
IV.6 Evaluation des performances de l'approche de segmentation et de classification du gliome basée SegResNet	95
IV.6.1 Evaluation des performances de la segmentation	95
IV.6.2 Evaluation des performances de la classification.	98
IV.7 Comparaison des performances de notre approches vs d'autres approches	101
IV.8 Présentation de l'interface de l'approche proposée	102
IV.8.1 L'interface du système de pronostic automatique du gliome basé SegResNet	102
IV.8.2 Notebook du système de pronostic automatique du gliome basé SwinU- netR	107
IV.9 Conclusion	110

TABLE DES FIGURES

I.1	Problème de classification à deux classes avec une séparation linéaire et une séparatrice non linéaire. Source [1]	11
I.2	Architecture d'un CNN. Source [2]	16
I.3	Pronostic automatique des IRM cérébrales basée sur un CNN. Source [3]	17
I.4	Principe de la convolution. Source [4]	18
I.5	Illustration de Max Pooling et Average Pooling. Source [2]	19
I.6	La fonction d'activation ReLU (Rectified Linear Unit). Source [5]	20
I.7	Architecture de U-Net. Source [6]	21
I.8	Architecture de SegNet. Source [7]	22
I.9	Architecture de VggNet. Source [8]	23
I.10	Architecture d'AlexNet. Source [9]	23
I.11	Architecture de ResNet. Source [10]	24
I.12	Architecture de Transformer. Source [11]	25
I.13	Illustration de l'approche à fenêtres variables. Source [12]	26
I.14	Architecture de Swin Transformer. Source [12]	26
I.15	Architecture de Swin UNET. Source [13]	27
I.16	Architecture de Swin UNETR. Source [14]	28
I.17	Architecture de TransUNet. Source [15]	28
I.18	Architecture de U-Netmer. Source [16]	29
I.19	Architecture de MMCFormer. Source [17]	30
I.20	Architecture de CoTr. Source [18]	30
I.21	Architecture de HiFormer. Source [19]	32
I.22	Architecture de UNetFormer. Source [20]	32

II.1 Exemples de tumeurs cérébrale. Source [21]	37
II.2 Tumeur LGG en A et HGG en B. Source [22]	38
II.3 Exemples d'IRM multimodale de tumeur cérébrale. Source [23]	40
II.4 Illustrations of different sub-regions for brain tumor segmentation. Source [24]	40
II.5 Fonctionnalités d'un système de pronostic automatique du cancer du cerveau. Source [25]	41
III.1 Architecture générale de l'approche proposée de segmentation et classification automatique du gliome	54
III.2 Résultats statistiques des bases de données les plus couramment utilisées. Source [24]	56
III.3 IRM cérébrales de la base de données BraTS 2023 avec leurs masques respectifs.	56
III.4 Prétraitement d'une IRM cérébrale de la base de données BraTS 2023 où en haut nous présentons les modalités originales et en bas suite à leur prétraitement par un filtre Gaussien suivi du skull stripping.	57
III.5 Architecture de Swin UnetR. Source [14]	58
III.6 Vue d'ensemble du mécanisme de fenêtrage décalé. Source [14]	59
III.7 Architecture de SegResNet. Source [26]	61
III.8 Pipeline adopté pour la classification des gliomes en LGG ou HGG.	63
III.9 Méthode proposée d'augmentation des données par la considération d'autres slices 80 et 85 et 90 respectivement.	64
III.10 Résultat du cropping d'une tumeur.	64
III.11 Courbes de Zipf et de Zipf inverse d'un gliome à haut grade (HGG).	66
III.12 Courbe de Zipf et de Zipf inverse d'un gliome à bas grade (LGG).	67
III.13 Application des filtres de Gabor dans le domaine fréquentiel.	69
III.14 Parties réelles des filtres de Gabor utilisés.	70
III.15 Parties réelles des tumeurs filtrées.	70
III.16 La magnitude de la réponse relative au filtre de Gabor suite à la convolution de la tumeur avec une banque de filtres de Gabor.	71
III.17 Structure générale de ResNet50. Source [27]	72
III.18 Application de la DWT 2D à une tumeur du gliome.	73
IV.1 Matrice de Confusion.	80
IV.2 Code source du modèle Transformer (SwinUnetR).	81
IV.3 Code source du ConvertToMultiChannelBasedOnBratsClassesd	82

IV.4 Utilisation de l'optimiseur AdamW.	83
IV.5 Code source du Skull stripping.	83
IV.6 Exemple du résultat généré par skull stripping sur une image IRM cérébrale. . .	84
IV.7 Evolution de la fonction d'Average Loss au long des époques par SwinUnetR. . .	84
IV.8 Evolution de la fonction de Val Mean Dice au long des époques.	85
IV.9 Evolution de la fonction de Val Mean Dice TC au long des époques.	86
IV.10 Evolution de la fonction de Val Mean Dice WT au long des époques.	86
IV.11 Evolution de la fonction de Val Mean Dice ET au long des époques.	87
IV.12 Exemples des résultats de la segmentation 3D du gliome basée Swin UnetR. . .	90
IV.13 Exemple du processus de cropping.	91
IV.14 Minimisation de loss function par le modèle SegResNet.	95
IV.15 Evolution de la fonction de Val Mean Dice au long des époques en utilisant le modèle SegResNet.	96
IV.16 Evolution de la fonction de Val Mean Dice TC, WT et ET au long des époques.	96
IV.17 Exemple de la segmentation du gliome basée SegResNet.	98
IV.18 Vue globale de l'interface du système de pronostic automatique du gliome basé SegResNet.	103
IV.19 Le choix d'une IRM cérébrale pour la réalisation du pronostic automatique. . .	103
IV.20 Un exemple de segmentation d'une IRM cérébrale.	104
IV.21 Un exemple de cropping de tumeur d'une IRM cérébrale.	104
IV.22 Pronostic de la tumeur gliome segmentée qui s'est avéré un gliome HGG. . . .	105
IV.23 Affichage de la tumeur gliome segmentée en 3D.	105
IV.24 Une autre vue 3D de la tumeur segmentée.	106
IV.25 Troisième vue 3D de la tumeur segmentée.	106
IV.26 Fonctionnalités du Notebook relatif au pronostic automatique du gliome basé SwinUnetR	109

LISTE DES TABLEAUX

II.1	Approches du pronostic automatique du cancer du cerveau.	47
IV.1	Caractéristiques du matériel local utilisées.	79
IV.2	Comparaison entre diverses architectures basées Transformer pour la segmenta- tion du gliome	90
IV.3	Précision de classification du gliome en HGG et LGG basée SwinUnetR et SVM.	92
IV.4	Précision de classification du gliome en HGG et LGG basée SwinUnetR et SVM ainsi que l'incorporation de l'ACP.	92
IV.5	Précision de classification du gliome en HGG et LGG basée SwinUnetR et ANN.	92
IV.6	les matrices de confusion relatives à la segmentation basée SwinUnetR et la classification du gliome en HGG et LGG basée SVM.	93
IV.7	les matrices de confusion relatives à la segmentation basée SwinUnetR et la classification du gliome en HGG et LGG basée ANN.	94
IV.8	Précision de classification du gliome en HGG et LGG par les SVM et issue d'une segmentation basée SegResNet.	99
IV.9	Précision de classification du gliome en HGG et LGG par les SVM et issue d'une segmentation basée SegResNet avec l'incorporation de l'ACP.	99
IV.10	Précision de classification du gliome en HGG et LGG basée SwinUnetR et ANN.	99
IV.11	Les matrices de confusion relatives à la classification du gliome en HGG et LGG basée SVM et issue d'une segmentation basée SegResNet.	100
IV.12	Les matrices de confusion relatives à la classification du gliome en HGG et LGG basée ANN et issues d'une segmentation basée SegResNet.	101
IV.13	Comparaison des performances de notre approches vs d'autres approches	102

INTRODUCTION GÉNÉRALE

La formation de cellules tissulaires irrégulières dans le cerveau humain est la principale cause de tumeur, autrement appelée néoplasme. Selon le rapport de « l’American Brain Tumor Association (ABTA) », chaque année, environ 6 00 000 personnes sont touchées par une tumeur cérébrale et entraînent la mort [22]. Le taux de mortalité dû aux tumeurs cérébrales augmente d’année en année dans le monde. Les tumeurs cérébrales touchent les hommes, les femmes et les enfants, quel que soit leur sexe. Les tumeurs cérébrales jusqu’à présent identifiées sont classées en fonction de leur localisation et de leur taux de croissance. Les tumeurs primaires et métastatiques sont les facteurs clés des tumeurs cérébrales localisées, tandis que les tumeurs bénignes et malignes sont les facteurs clés du taux de croissance des cellules cérébrales dans le cerveau humain. Les plus petites caractéristiques des cellules cérébrales sont clairement analysées par imagerie par résonance magnétique (IRM).

En effet, compte tenu de l’exactitude et l’intuition de l’IRM, l’examen des tumeurs cérébrales basé sur l’IRM est devenu le plus schéma couramment utilisé [24]. Dans ce travail, l’IRM est préférée à la tomодensitométrie en raison de la nécessité de connaître les plus petites caractéristiques des cellules cérébrales. De plus, l’identification des anomalies est clairement localisée grâce à l’IRM pour le processus de détection des tumeurs cérébrales.

Selon la société nationale des tumeurs cérébrales (NBTS) des États-Unis d’Amérique, il existe plus de 130 tumeurs cérébrales distinctes. Parmi eux, les gliomes sont fréquemment observés chez l’adulte. Généralement, les gliomes sont des tumeurs cérébrales primaires qui surviennent en raison de l’environnement (rayonnements des téléphones portables, consommation d’alcool, organismes pathologiques, tabagisme et conditions allergiques) et génétiques (modifications de l’isocitrate déshydrogénase 1 (IDH1) ou de l’isocitrate déshydrogénase 2 (IDH2)) [28]. Généralement, les gliomes sont classés en bas grade (LGG) et haut grade (HGG). Les tumeurs LGG se développent lentement tandis que les tumeurs HGG se développent rapidement. Par conséquent, le taux de survie d’un patient atteint de tumeurs HGG est inférieur à celui atteint d’une tumeur LGG. S’ils ne sont pas identifiés dans un délai imparti, cela aura de graves conséquences sur l’état de santé du patient et peut parfois même entraîner la mort. En effet, les gliomes de haut grade continuent à se propager l’environnement avec les fibres nerveuses, et peuvent même métastaser à d’autres tissus avec épanchement cérébral, sang, lymphe. Les patients ont généralement besoin d’un traitement complet tel qu’un traitement chirurgical, radiothérapie et chimiothérapie [29]. Dans ce sens, un diagnostic précoce et précis des tumeurs est essentiel où le pronostic automatique basée sur les paradigmes de l’intelligence artificielle a donné d’impressionnants résultats en les considérant comme deuxième avis aux radiologues durant l’élaboration de leur diagnostic. En effet, ceci désigne la problématique principale de ce

travail où nous nous trouvons faces à différentes contraintes comme suit :

- Les réseaux de segmentation basés sur CNN ne peuvent pas capturer efficacement les informations contextuelles à longue distance [30].
- La complexité informatique des méthodes basées sur Transformer est énorme, en particulier dans les tâches de segmentation 3D.
- Les Transformers standard sont coûteux durant le training et ils ne fonctionnent généralement pas aussi bien que les CNN en cas de données insuffisantes.
- La plupart des réseaux basés sur des Transformers existants sont conçus pour des tâches de segmentation 2D, qui ne peuvent pas être bien étendues à la segmentation 3D.

Compte tenu de ces points problématiques où les opérations convolutives se concentrent davantage sur les textures locales et que d'autre part, le Transformer excelle dans ses puissantes capacités de modélisation globale, compensant le champ de réception limité des CNN, nous proposons un système de segmentation et classification automatique des tumeurs cérébrales basé SwinUnetR pour la tâche de segmentation et les machines à vecteurs de support ainsi que les réseaux de neurones artificiels pour la classification. En effet, nous opérerons une segmentation 3D par le biais d'une IRM multimodale (T1, T1C, T2, FLAIR) où la fusion des images IRM multimodales va considérablement augmenter la capacité des cliniciens à évaluer précisément les tumeurs cérébrales [29]. En outre, nous exposerons le résultat de segmentation des tumeurs cérébrales en format 3D.

La structuration de ce mémoire est la suivante :

Chapitre I : *Les paradigmes de l'intelligence artificielle : l'apprentissage automatique, l'apprentissage profond et les Transformers*

Au long du premier chapitre, nous allons évoquer les définitions de l'analyse d'image et de la vision par ordinateur avec une extension vers l'apprentissage automatique ainsi que l'apprentissage profond et les Transformers bien évidemment.

Chapitre II : *Exploration des paradigmes de l'intelligence artificielle pour le diagnostic automatique du cancer du cerveau : problématique et état de l'art*

Quant au chapitre II il sera consacré à une initiation à la fatale maladie du cancer du cerveau où nous nous baserons sur les paradigmes de l'intelligence artificielle pour le pronostic automatique de cette maladie tout en soulignant notre problématique.

Chapitre III : *Approche proposée pour la segmentation 3D du gliome et le pronostic automatique en LGG ou HGG*

Le chapitre III détaillera l'approche proposée en réponse à la problématique.

Chapitre IV : Evaluation des performances de l'approche proposée

Quant au chapitre IV, nous exposerons une évaluation de pointe de l'approche proposée. Nous achèverons ce mémoire par une conclusion générale et des perspectives.

CHAPITRE I

LES PARADIGMES DE L'INTELLIGENCE
ARTIFICIELLE : L'APPRENTISSAGE
AUTOMATIQUE, L'APPRENTISSAGE
PROFOND ET LES TRANSFORMERS

I.1 Introduction

Le monde a connu un développement sans précédent, notamment ces dernières années, grâce à l'intelligence artificielle, qui a touché tous les domaines, notamment l'économie, la production, l'agriculture, et surtout la médecine et la santé à travers l'analyse des images médicales et le diagnostic précoce des maladies. Il est parfois difficile de diagnostiquer l'imagerie cérébrale (IRM) en raison du faible contraste, ce qui rend difficile la distinction entre les lésions et les tissus à l'œil nu, ce qui entraîne une augmentation des erreurs médicales. Contrairement, les approches de l'intelligence artificielle (IA) analysent les images IRM avec une plus grande précision comparées aux méthodes traditionnelles.

Dans ce chapitre, nous exposerons les différents modèles d'intelligence artificielle utilisés pour analyser les images et extraire des caractéristiques distinctives durant le processus de classification où nous nous intéressons particulièrement à l'imagerie médicale.

I.2 Définition de la vision par ordinateur

La Computer Vision ou Vision par Ordinateur est une technologie d'intelligence artificielle qui analyse une ou des vidéos puis les interprète. Au même titre que le NLP pour le texte, les méthodes de Computer Vision ont pour objectif d'interpréter des images ou des vidéos de la même manière qu'un humain. Autrement dit, elles cherchent à donner à un ordinateur la capacité de voir[31].

Des filtres de Snapchat aux voitures autonomes, en passant par la détection de cancers, la vision par ordinateur est aujourd'hui partout autour de nous. Elle est aussi efficace qu'elle balaie des domaines variés .

L'apprentissage automatique utilise des modèles algorithmiques qui permettent à un ordinateur de se former seul au contexte des données visuelles. Si suffisamment de données sont introduites dans le modèle, l'ordinateur « regardera » les données et apprendra à distinguer une image d'une autre. Les algorithmes permettent à la machine d'apprendre par elle-même, plutôt que quelqu'un la programme pour reconnaître une image[31].

Un réseau de neurones convolutifs (CNN) aide un modèle d'apprentissage automatique ou d'apprentissage en profondeur à « regarder » en décomposant les images en pixels auxquels sont attribuées des balises ou des libellés [31].

I.3 Définition de l'analyse d'image et ses divers niveaux

L'analyse d'images est un domaine interdisciplinaire qui consiste à extraire des informations utiles à partir d'images numériques. Cette spécialité est utilisée dans de nombreux domaines tels que la vision par ordinateur, l'imagerie médicale, la reconnaissance de formes, la vidéo-surveillance, etc. Différentes techniques d'analyse d'images peuvent être classées en plusieurs niveaux en fonction du type d'informations requises et des méthodes utilisées pour les extraire. L'analyse d'image comporte plusieurs étapes : le traitement de l'image, l'extraction de caractéristiques visuelles, de textures et de motifs ainsi que la segmentation de l'image. Les méthodes utilisées peuvent inclure des algorithmes d'apprentissage automatique, des réseaux de neurones et un traitement avancé du signal [32].

L'analyse d'images a ses propres défis et applications. Les progrès récents en matière d'apprentissage automatique et d'intelligence artificielle ont considérablement enrichi les capacités d'analyse d'images, permettant des applications plus avancées dans divers domaines [32].

I.3.1 Analyse de bas niveau d'image

L'analyse d'image de bas niveau présente un rôle crucial dans le pipeline global d'analyse d'image. En fournissant des informations de base telles que la segmentation, le filtrage et l'extraction de caractéristiques, cette phase de traitement prépare l'image pour des analyses plus avancées et complexes.

Dans de nombreuses applications, notamment dans le domaine médical, la télédétection et la vision par ordinateur, l'analyse d'image de bas niveau est une étape incontournable. Elle permet de filtrer et de préparer les données d'image, en extrayant les éléments clés et en réduisant le bruit ou les distractions qui pourraient entraver les analyses ultérieures.

La précision et la performance des techniques d'analyse de niveau supérieur dépendent fortement de la qualité des informations fournies par l'analyse d'image de bas niveau. Des erreurs ou des imprécisions dans cette phase peuvent entraîner des résultats incorrects ou biaisés dans les étapes suivantes de l'analyse.

Par conséquent, investir dans des techniques efficaces d'analyse d'image de bas niveau est essentiel pour garantir la fiabilité et l'exactitude des processus d'analyse d'image dans un large éventail d'applications. Cela permet également de maximiser le potentiel des algorithmes d'analyse de niveau supérieur en leur fournissant des données d'entrée de haute qualité et bien préparées.

I.3.2 Analyse de haut niveau d'image

L'analyse d'images de haut niveau va au-delà des techniques de traitement d'image de base et implique des algorithmes et des méthodes plus avancés pour extraire des informations complexes et comprendre le contenu de l'image (la reconnaissance). Cette phase vise à interpréter et à extraire des caractéristiques significatives de l'image qui permettent une compréhension approfondie de son contenu[33]. Nous allons citer des techniques couramment utilisées dans l'analyse d'images de haut niveau :

- Reconnaissance d'objets.
- Interprétation de scènes.
- Apprentissage en profondeur (Deep Learning).
- Segmentation sémantique.
- Génération d'images.

L'analyse d'images de haut niveau repose sur des techniques avancées qui permettent d'extraire des informations riches et complexes à partir d'images numériques[33].

I.4 La texture d'image et son analyse

L'analyse de la texture de l'image est une technique utilisée dans le traitement d'image pour extraire des informations sur la distribution spatiale des niveaux de gris ou de couleur dans une région spécifique d'une image. Elle vise à identifier les différences locales de densité de pixels, qui peuvent être utilisées pour identifier et distinguer différents motifs et structures dans l'image.

Nous distinguons divers points clés liés à l'analyse de la texture de l'image : détection de motifs et de structures, caractérisation de rugosité et de régularité, classification et segmentation d'images et beaucoup d'autres[34].

Il existe également des techniques couramment utilisées pour analyser la texture de l'image [35] :

- Matrices de cooccurrence (GLCM).
- Transformation en ondelettes d'image.
- Graphiques de gradients dirigés (HOG).
- Filtres en tissu.
- Transformée de Fourier.

- Réseaux de neurones convolutifs (CNN).
- Filtre de Gabor.
- Les lois de Zipf et Zipf inverse.

Ces techniques peuvent être utilisées seules ou en combinaison entre elles en fonction des besoins spécifiques de l'application d'analyse de texture d'image. Chacune a ses avantages et ses limites, et le choix de la technique dépend souvent de la nature des données et des objectifs de l'analyse. L'analyse de la texture de l'image est une technique puissante pour extraire des informations pertinentes sur les motifs et les structures d'une image. Elle trouve des applications dans de nombreux domaines et peut être utilisée pour la classification des textures, la segmentation, la détection d'anomalies et dans autres tâches d'analyse d'images.

I.5 Les atouts de l'intelligence artificielle : l'apprentissage automatique et l'apprentissage profond

L'apprentissage automatique (machine Learning) et l'apprentissage profond (Deep Learning) sont une branche de l'intelligence artificielle qui fonctionne de manière différente mais quelque peu complémentaire et est principalement utilisée pour résoudre des problèmes spécifiques. Ces technologies ont été utilisées dans de nombreux domaines dont les plus importants sont :

- Vision par ordinateur (comme la classification d'images, la détection d'objets, le suivi, la reconnaissance faciale et l'imagerie médicale).
- Traitement du langage naturel (comme la génération de texte et la traduction automatique).
- Robotique (comme Entraîner un robot à s'adapter à son environnement).

Ces modèles peuvent également reconnaître des modèles complexes dans les images, le texte, les sons et d'autres données pour produire des informations et des prédictions précises. Il est important de mentionner que l'apprentissage automatique et l'apprentissage profond se sont largement répandus dans tous les domaines et connaissent un développement rapide.

I.5.1 L'apprentissage automatique

Le machine Learning (ML) est une forme d'intelligence artificielle (IA) axée sur la création de systèmes qui apprennent, ou améliorent leurs performances, en fonction des données qu'ils

traitent [36]. De manière générale, deux principaux types d'algorithmes d'apprentissage automatique sont utilisés aujourd'hui : l'apprentissage supervisé et l'apprentissage non supervisé. Dans l'apprentissage automatique supervisé, un algorithme apprend à l'aide d'un ensemble de données déjà étiqueté et dont le résultat est prédéterminé. Des exemples d'apprentissage automatique supervisé incluent des algorithmes tels que la régression linéaire et logistique, la classification multi classe et les machines à vecteur de support. Quant à l'apprentissage automatique non supervisé, il utilise une approche plus indépendante dans laquelle l'ordinateur apprend à identifier des processus et des modèles complexes sans aucune guidance humaine cohérente et stricte. L'apprentissage automatique non supervisé implique une formation basée sur les données, sans étiquettes ni résultats spécifiques.

Le choix entre un algorithme d'apprentissage automatique supervisé ou non supervisé dépend généralement de facteurs liés à la structure et à la taille de données, ainsi que du cas d'utilisation prévu. La différence entre les deux est déterminée par la méthode utilisée pour traiter les données afin de faire des prédictions [36].

Nous allons citer quelques algorithmes d'apprentissage automatique comme suit :

- **Machines à Vecteurs de Support**

Les machines à vecteurs de support (SVM) comme mentionné sur la figure I.1 sont une classe de modèles d'apprentissage supervisé utilisés pour la classification et la régression. Il est particulièrement populaire dans l'apprentissage automatique en raison de sa capacité à bien généraliser à partir d'un ensemble de données relativement petit. Quelques points de base sur les SVM : Premièrement, les SVM cherchent à trouver l'hyperplan qui sépare le mieux les données en différentes classes.

Deuxièmement, la fonction de décision SVM : cette fonction attribue une classe à chaque point de données en fonction de son emplacement par rapport à l'hyperplan.

Troisièmement, bien que les SVM trouvent initialement un niveau hyperlinéaire de séparation des données, ils peuvent être étendus pour gérer les problèmes de classification non linéaire en utilisant des techniques telles que l'astuce du noyau.

Quatrièmement, la formation SVM consiste à résoudre un problème d'optimisation convexe visant à maximiser la marge tout en minimisant l'erreur de classification. Utiliser des techniques telles que la descente de gradient et des algorithmes de programmation quadratique[37].

Les SVM sont également intrinsèquement régularisés, ce qui signifie qu'ils sont robustes contre le surajustement. Cela est dû à leur objectif de maximisation de la marge, qui privilégie des

limites de décision simples et générales.

Les machines à vecteurs de support (SVM) présentent plusieurs avantages pour les problèmes de classification et de régression :

- Les SVM fonctionnent bien même dans des espaces de grande dimension, ce qui en fait un choix approprié pour les problèmes de classification impliquant de nombreux attributs ou fonctionnalités[38].
- Les SVM sont naturellement régularisés, ce qui signifie qu'ils sont moins sujets au sur-ajustement.
- Les SVM sont également utilisés pour la régression en ajustant les marges autour de la ligne de régression plutôt que de simplement la séparer. Cela le rend polyvalent et peut être appliqué à une variété de tâches[39].
- Grâce à l'astuce du noyau, les SVM peuvent être utilisées efficacement pour résoudre des problèmes de classification non linéaire. Cela permet de capturer les relations complexes entre les variables d'entrée et les étiquettes de sortie [39].
- Les SVM sont efficaces pour les ensembles de données de taille moyenne à grande[39].
- Les SVM offrent la possibilité de choisir différentes fonctions du noyau en fonction de la nature des données et du problème à résoudre. Cela permet d'adapter le modèle aux caractéristiques spécifiques des données.
- Il est capable de s'adapter à des problèmes non linéaires, ce qui en fait un outil précieux pour de nombreuses applications de classification et de régression [40].

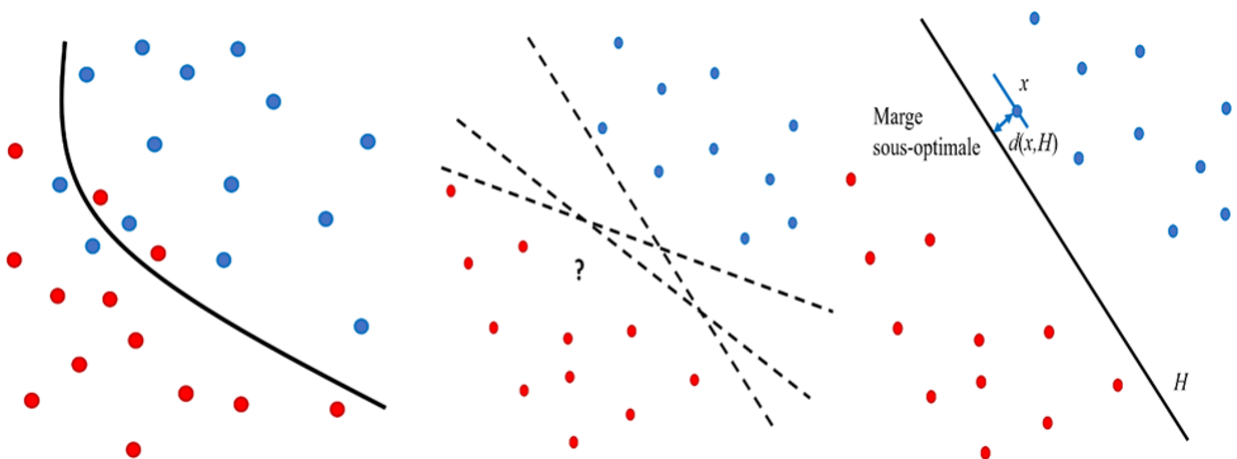


FIGURE I.1 – Problème de classification à deux classes avec une séparation linéaire et une séparatrice non linéaire. Source [1]

- **Les réseaux de neurones**

Les réseaux de neurones sont un ensemble de modèles d'apprentissage automatique inspirés du fonctionnement du cerveau humain [41]. Il se compose de plusieurs couches de neurones interconnectés qui transforment les entrées en sorties requises par le processus d'apprentissage. Les réseaux de neurones offrent une combinaison unique de puissance, de flexibilité et d'efficacité, ce qui en fait des outils précieux pour de nombreuses applications d'apprentissage automatique et de traitement de données.

En fait, les réseaux de neurones présentent de nombreux avantages, ce qui explique leur popularité croissante dans divers domaines. Voici quelques-uns des principaux avantages des réseaux de neurones :

- Les réseaux de neurones peuvent apprendre des relations complexes et non linéaires entre les données à l'aide d'architectures approfondies et d'algorithmes d'apprentissage personnalisés [42].
- Les réseaux de neurones peuvent être utilisés pour diverses tâches d'apprentissage automatique, notamment la classification, la régression, la génération de texte, la détection d'objets et la segmentation d'images.
- Les réseaux de neurones sont particulièrement adaptés au traitement de données non structurées telles que les images, le texte et l'audio.
- Les réseaux de neurones peuvent être mis en œuvre de manière à tirer parti du parallélisme massif fourni par les unités de traitement graphique (GPU) et les systèmes distribués.
- Les réseaux de neurones peuvent être constamment mis à jour et améliorés à mesure que de nouvelles données deviennent disponibles. Cela leur permet de s'adapter à des environnements changeants et d'améliorer leurs performances au fil du temps [42].

Les réseaux de neurones peuvent apprendre automatiquement les fonctionnalités pertinentes à partir des données brutes, éliminant ainsi le besoin d'ingénierie manuelle des fonctionnalités. Cela simplifie le processus de modélisation et peut conduire à de meilleures performances sur des ensembles de données complexes [43].

Bien entendu, malgré leurs nombreux avantages, les réseaux de neurones présentent également certains inconvénients, parmi lesquels :

- Les réseaux de neurones, en particulier les architectures profondes, nécessitent souvent de grandes quantités de données pour s'entraîner efficacement [44].

- Les réseaux de neurones, notamment les modèles profonds, peuvent être difficiles à interpréter en raison de leur complexité [44].
- Les réseaux de neurones peuvent être vulnérables au surapprentissage, en particulier lorsqu'ils sont formés sur de petits ensembles de données [44].

• Les K plus proches voisins (K-NN : K Nearest Neighbors)

L'algorithme des k plus proches voisins, également connu sous le nom de KNN ou k-NN, est un discriminant d'apprentissage supervisé non paramétrique, qui utilise la proximité pour effectuer des classifications ou des prédictions sur le regroupement d'un point de données individuel [45]. Bien qu'il puisse être utilisé pour des problèmes de régression ou de classification, il est généralement utilisé comme algorithme de classification, en partant de l'hypothèse que des points similaires peuvent être trouvés les uns à côté des autres [45].

Voici quelques avantages associés à l'algorithme K-Nearest Neighbours :

- Regardez simplement les étiquettes des K points les plus proches dans l'espace des fonctionnalités et prenez la classe majoritaire [45].
- Contrairement à certains autres algorithmes d'apprentissage supervisé, tels que les modèles linéaires, les K-NN ne font aucune hypothèse sur la distribution des données [45].
- Les K-NN peuvent modéliser des limites de décision complexes qui peuvent être non linéaires et avoir des formes arbitraires.
- Il peut s'adapter rapidement aux nouveaux exemples de données ajoutés à l'ensemble de données sans avoir à recycler l'intégralité du modèle.

Bien que les K-NN présentent de nombreux avantages, ils présentent également des limites, telles que leur sensibilité aux données bruitées et la présence de fonctionnalités redondantes, ainsi que leur coût de calcul élevé lors de la prévision de grands ensembles de données.

• Les arbres de décision

Les arbres de décision sont des modèles d'apprentissage supervisé utilisés pour la classification et la régression [46]. Ils ont été particulièrement appréciés pour leur simplicité conceptuelle et leur facilité d'interprétation. Ces règles sont représentées sous forme d'arbre, permettant une visualisation facile de la logique de décision. Cela les rend adaptées à une variété de tâches où les données sont hétérogènes. Voici quelques-uns des avantages associés aux arbres de décision :

- Les arbres de décision produisent des règles de décision simples et intuitives que les humains peuvent facilement interpréter et comprendre.

- Il gère efficacement des ensembles de données avec différents types de fonctionnalités (continues et catégorielles) sans nécessiter de prétraitement complexe.
- Les arbres de décision peuvent gérer efficacement les valeurs manquantes dans les données sans nécessiter de prétraitement supplémentaire.
- Les arbres de décision peuvent modéliser des relations non linéaires entre les fonctionnalités et l'objectif [46].
- Différentes techniques de régularisation telles que la taille maximale de l'arbre (profondeur maximale) ou le partitionnement minimum des nœuds peuvent être utilisées pour contrôler cet effet et améliorer la généralisation.

En résumé, les arbres de décision sont des modèles polyvalents et interprétables qui offrent une grande flexibilité pour modéliser des problèmes complexes dans différents domaines. Ils sont largement utilisés dans de nombreuses applications, notamment la classification, la régression, le marketing, la finance et la médecine.

En fait, même si les arbres de décision présentent de nombreux avantages, ils présentent également certains inconvénients. Certains des principaux inconvénients associés aux arbres de décision sont les suivants :

- Le surajustement des données d'entraînement, en particulier lorsque l'arbre est profond, affecte les performances [47].
- Cela crée des modèles trop interprétables, ce qui peut entraîner une perte de généralisation et de mauvaises performances sur les données de test [47].

I.5.2 L'apprentissage profond

C'est est une technique d'intelligence artificielle qui apprend aux ordinateurs à traiter les données d'une manière inspirée du cerveau humain. Les modèles d'apprentissage profond peuvent reconnaître des modèles complexes dans les images, le texte, les sons et d'autres données pour produire des informations et des prédictions précises. Des méthodes d'apprentissage profond peuvent être utilisées pour automatiser des tâches qui nécessitent généralement l'intelligence humaine, comme la description d'images ou la conversion d'un fichier audio en texte [48].

La technologie d'apprentissage profond alimente de nombreuses applications d'IA utilisées dans les produits du quotidien. Le deep learning est une branche de l'apprentissage automatique qui implique l'utilisation de réseaux de neurones artificiels composés de plusieurs couches de

traitement pour apprendre des représentations hiérarchiques de données [48]. L'apprentissage profond est flexible dans le traitement de différents types de données et de tâches et puissant pour s'adapter à de grandes quantités de données. Ces avantages le font considérer comme un outil puissant pour de nombreuses applications d'apprentissage automatique dans divers domaines. Citons d'autres avantages associés au deep learning :

- Les réseaux de neurones profonds peuvent apprendre des représentations hiérarchiques de données à partir de données brutes.
- Il a la capacité de s'adapter aux données non structurées.
- Ses performances sont élevées lorsqu'il s'agit de traiter de grandes quantités de données.
- Flexibilité et adaptabilité à diverses tâches.

Les CNN, ou réseaux de neurones convolutifs plus couramment appliqués en apprentissage profond, sont une classe de réseaux de neurones profonds spécifiquement conçus pour le traitement d'images [3]. Ils sont largement utilisés dans des tâches telles que la reconnaissance d'objets, la classification d'images, la détection d'objets, la segmentation sémantique et bien d'autres encore. Les CNN sont capables d'apprendre automatiquement les fonctionnalités pertinentes à partir des données d'entrée. Les couches de convolution et de pooling de CNN sont utilisées pour extraire les caractéristiques hiérarchiques des images. Les CNN sont robustes aux variations de position, de taille et d'orientation des objets dans les images. Ceci est rendu possible grâce à l'utilisation de couches de convolution, qui appliquent des filtres locaux aux régions de l'image, et à des couches de regroupement, qui réduisent les dimensions tout en préservant les caractéristiques importantes. Les CNN peuvent traiter de grandes images avec une efficacité de calcul élevée grâce à la stratégie de partage de poids des filtres de convolution. Cela réduit considérablement le nombre de paramètres du modèle par rapport aux architectures entièrement connectées. Les CNN pré-entraînés sur de grands ensembles de données, tels qu'ImageNet, peuvent être utilisés comme points de départ pour de nouvelles tâches d'apprentissage sur des ensembles de données plus petits. Grâce à des techniques telles que le réglage fin, les CNN peuvent être utilisés pour diverses tâches liées à la vision par ordinateur, telles que la classification, la détection d'objets, la segmentation sémantique, la reconnaissance faciale, etc. Les CNN ont démontré des performances de pointe dans de nombreux concours et défis de vision par ordinateur, surpassant souvent les méthodes traditionnelles sur des tâches telles que la classification d'images et la détection d'objets [3]. Les CNN offrent de nombreux avantages pour le traitement d'images, notamment :

- Sa capacité à extraire automatiquement les fonctionnalités pertinentes [49].

- Son pouvoir pour gérer les différences d'objets.
- Son efficacité dans le traitement des grandes images.
- Son adaptabilité à différentes tâches et ses performances les plus élevées dans de nombreuses applications [49].

Nous présentons sur la figure I.2 une architecture générale d'un CNN.

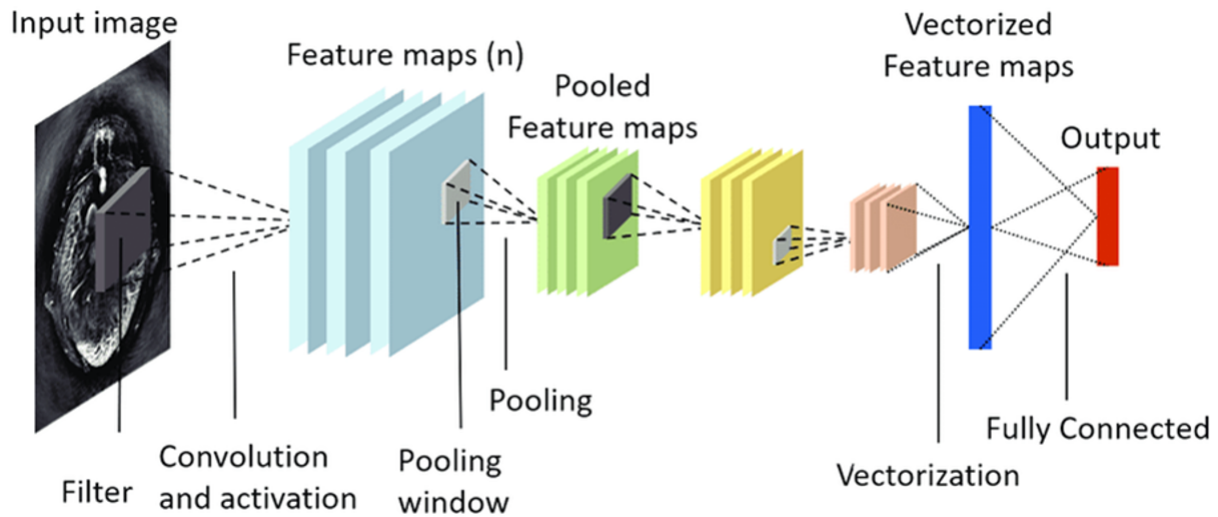


FIGURE I.2 – Architecture d'un CNN. Source [2]

Nous présentons sur la figure I.3 un pronostic automatique des IRM cérébrales basée sur un CNN.

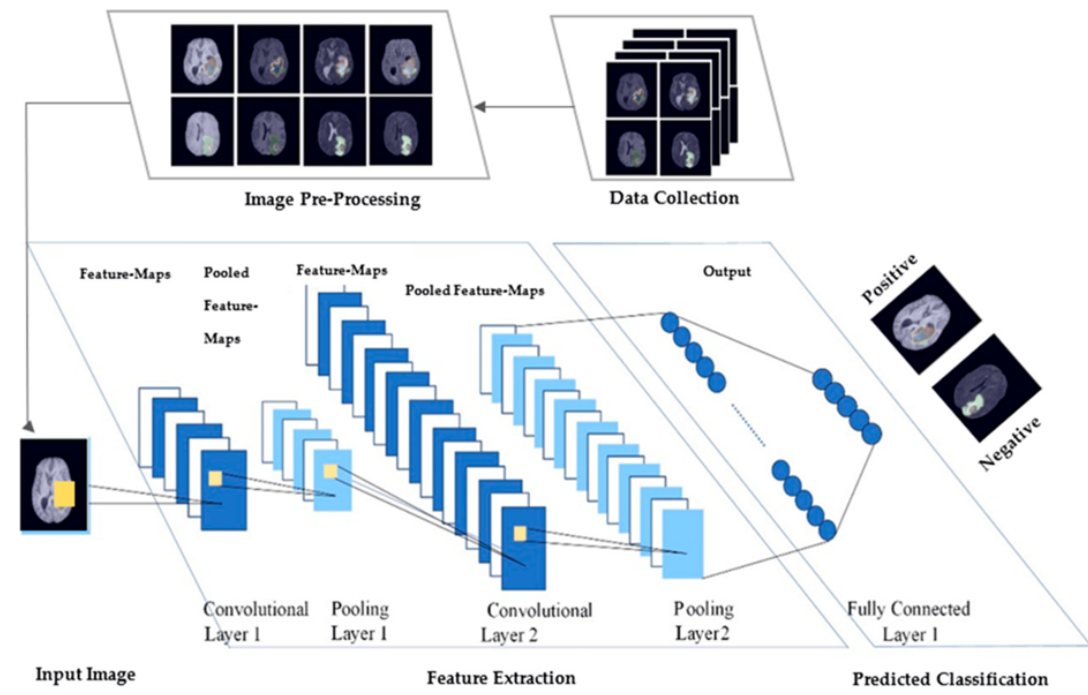


FIGURE I.3 – Pronostic automatique des IRM cérébrales basée sur un CNN. Source [3]

Nous allons définir le rôle de chaque couche comme suit :

- **La couche convolutive**

La couche convolutive comme mentionnée sur la figure I.4, est un composant clé des réseaux de neurones convolutifs (CNN). Cette couche est chargée d'extraire les caractéristiques importantes des données d'entrée, en particulier des images, via des opérations de convolution. La couche de convolution applique des filtres (ou noyaux) à la région locale des images. Les filtres de convolution agissent comme des détecteurs de caractéristiques en capturant des motifs locaux tels que les bords, les textures, et les motifs dans l'image. Ces filtres sont appris pendant la phase de formation du réseau neuronal, permettant au réseau d'apprendre à extraire les fonctionnalités pertinentes pour la tâche à accomplir [50].

La couche convolutive est un composant essentiel des réseaux de neurones convolutifs, qui joue un rôle crucial dans l'extraction et la représentation des caractéristiques des données d'entrée, en particulier des images, pour les tâches de vision par ordinateur telles que la classification, la détection d'objets et la segmentation sémantique [50].

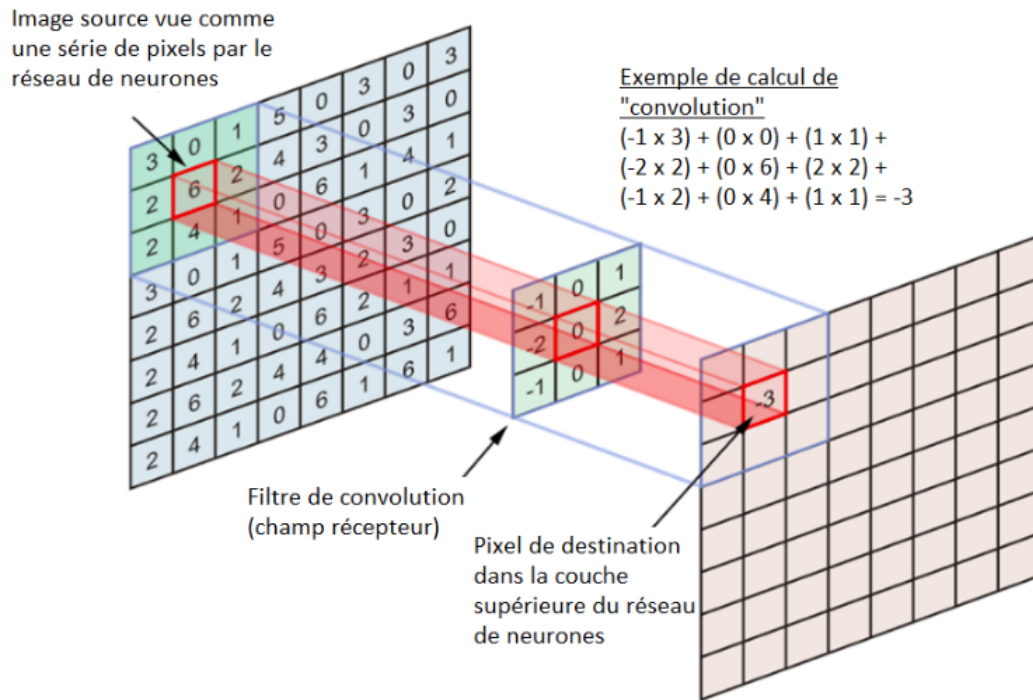


FIGURE I.4 – Principe de la convolution. Source [4]

- La Couche de pooling

Il existe également le processus de traitement par lots, où ce processus est utilisé pour convertir l'image d'entrée en une image plus petite comme mentionné sur la figure I.5.

Il vise à réduire spatialement l'entrée (en minimisant le nombre de paramètres) grâce à un processus de sous-échantillonnage. Il existe deux opérations de couche de pooling utilisées dans l'architecture Inception-v3, à savoir Max Pooling et Agrégation moyenne. La sommation maximale est une opération mathématique qui fonctionne en prenant la plus grande valeur une partie d'une image a une certaine taille, tandis que le regroupement moyen est une opération mathématique réussie, En prenant la valeur moyenne d'une partie d'une image d'une taille donnée.

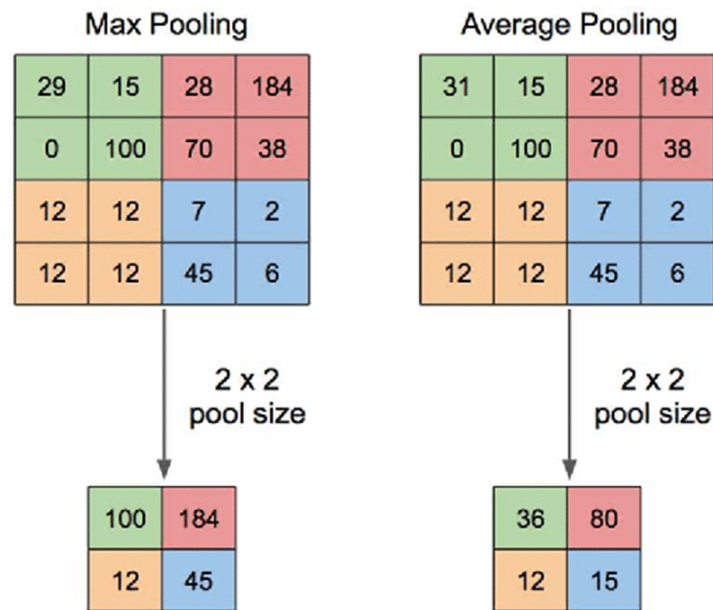


FIGURE I.5 – Illustration de Max Pooling et Average Pooling. Source [2]

• La Couche de correction ReLU

ReLU (Rectified Linear Unit) comme mentionné sur la figure I.6 est une fonction d'activation non linéaire à grande échelle utilisée dans les réseaux de neurones, y compris les réseaux de neurones convolutifs (CNN) [51]. Cette fonction s'applique en partie à la transformation d'une surface maillée ou de toute autre surface maillée linéaire. Cela présente certains avantages :

- Toutes les valeurs négatives sont converties en 0, donc aucune valeur négative n'est disponible [51].
- La valeur seuil maximale est l'infini, il n'y a donc pas de problème de disparition du gradient, de sorte que la précision et l'efficacité de la prédiction de sortie atteignent le maximum [51].
- La vitesse est rapide par rapport aux autres fonctions d'activation [51].

La fonction ReLU ne souffre pas du problème de mise à l'échelle du contraste et n'est pas sans inconvénients. Dans le cas d'entrées négatives, le gradient de la fonction ReLU asymétrique peut disparaître de zéro [51]. Quelques variables ajoutées :

- Leaky ReLU est une version de ReLU qui introduit une vision négative des valeurs négatives pour résoudre le problème des neurones morts [51].
- Parametric ReLU (PReLU) L'extension ReLU du paramètre Pente a été appliquée à la formation [51].
- L'unité linéaire exponentielle (ELU) est une autre ReLU qui offre des avantages en termes de stabilité et de vitesse de convergence [51].

ReLU est une fonction d'activation couramment utilisée dans les réseaux de neurones en raison de sa simplicité, de son efficacité et de son efficacité dans la résolution du problème de variance de gradient.

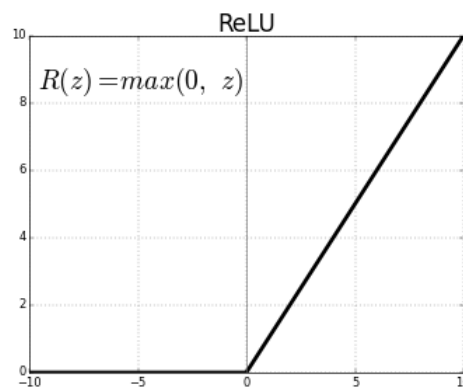


FIGURE I.6 – La fonction d'activation ReLU (Rectified Linear Unit). Source [5]

- **La couche entièrement connectée**

Généralement appelée couche de sortie, est une couche de neurones où chacun d'entre eux fut connecté avec l'intégralité des neurones de la couche précédente. Le rôle principal de cette couche est la classification ou la régression où les caractéristiques extraites à partir des couches précédentes sont considérées comme entrées pour et générer une sortie consistant en la classe de la cible. Donc, cette une sortie correspondra à la classification ou à la régression.

I.6 Les différents modèles de réseaux de neurone convolutifs

I.6.1 U-Net

En segmentation d'image, il faut non seulement la convertir en vecteur, mais aussi reconstruire l'image à partir de ce vecteur. C'est une tâche énorme, car convertir un vecteur en image est beaucoup plus difficile que l'inverse. Toute l'idée d'U-Net tourne autour de ce problème[6]. U-Net, dérivé du réseau neuronal traditionnel CNN, a été conçu et appliqué pour la première fois en 2015 pour le traitement d'images biomédicales. Il est capable d'identifier et de distinguer les limites des composants d'une image donnée en effectuant une classification sur chaque pixel. Entrons dans le détail de l'architecture de ce modèle pour comprendre comment il fonctionne et ce qui le rend si précis et adaptable à ces situations complexes. Nous présentons sur la figure I.7 l'architecture de U-Net.

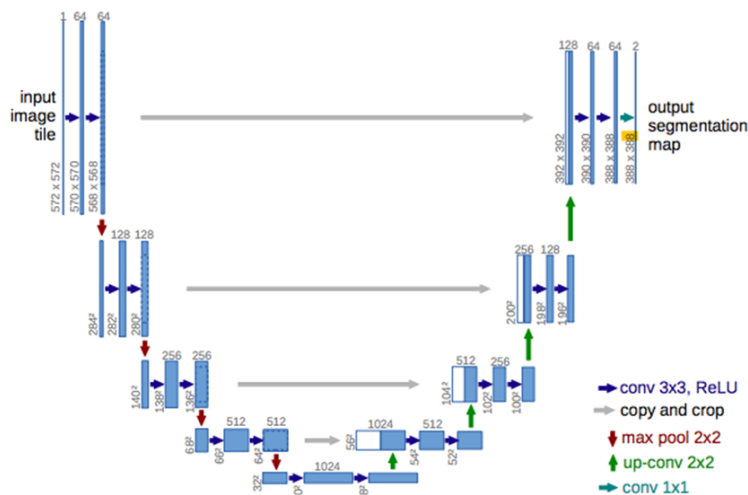


FIGURE I.7 – Architecture de U-Net. Source [6]

Visuellement, il a une forme en « U ». L'architecture est symétrique et se compose de trois sections : la section de contraction, la section de bottleneck et la section d'expansion. U-Net utilise une fonction de perte pour chaque pixel de l'image [6]. Une fonction Softmax est appliquée à chaque pixel suivie d'une fonction de perte. Cela transforme le problème de segmentation en un problème de classification où nous devons classer chaque pixel dans l'une des classes [6]. Une caractéristique très importante d'U-Net est le fait qu'il est capable d'effectuer une segmentation d'image en prédisant l'image pixel par pixel [6].

Le réseau est suffisamment puissant pour faire de bonnes prédictions basées sur peu de données

d'entraînement en utilisant des techniques supérieures d'augmentation des données. Ils offrent également une plus grande précision que les modèles traditionnels. En plus de sa robustesse, U-Net dispose d'une architecture simple basée sur des auto-encodeurs et des couches convolutives faciles à mettre en œuvre [6].

Il existe de nombreuses variantes de ce réseau qui s'adaptent à différentes situations, certaines très précises, d'autres très rapides.

U-Net est une architecture réseau traditionnelle pour une segmentation d'image rapide et précise. Elle a jusqu'à présent surpassé toutes les méthodes précédentes, notamment dans le domaine médical et dans la détection d'anomalies et de tumeurs[6]. Il existe de nombreuses applications d'U-Net dans la segmentation d'images biomédicales, telles que la segmentation d'images du cerveau (« BRATS ») etc.

I.6.2 SegNet, VggNet et AlexNet

Nous présentons sur les figures I.8,I.9 etI.10 respectivement les architectures relatives à SegNet, VggNet et AlexNet.

— SegNet

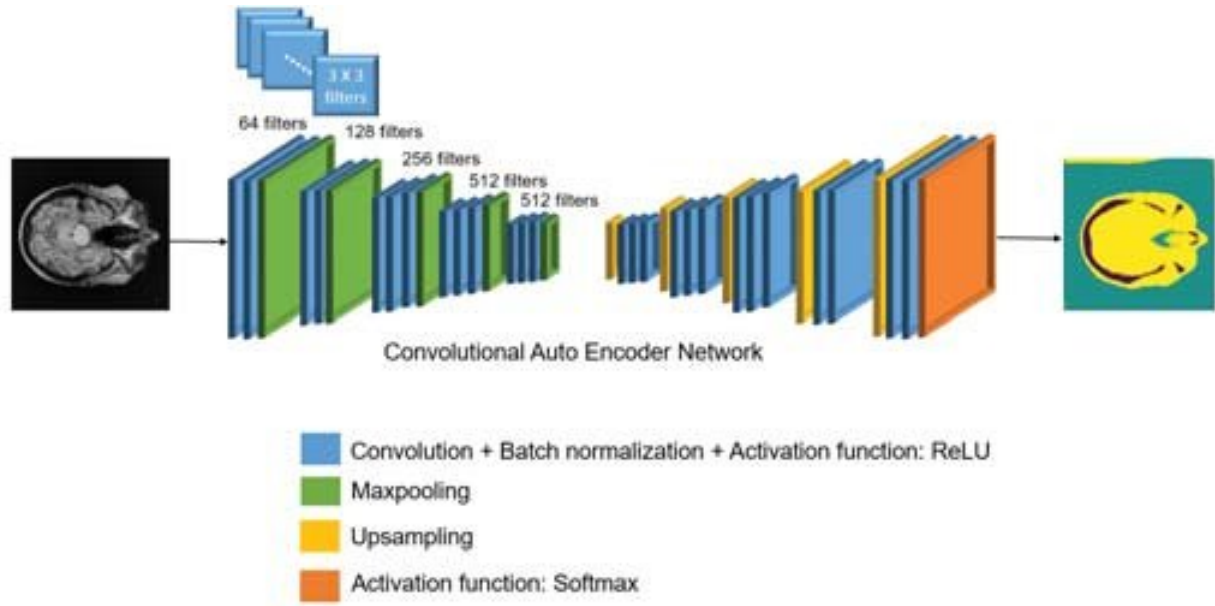


FIGURE I.8 – Architecture de SegNet. Source[7]

— VggNet

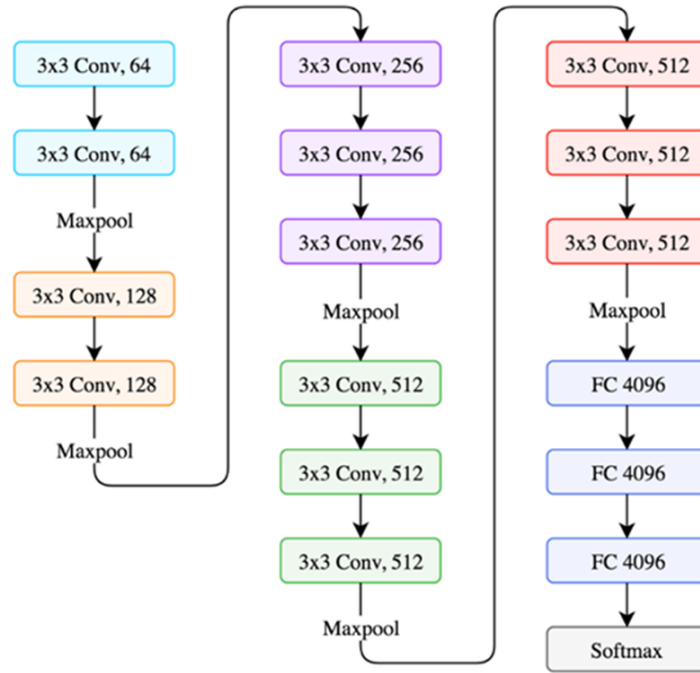


FIGURE I.9 – Architecture de VggNet. Source[8]

— AlexNet

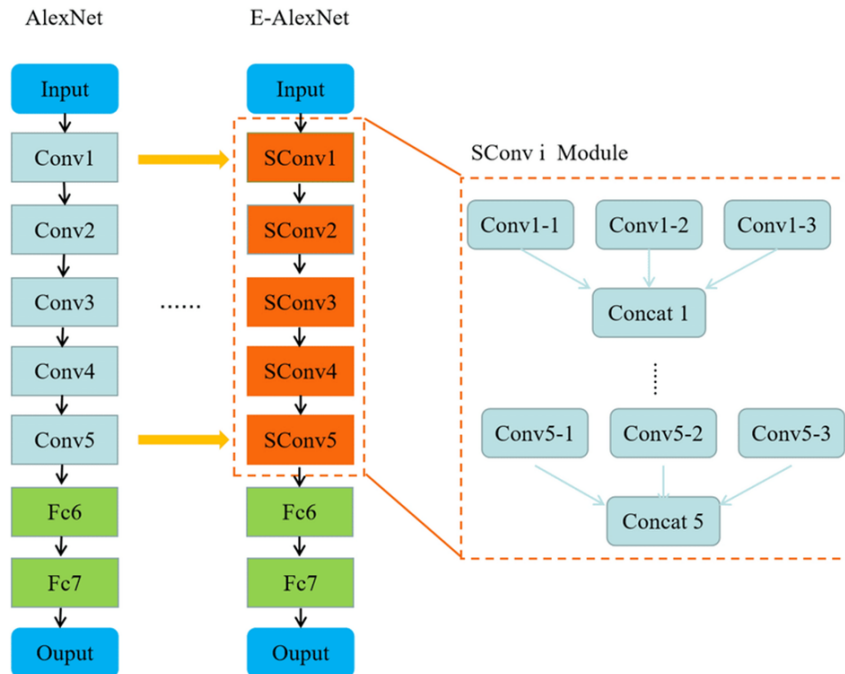


FIGURE I.10 – Architecture d'AlexNet. Source[9]

I.6.3 ResNet

Les réseaux résiduels sont un modèle d'apprentissage profond proposé en 2015 par des chercheurs de Microsoft, dans le but principal de surmonter les problèmes de disparition et d'explosion du gradient auxquels sont confrontés les réseaux neuronaux profonds.

Les réseaux résiduels sont l'un des modèles d'apprentissage profond connus pour leurs capacités exceptionnelles d'apprentissage et d'extraction de caractéristiques malgré leur faible complexité par rapport à d'autres modèles. Les ResNet peuvent comporter jusqu'à 152 couches, soit 8 fois plus que les réseaux VGG [52].

ResNet a été largement évalué sur une variété d'ensembles de données ouverts, y compris CIFAR 10, la classification ImageNet et COCO. Le mécanisme d'apprentissage résiduel est une méthode d'apprentissage générique facile à optimiser et efficace sur les architectures CNN profondes [53].

La figure I.11 présente l'architecture de ResNet.

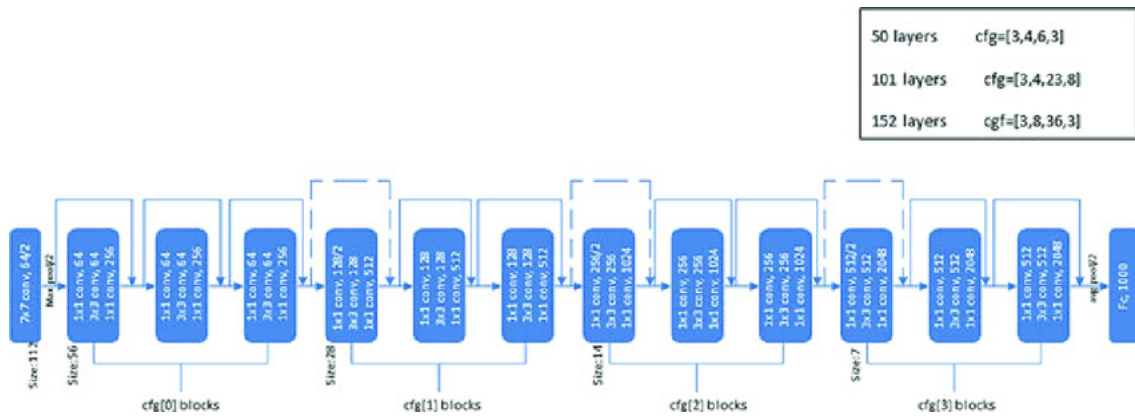


FIGURE I.11 – Architecture de ResNet. Source [10]

I.7 Transformer

Le Transformer est un modèle de Deep Learning (donc un réseau de neurones) qui a la particularité de n'utiliser que le mécanisme d'attention et aucun réseau récurrent ou convolutif. Le Transformer est présenté dans le papier « Attention Is All You Need » qui décrit son architecture et ses performances impressionnantes sur plusieurs jeux de données de traduction. Un modèle séquence à séquence est un modèle qui prend en entrée une séquence (une suite d'éléments du même type) et renvoie une séquence en sortie. L'exemple par excellence pour ce type de modèle est la traduction de texte.

Pour faire des modèles séquence à séquence avant la venue du Transformer, il fallait faire recours

au fameux LSTM (ou GRU) qu'on utilisait dans une architecture Encoder-Decoder. Nous allons présenter quelques points clés sur les Transformers :

- L'architecture du Transformer a du pattern Encoder-Decoder.
- L'idée du concept d'attention est de mesurer à quel point deux éléments de deux séquences sont liés.
- Self-attention ou Intra-attention.

L'un des modèles les plus importants de génération de langage naturel (NLP) dans lequel Transformer a été utilisé et qui est devenu célèbre dans le monde est ChatGPT (Generative Pre-trained Transformer), qui a utilisé la technologie Transformer pour représenter des textes, générer des réponses et prêter attention à se concentrer sur des parties spécifiques du contexte tout en générant le texte. Les transformateurs contiennent plusieurs éléments de base, dont les plus importants sont : l'attention, l'encodage, le décodage et les couches profondes. Nous présentons sur la figure I.12 l'architecture générale des Transformers.

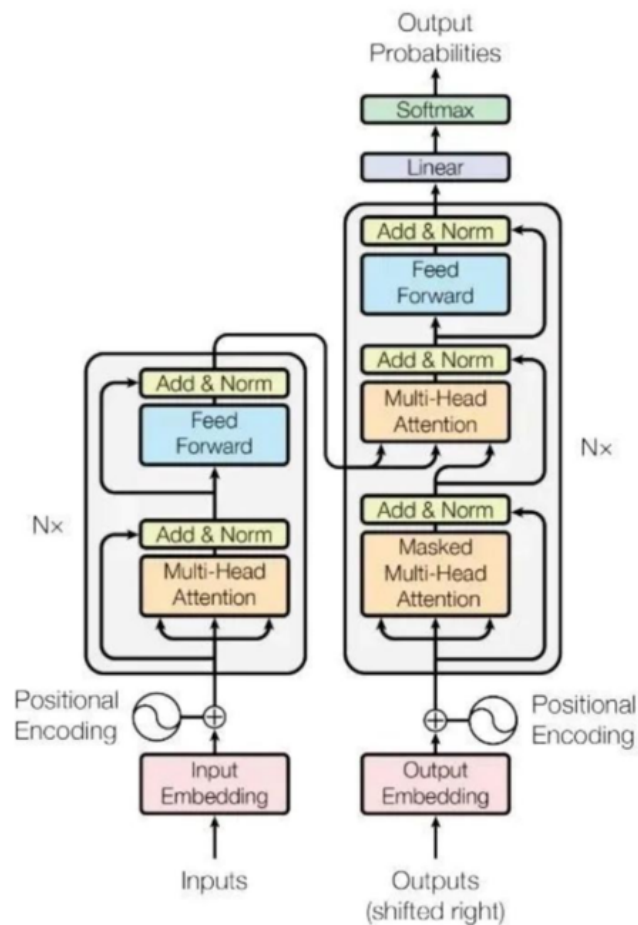


FIGURE I.12 – Architecture de Transformer. Source [11]

I.7.1 Swin Transformer

C'est un modèle qui se distingue des autres par la façon dont il divise les nouvelles fenêtres, car elles déplacent la division des fenêtres, ce qui entraîne l'émergence de nouvelles fenêtres, et ainsi une nouvelle auto-attention est calculée au sein de chaque fenêtre comme mentionné sur la figure I.13.

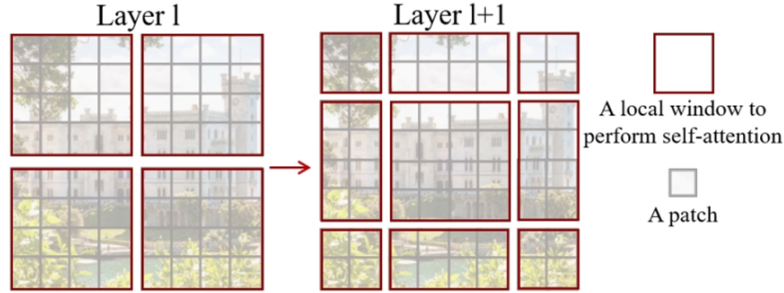


FIGURE I.13 – Illustration de l'approche à fenêtres variables. Source [12]

Un aperçu de l'architecture de Swin Transformer est présenté dans la figure 13, qui montre la petite version (SwinT). Il divise d'abord l'image RGB d'entrée en patches qui ne se chevauchent pas à l'aide d'un séparateur de patch, tel que ViT. Chaque patch est traité comme un « jeton » et sa caractéristique est mappée comme une séquence de valeurs RVB de pixels bruts. Dans notre implémentation, nous utilisons une taille de patch de 4×4 et donc la dimension de chaque patch est de $4 \times 4 \times 3 = 48$ (Il dispose de trois canaux de couleur : rouge, vert, bleu). Une couche d'intégration linéaire est appliquée à cette fonctionnalité de valeur initiale pour la projeter dans une dimension arbitraire (noté C) [12]. Plusieurs blocs de transformation avec un calcul d'auto-attention modifié (blocs de transformation Swin) sont appliqués à ces codes de débogage [12].

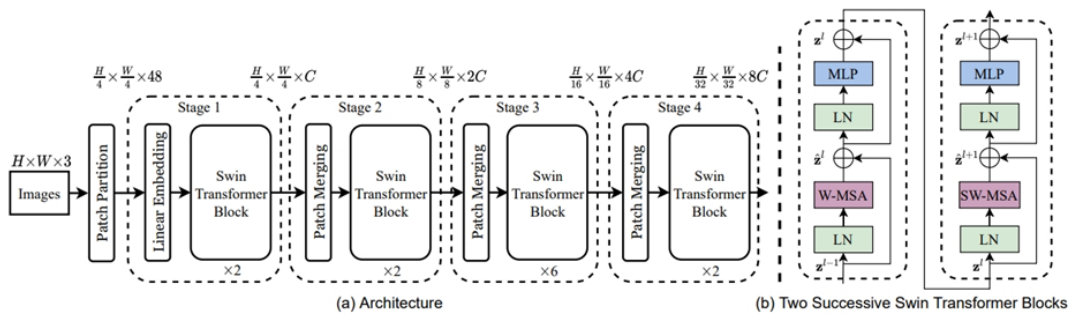


FIGURE I.14 – Architecture de Swin Transformer. Source [12]

I.7.2 Swin UNET

Swin-Unet est une architecture basée sur un commutateur qui comprend un encodeur, bottleneck, des décodeurs et des communications de mobilité (skip connections). L'encodeur, bottleneck et le décodeur sont tous connectés à la base du bloc Swin Transforme, Swin-Unet est une première architecture en forme de U pure basée sur un transformateur, C'est une architecture inspirée d'U-NET [13]. Nous présentons sur la figure I.15 l'architecture de Swin UNET.

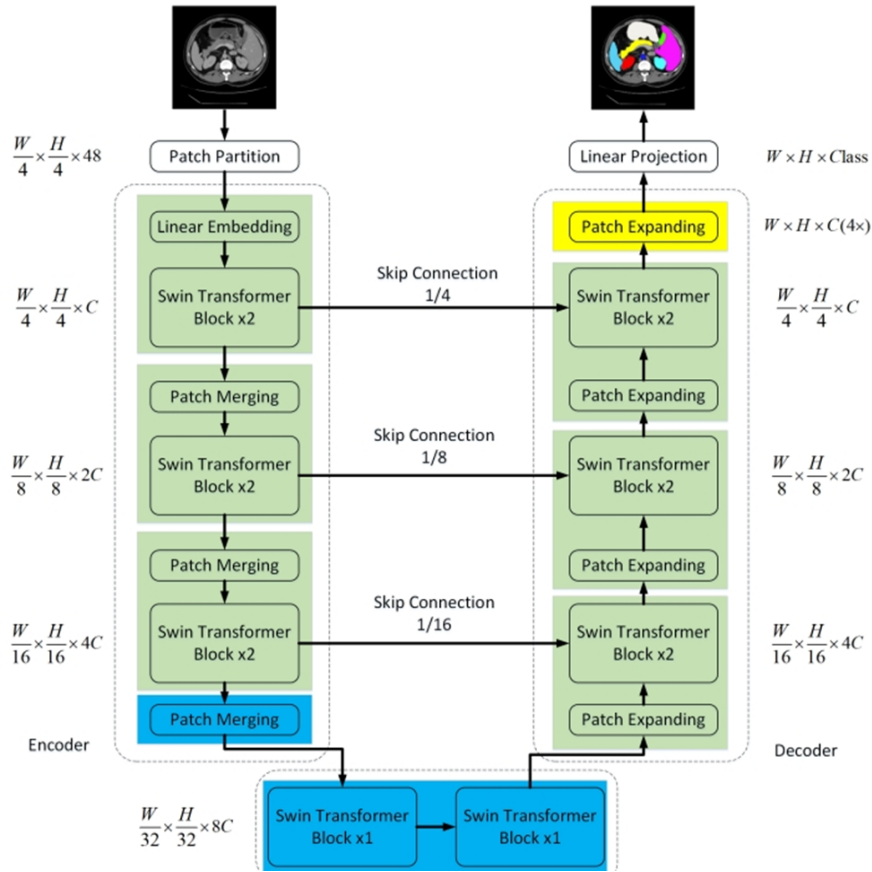


FIGURE I.15 – Architecture de Swin UNET. Source [13]

I.7.3 Swin UNETR

Swin UNETR Transformers (Swin UNETR), dont l'architecture utilise un réseau en forme de U dans lequel la technologie de Transformer Swin est utilisée comme encodeur et connectée à un décodeur basé sur CNN à différentes résolutions via des connexions sautées [14]. Nous présentons sur la figure I.16 l'architecture de Swin UNETR.

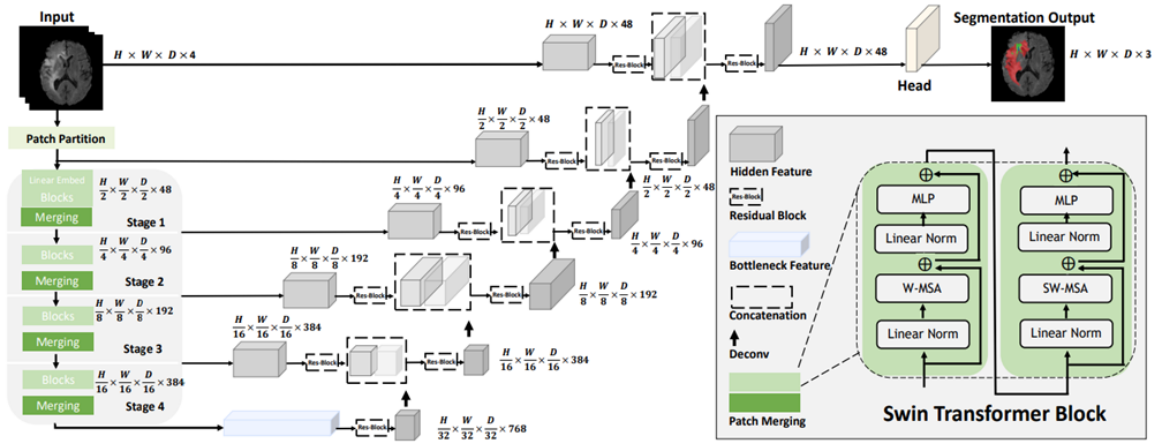


FIGURE I.16 – Architecture de Swin UNETR. Source [14]

I.7.4 TransUNet

TransUNet est le premier cadre de segmentation d'images médicales inspiré de l'architecture en forme de U [15]. Il utilise une architecture hybride CNN-Transformer pour exploiter les informations spatiales détaillées haute résolution des fonctionnalités CNN [15]. Il regroupe la fonction d'auto-attention codée par Transformers pour la combiner avec diverses fonctionnalités CNN haute résolution qui ont été ignorées du chemin de cryptage [15]. Nous présentons sur la figure I.17 l'architecture de TransUNet.

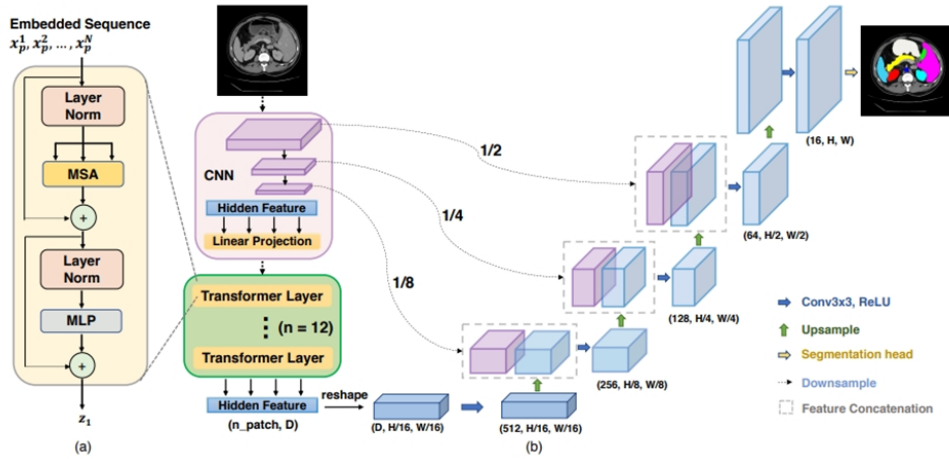


FIGURE I.17 – Architecture de TransUNet. Source [15]

I.7.5 U-Netmer : U-Net meets Transformer

U-Netmer est une nouvelle architecture de réseau neuronal simple permettant de combiner de manière optimale U-Net et le transformateur pour la segmentation. Il divise explicitement l'image d'entrée en patches locaux et utilise le transformateur pour combiner les informations

locales entre ces patches [16].

U-Netmer utilisent la technique d'auto-attention du transformateur et l'épine dorsale d'un réseau neuronal de segmentation standard (tel que U-Net) pour effectuer une segmentation sur des correctifs locaux plutôt que d'aplatir chaque correctif en jetons unidimensionnels distincts [16]. Nous présentons sur la figure I.18 l'architecture de U-Netmer.

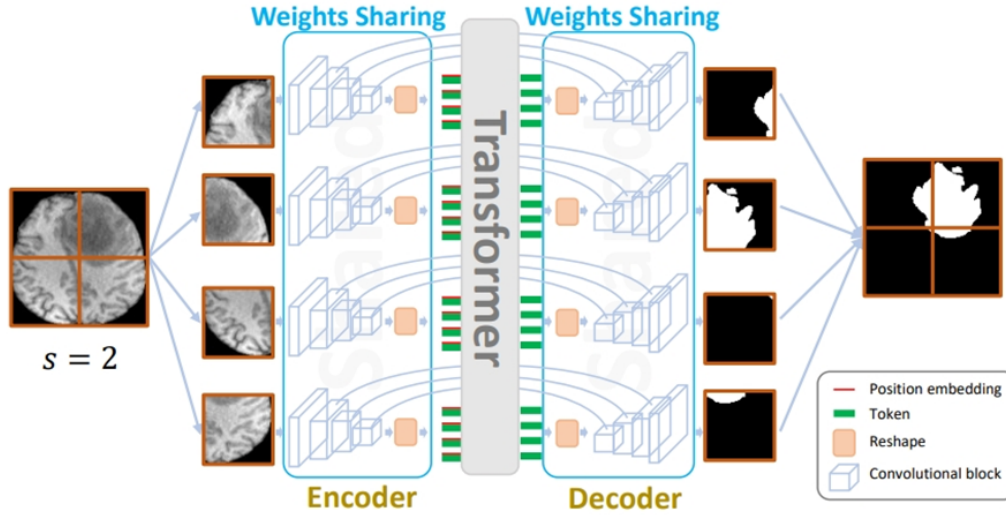


FIGURE I.18 – Architecture de U-Netmer. Source [16]

I.7.6 MMCFormer : Missing Modality Compensation Transformer

MMCFormer est un nouveau réseau pour compenser la méthode manquante. Cette stratégie est basée sur des blocs de transformer 3D efficaces et utilise une stratégie de formation commune pour former efficacement le réseau de méthodes manquantes [17]. Pour garantir la cohérence des fonctionnalités entre domaines, MMCFormer utilise des unités d'accord contextuelles global dans chaque domaine d'encodeur [17]. De plus, pour transférer des représentations spécifiques à la méthode, nous proposons d'incorporer des codes auxiliaires dans la phase de goulot d'étranglement du modèle d'interaction entre les chemins de méthode complets et manquants [17]. Nous présentons sur la figure I.19 l'architecture de MMCFormer.

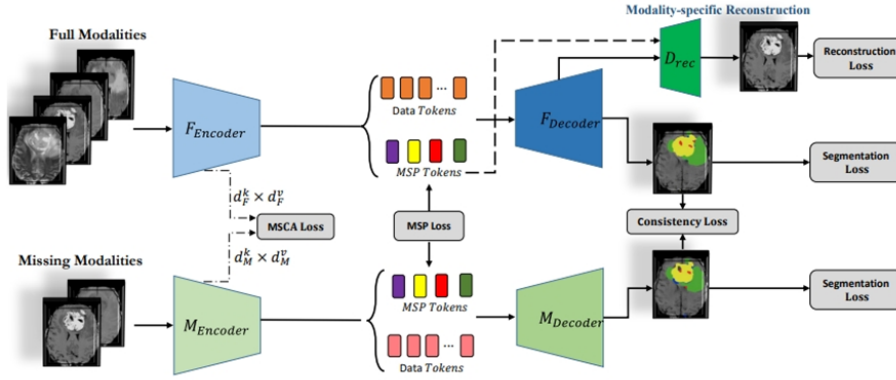


FIGURE I.19 – Architecture de MMCFormer. Source [17]

I.7.7 CoTr

CoTr est une architecture hybride entre un réseau neuronal convolutif CNN et un Transformer, un CNN est créé pour extraire les représentations de caractéristiques et un Transformer déformable efficace (DeTrans) est créé pour modéliser la dépendance à long terme des cartes de caractéristiques extraites [18]. L'avantage de DeTrans est qu'il ne prête attention qu'à un petit ensemble de postes clés en introduisant un mécanisme d'auto-attention déformable. Grâce à cela, la complexité informatique et spatiale de DeTrans est considérablement réduite, ce qui est très approprié pour la formation aux images médicales 3D car il permet de traiter des cartes de caractéristiques multi-échelles à haute résolution [18]. Nous présentons sur la figure I.20 l'architecture de CoTr.

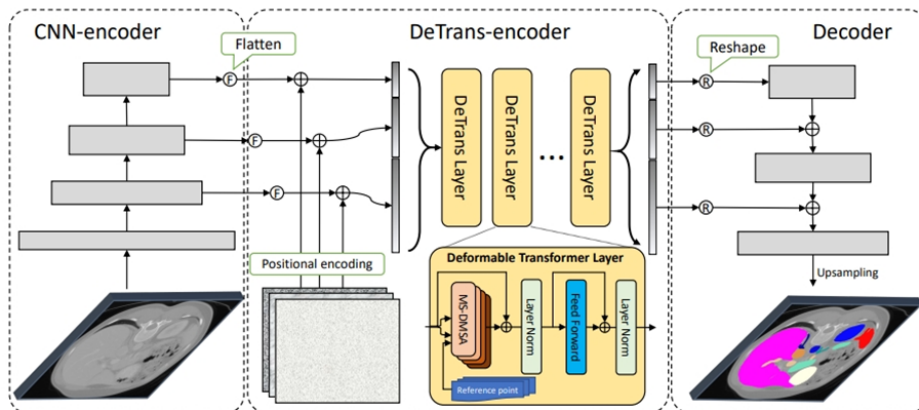


FIGURE I.20 – Architecture de CoTr. Source [18]

I.7.8 HiFormer : Hierarchical Multi-scale Representations Using Transformers for Medical Image Segmentation

Il s'agit d'une architecture hybride entre CNN et Transformer, adaptée à la segmentation fine des images médicales, exploitant efficacement les relations globales à longue portée des Transformers et les représentations de caractéristiques locales des CNN [19].

L'encodeur contient 3 modules : les modules hiérarchiques CNN et Swin Transformer et le module DLF. Chacun des modules Swin Transformer et CNN comporte trois niveaux. Tout d'abord, l'image d'entrée est introduite dans le module CNN pour apprendre sa représentation sémantique locale. Pour compenser le manque de représentation globale, un module Swin Transformer est appliqué au-dessus des fonctionnalités CNN peu profondes pour capturer les dépendances à longue portée. Ensuite, une pyramide de modules Swin Transformer avec différentes tailles de fenêtre est utilisée pour l'apprentissage par interaction multi-échelle. Pour encourager la réutilisation des fonctionnalités à vie et fournir des informations de localisation, un module de communication de saut est conçu pour transférer les fonctionnalités CNN locales vers les blocs Transformer [19].

La représentation résultante des niveaux de pyramide les plus petits et les plus grands est ensuite entrée dans le module DLF. Le nouveau module DLF proposé est un Transformer de vision multi-échelle qui fusionne deux cartes de caractéristiques obtenues à l'aide d'un mécanisme d'attention croisée [19].

Enfin, les deux cartes de caractéristiques recalibrées sont transmises au bloc décodeur pour produire le masque de segmentation final. Notre HiFormer proposé atténue non seulement le problème mentionné ci-dessus, mais surpasse également tous ses homologues en termes de différentes mesures d'évaluation. Nos principales contributions [19] :

- Une nouvelle méthode hybride qui fusionne les interactions contextuelles à longue portée du Transformer et les informations sémantiques locales de CNN [19].
- Un module DLF pour établir une fusion efficace des caractéristiques entre les représentations de caractéristiques grossières et fines [19].
- Les résultats expérimentaux démontrent l'efficacité et la supériorité du HiFormer proposé par rapport aux méthodes concurrentes sur les ensembles de données de segmentation d'images médicales [19].

Nous présentons sur la figure I.21 l'architecture de HiFormer.

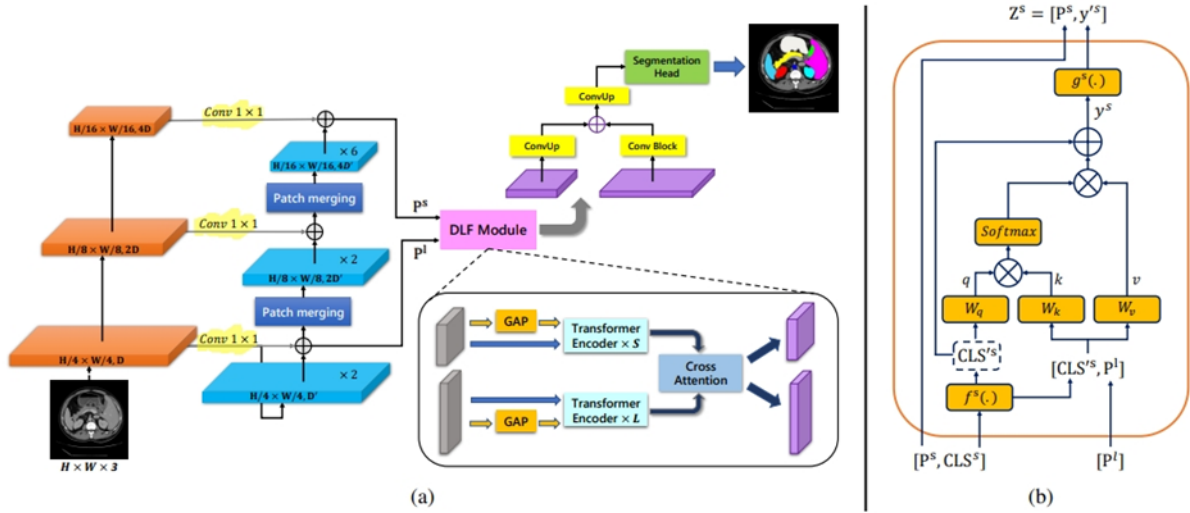


FIGURE I.21 – Architecture de HiFormer. Source [19]

I.7.9 UNetFormer

UNetFormer est une architecture hybride basée sur un transformateur, utilisant respectivement 3D Swin Transformer comme encodeur et Swin Transformer basé et CNN basé comme décodeurs. Utilisé pour former des images médicales 3D [20]. Nous présentons sur la figure I.22 l'architecture de UNetFormer.

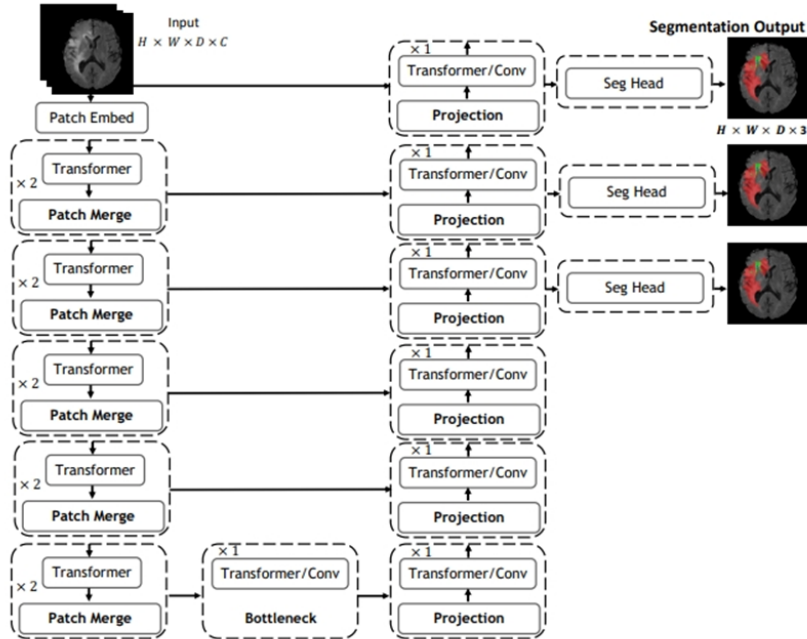


FIGURE I.22 – Architecture de UNetFormer. Source [20]

I.8 Conclusion

Les paradigmes de l'intelligence artificielle ouvrent de nouvelles frontières passionnantes dans le domaine de l'analyse d'images, offrant des opportunités de traitement multimodal et une compréhension plus approfondie du contenu visuel. Les convertisseurs enrichissent les caractéristiques des systèmes d'analyse d'images, permettant des applications plus sophistiquées telles que l'annotation d'images, la détection d'images et des objets, la segmentation sémantique et bien plus encore.

Dans ce chapitre, nous avons identifié les concepts les plus importants liés au machine learning et au deep learning du fait que nous nous intéressons aux réseaux de neurones convolutifs et aux Transformers pour le diagnostic automatique du cancer du cerveau, plus précisément des gliomes.

Dans le deuxième chapitre, nous aborderons les nouveaux développements liés au diagnostic automatique du cancer du cerveau, la détection des gliomes faisant l'objet de ces travaux.

Cependant, il est important de noter que les CNN sont encore largement utilisés et restent très efficaces pour de nombreuses tâches d'analyse d'images. L'intégration des CNN et des Transformers ouvre la voie à des systèmes plus complets et plus efficaces, exploitant les avantages de chaque architecture pour obtenir des performances optimales dans une large gamme d'applications visuelles.

CHAPITRE II

EXPLORATION DES PARADIGMES DE
L'INTELLIGENCE ARTIFICIELLE POUR LE
DIAGNOSTIC AUTOMATIQUE DU
CANCER DU CERVEAU :
PROBLÉMATIQUE ET ÉTAT DE L'ART

II.1 Introduction

Le cancer du cerveau est l'une des maladies les plus graves qui nécessite un diagnostic précis et des techniques d'imagerie exactes, et grâce aux progrès rapides dans le domaine de l'intelligence artificielle, il est possible d'améliorer la précision et la rapidité du diagnostic médical, et donc d'obtenir un plan de traitement approprié. Le développement de modèles d'intelligence artificielle (machine Learning - Deep Learning -Transformer) pour le diagnostic automatique du cancer du cerveau est un domaine de recherche prometteur et un développement important dans le domaine de la médecine qui contribue à améliorer la précision du diagnostic et la rapidité de la réponse, permettant aux patients de recevoir de meilleurs soins de santé. Dans ce chapitre, nous parlerons du cancer du cerveau et des problèmes de diagnostic qui existent.

II.2 Le cancer du cerveau

II.2.1 Définition du cancer du cerveau

La progression irrégulière des cellules dans le crâne cérébral est identifiée comme une tumeur cérébrale [54]. Le cerveau sert de centre de commande pour le corps humain ; par conséquent, les tumeurs du crâne peuvent gravement affecter la santé humaine. Selon les statistiques...

- Les tumeurs cérébrales sont responsables de 85 à 90 % de l'ensemble des tumeurs du système nerveux central [54].
- Les tumeurs cérébrales représentent plus de 85 % de toutes les tumeurs primaires du système nerveux central et environ 2 à 3 % des décès liés au cancer dans le monde [30], ce qui représente une grande menace pour la santé humaine.
- L'Organisation Mondiale de la Santé (OMS) classe les tumeurs cérébrales comme la dixième cause de décès dans le monde [55]. La maladie fait environ 97 500 morts chaque année. Le liquide céphalorachidien (LCR), la substance blanche (MW) et la matière grise (GM) sont tous interconnectés dans les tissus cérébraux sains. La croissance cellulaire incontrôlée et l'échec des mécanismes de contrôle cérébral provoquent des tumeurs cérébrales.
- Les tumeurs cérébrales malignes augmentent avec l'âge. Au cours des dernières décennies, environ 120 types différents de tumeurs cérébrales ont été identifiés [55]. En général, il existe deux grandes catégories de tumeurs cérébrales : primaires et secondaires. Un indice de gravité de 1 à 4 est attribué aux gliomes et aux méningiomes.

— Nous estimons que 251 329 personnes sont décédées des suites de tumeurs cancéreuses primaires du cerveau et du système nerveux central en 2020 [56].

Il est utile de mentionner que les tumeurs cérébrales humaines et plus particulièrement les gliomes sont parmi les plus potentiellement mortelles qui résultent généralement d'une croissance anormale des cellules souches gliales [17]. La détection précoce des tumeurs cérébrales peut améliorer fortement l'efficacité du traitement et la survie des patients atteints d'une tumeur cérébrale.

II.2.2 Types des tumeurs cérébrales

Divers types de tumeurs cérébrales existent, citons, le gliome, le schwannome, tumeur du tronc cérébral, le méningiome, l'adénome hypophysaire, les tumeurs métastatiques.

Le gliome provenant de cellules gliales est la tumeur cérébrale primitive maligne la plus fréquente chez les adultes.

C'est extrêmement invasif et constitue la tumeur maligne mortelle la plus courante dans le système nerveux, représentant 35 % des tumeurs intracrâniennes [29]. En fait, le gliome est un terme général désignant différents types de gliomes, et il en existe environ 140 types.

Les gliomes sont causés par l'interaction entre facteurs génétiques congénitaux à haut risque et des facteurs cancérigènes environnementaux [24]. La signification clinique du gliome est qu'il s'agit d'une maladie dangereuse et une tumeur cérébrale mortelle, hautement maligne et agressive, entraînant divers symptômes, tels que des convulsions, des maux de tête, des problèmes de vision et des changements dans l'élocution et le comportement. En général, les emplacements, les formes et la taille des tumeurs cérébrales ont des impacts significatifs sur le degré et la nature de ces symptômes diagnostiqués par les médecins, ainsi que la formulation du traitement et les plans de chirurgie. Suivant les critères de la classification de l'Organisation Mondiale de la Santé (OMS) (World Health Organization (WHO)), les gliomes sont classés histopathologiquement et cliniquement en grades II, III (gliome de bas grade, LGG) et IV (glioblastome, GBM). Les gliomes de différents grades ont un large éventail de pronostics.

La survie médiane des gliomes de bas grade (grades II et III) est plus élevée que 7 ans. Cependant, la survie médiane des gliomes de grade IV de l'OMS ou glioblastomes (GBM) est de 14 mois [57].

Le Glioblastome (GBM), les métastases cérébrales isolées (SBM) et le lymphome primitif du système nerveux central (PCNSL) possèdent un haut niveau de similarité en termes d'histomorphologie et de manifestations cliniques en IRM multimodale. Ces similitudes ont conduit à

des défis dans le diagnostic clinique de ces trois tumeurs malignes [21]. En effet, ces tumeurs cérébrales malignes sont similaires et complexes dans leur organisation biologique et leurs caractéristiques cliniques typiques, comme mentionné sur la figure, posant de grands défis aux cliniciens dans le diagnostic et le traitement de ces tumeurs. Pour de nombreux patients, une procédure de biopsie invasive est nécessaire en obtenant un échantillon de tissu pour déterminer le type de tumeur. Cependant, les biopsies peuvent entraîner une hémorragie tumorale/des tissus cérébraux, infection, troubles neurologiques, convulsions et même décès.

Avec le développement d'une technologie d'imagerie médicale moderne, multimodale l'imagerie par résonance magnétique est devenue l'outil le plus couramment utilisé pour le diagnostic et le traitement des tumeurs cérébrales en raison de la richesse de ses informations sur le contraste des tissus, de l'absence de rayonnements ionisants pendant l'imagerie, et sa faible susceptibilité aux erreurs d'échantillonnage.

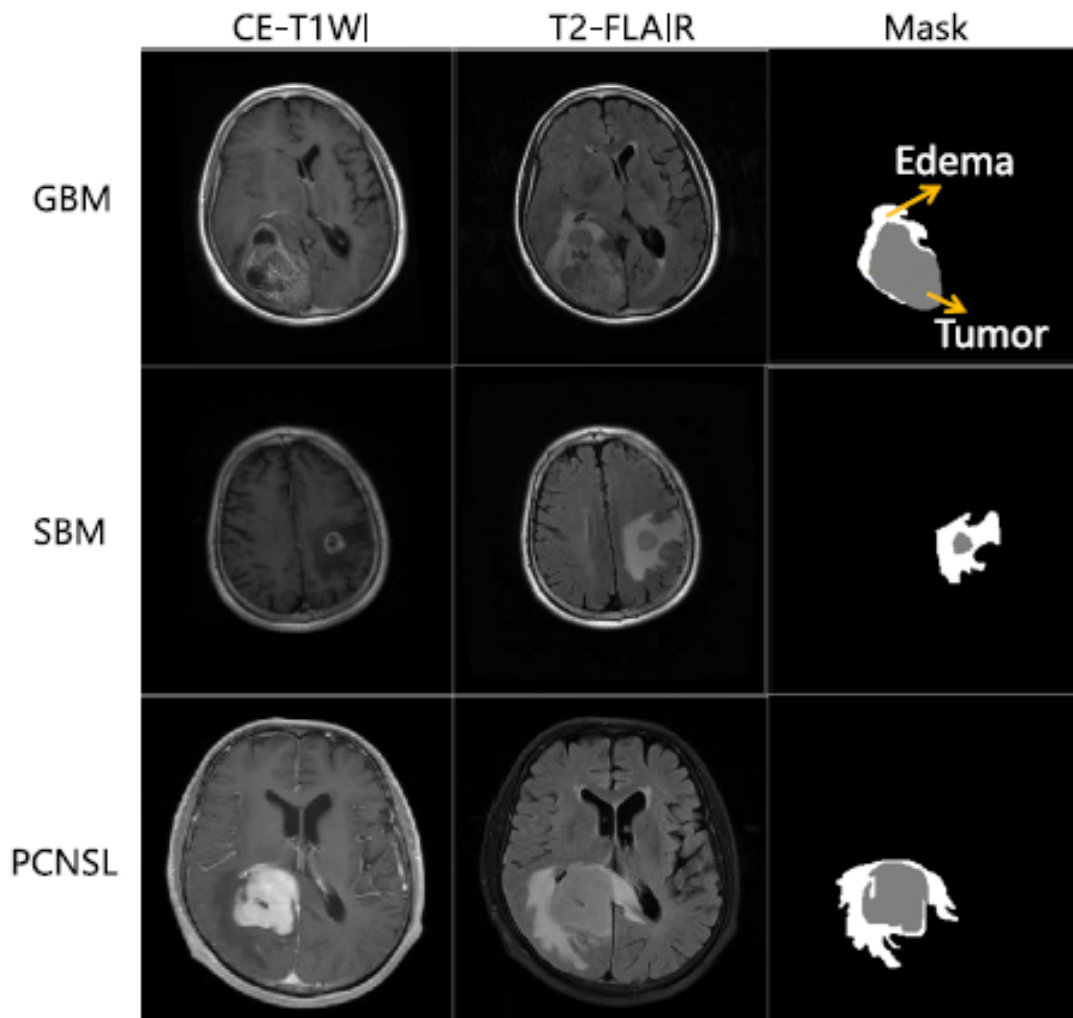


FIGURE II.1 – Exemples de tumeurs cérébrale. Source [21]

Dans ce travail, nous nous intéressons au LGG autrement appelé bénin et au HGG autrement

appelé malin. Dans le cas de LGG comme illustré sur la figure II.2 (a), le taux d'agressivité des tissus tumoraux dans les zones touchées est faible et il est élevé dans le cas de HGG, comme illustré sur la figure II.2 (b). Par conséquent, le HGG est plus important que le LGG [22].

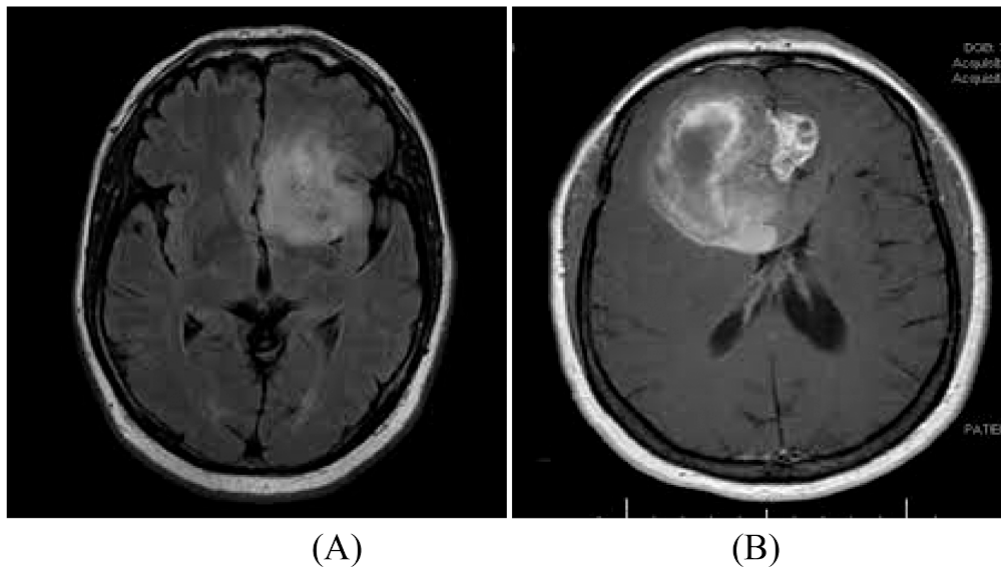


FIGURE II.2 – Tumeur LGG en A et HGG en B. Source [22]

II.2.3 Dépistage du cancer du cerveau basé imagerie par résonance magnétique (IRM)

Il existe diverses méthodes utilisées par les oncologues pour évaluer l'existence de tumeurs cérébrales, telles que des analyses de sang ou des évaluations visuelles. De plus, la technique non invasive d'imagerie par résonance magnétique (IRM) sans rayonnement ionisant a été couramment utilisée pour le diagnostic [54]. En effet, les radiologues ont utilisé de nombreuses techniques d'imagerie, notamment la tomodensitométrie, pour détecter les tumeurs cérébrales. De plus, les IRM présentent une imagerie avec une résolution plus élevée que les tomodensitogrammes (CT) [54]. En raison de la nature astronomique de l'IRM, elle est largement utilisée pour détecter les tumeurs cérébrales malignes.

Nous distinguons d'autres imageries médicales pour le diagnostic précoce des patients comprenant l'imagerie échographique et l'imagerie à rayons X [29]. La technologie d'imagerie Ultrason peut être utilisée pour juger de la position et forme des organes du corps et aide les médecins à localiser les lésions avec précision. La technologie d'imagerie CT peut offrir aux médecins davantage d'informations détaillées sur les lésions du patient et est souvent utilisée dans le diagnostic des maladies du système nerveux central et des maladies pulmonaires. La technologie d'imagerie à rayons X (X-ray) est principalement utilisée dans l'inspection des fractures et

l'analyse de la densité osseuse. Cependant, les examens de la CT et des rayons X manquent de bon contraste pour les tissus mous tels que les tumeurs et les vaisseaux sanguins, et les rayons X générés pendant l'acquisition d'images, provoqueront certains dommages causés par les rayonnements ionisants au corps humain.

Avec le développement et la mise à niveau continus de l'équipement d'imagerie médicale, la technologie d'imagerie IRM a progressivement été largement utilisée. L'IRM est non invasive et peut fournir une image détaillée sur la forme, la taille et l'emplacement sans que le corps ne reçoive de niveaux élevés de rayonnements ionisants [29]. Le principe de base de la technologie IRM consiste à placer le corps humain sous un champ magnétique très puissant, et l'appareil envoie des impulsions de radiofréquence pour stimuler les atomes d'hydrogène dans les cellules humaines dans le but de générer des effets de résonance magnétique. Enfin, l'ordinateur traite et reconstruit davantage les signaux électromagnétiques collectés pour générer l'image d'IRM. Les tumeurs cérébrales sont riches en nouveaux vaisseaux sanguins, de plus, l'IRM présente un contraste plus élevé avec les tissus mous, ce qui peut montrer clairement les tumeurs cérébrales et leur relation anatomique avec les tissus environnants.

De plus, la technologie IRM peut exploiter diverses séquences d'imagerie pour discerner distinctes caractéristiques de tissus au sein de la région tumorale, ce qui entraîne l'acquisition d'images multimodales. Cette approche permet au praticien médical d'avoir une compréhension plus holistique des attributs de lésions.

A cet effet, l'imagerie par résonance magnétique (IRM) est une technique couramment utilisée de diagnostic non invasive en pratique clinique [23]. Elle offre plusieurs modalités d'imagerie. Nous distinguons : pondérée en T1 (T1), T1 avec contraste améliorant le gadolinium (T1c), pondérée en T2 (T2) et récupération d'inversion atténuée par les fluides (FLAIR). Les images pondérées en T1 ont la capacité d'identifier avec davantage de précision la position et le contour du noyau tumoral, quant aux images pondérées en T2, elles ont la plus grande valeur pour localiser l'œdème. Les images T1c peuvent détecter avec plus de précision s'il existe une tumeur renforcée au cœur de la tumeur, et les images FLAIR peuvent délimiter plus clairement le contour général de la tumeur.

Comme mentionné sur les figures, les tumeurs cérébrales se composent de trois régions qui se chevauchent : la région de l'œdème (ED), la région tumorale rehaussée (ET), la région centrale nécrotique et la région tumorale non rehaussée (NCR/NET). Chaque région reflétant différentes caractéristiques biologiques. Ces trois sous-régions constituent les régions d'intérêt (ROI) à segmenter, à savoir la tumeur entière (WT), le noyau tumoral (TC) et l'ET. Mentionnons que WT représente l'union des trois sous-régions, tandis que TC comprend NCR/NET et ET.

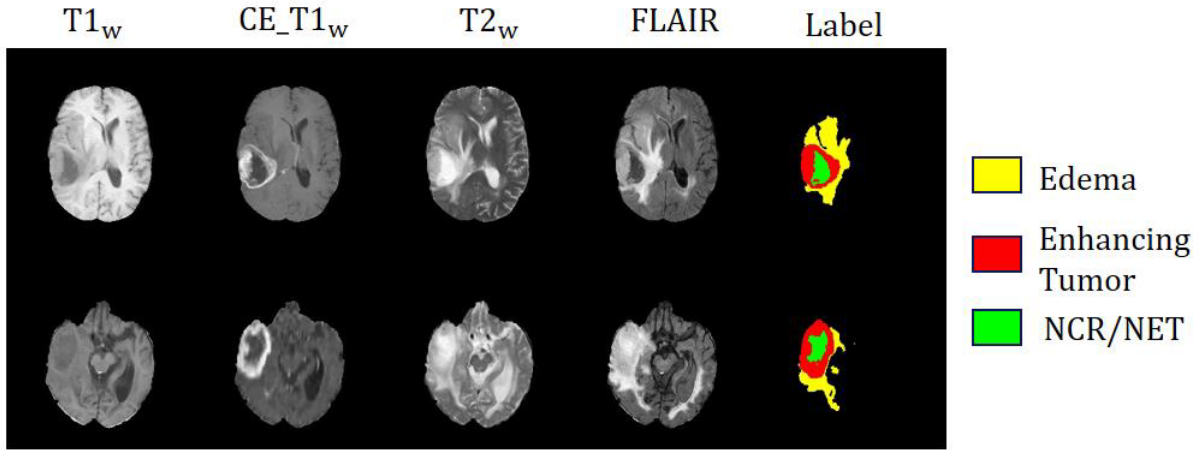


FIGURE II.3 – Exemples d'IRM multimodale de tumeur cérébrale. Source [23]

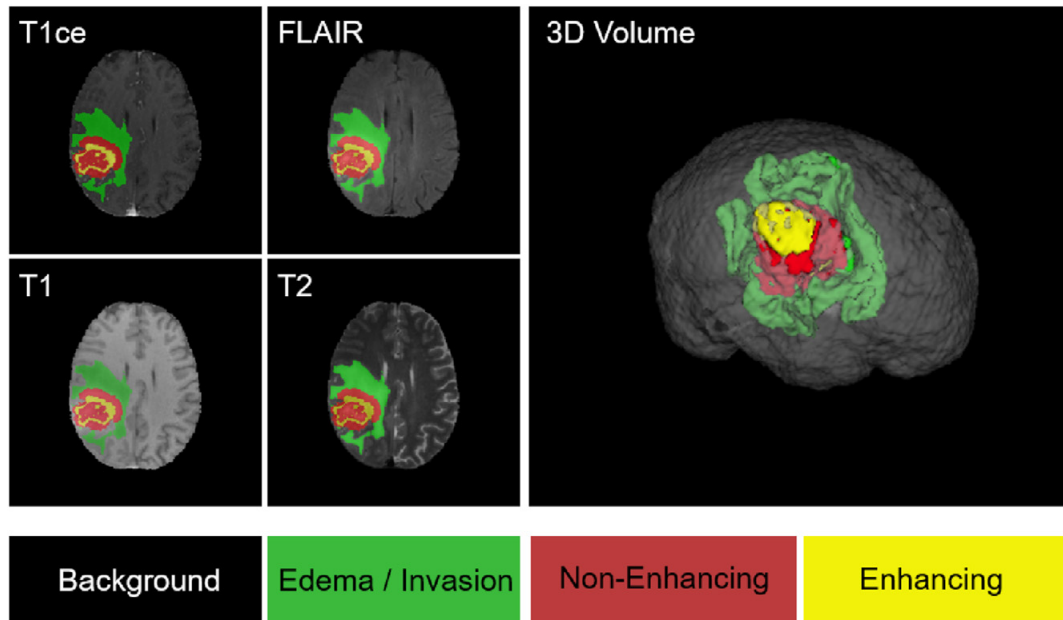


FIGURE II.4 – Illustrations of different sub-regions for brain tumor segmentation. Source [24]

II.3 Diagnostic automatique des tumeurs cérébrales

Les cliniciens diagnostiquent principalement en lisant les images IRM et en segmentant les régions tumorales à partir d'images du cerveau entier. Cependant, lire une IRM avec un certain nombre de tranches prend du temps et est laborieux [24]. Les méthodes de segmentation automatique basées sur l'apprentissage profond peuvent segmenter l'intégralité de la tumeur et les diagnostiqueurs peuvent considérablement réduire le temps nécessaire à l'analyse. Cependant, la plupart des méthodes basées sur l'apprentissage nécessitent un certain nombre d'annotations manuelles comme supervision, ce qui coûte quand même cher. Dans ce sens, des approches

de diagnostic automatique (computer aided diagnosis : CAD) ont été développés dans le but d'aider les neuro-oncologues durant le processus de reconnaissance des tumeurs cérébrales. En effet, les radiologues obtiendront un deuxième avis pour qu'ils puissent déterminer l'intensité, le diamètre ainsi que l'emplacement et le type des tumeurs et en aucun cas ces systèmes désigneront un substitut à l'expertise humaine.

Nous exposons sur la figure, les principales étapes des outils de système de diagnostic automatique du cancer du cerveau : le prétraitement, la segmentation de la région d'intérêt, l'extraction des descripteurs et la classification.

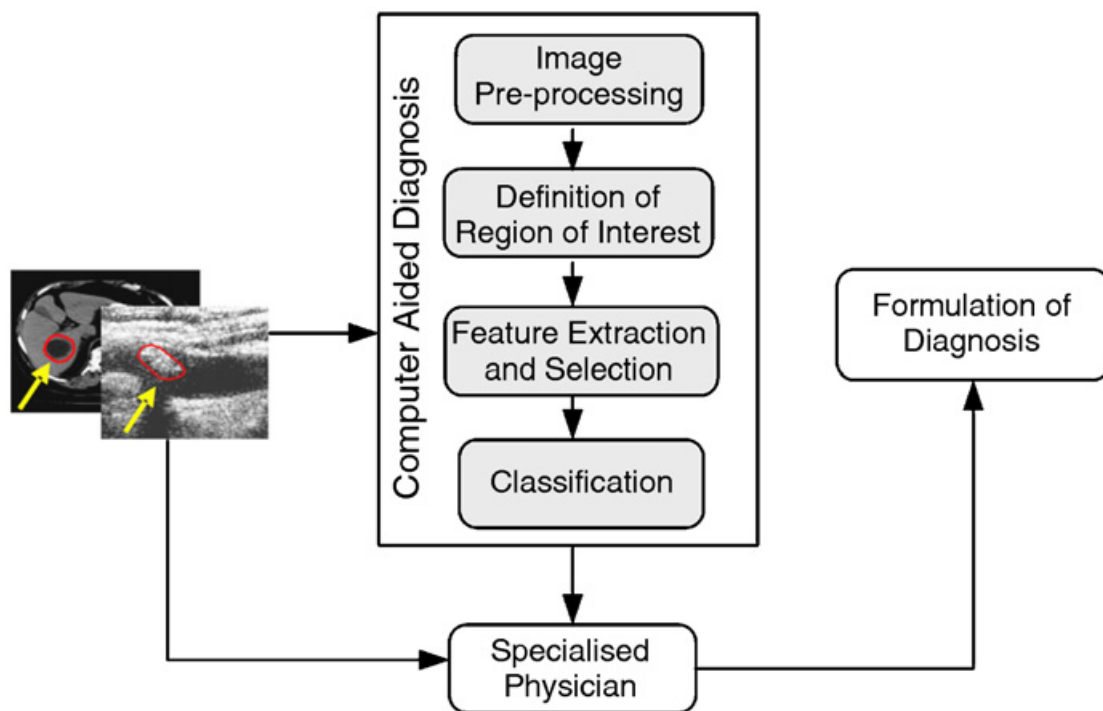


FIGURE II.5 – Fonctionnalités d'un système de pronostic automatique du cancer du cerveau.
Source [25]

La segmentation des tumeurs cérébrales est donc propice à une localisation et une reconnaissance précise et efficace des gliomes, ce qui peut aider les médecins à améliorer le diagnostic et le pronostic au sein des applications cliniques [24].

Au cours des dernières décennies, les chercheurs ont mené de vastes recherches fondamentales sur les tumeurs cérébrales. Les premières recherches visent à comprendre les propriétés biologiques des cellules gliales et comment elles se transforment en cellules malignes de tumeurs malignes. Au fil du temps, les chercheurs acquièrent une meilleure compréhension des changements génétiques et moléculaires qui se produisent dans les gliomes. Ces recherches ont favorisé le développement de nouvelles méthodes de diagnostic et de traitement pour les tumeurs cé-

rébrales, comme la détermination de la classe de la tumeur cérébrale, de l'hérédité et thérapie ciblée grâce aux données génomiques.

Plus précisément, la segmentation des tumeurs cérébrales vise à localiser plusieurs régions tumorales en modifiant les représentations de l'IRM, où les résultats de la segmentation des tumeurs cérébrales seront utilisés pour le pronostic ultérieur de la nature de la tumeur et la prédiction de la survie.

Concrètement, une segmentation précise de la tumeur cérébrale en imagerie par résonance magnétique (IRM) multimodale est essentielle et importante puisque la plupart des tissus ne sont pas résécables lors de la chirurgie des tumeurs cérébrales [24].

Au cours des dernières années, le développement rapide du deep learning a amélioré efficacement les performances du diagnostic assisté par ordinateur.

De nombreuses avancées techniques ont été réalisées dans le domaine de la segmentation tumorale multimodale du cerveau, conduisant à un nombre croissant de méthodes capable d'accomplir cette tâche avec différents degrés de précision et de rapidité. Les méthodes les plus anciennes et les plus simples de segmentation des tumeurs cérébrales se basaient sur le traçage manuel, où un clinicien expert dessine un contour autour de la tumeur dans les images [24]. Cependant, le traçage manuel prend du temps et peut être sujet à une variabilité inter et intra-observateur. Avec l'avènement des techniques de vision par ordinateur et d'apprentissage automatique, nombreuses méthodes automatisées ont été développées pour effectuer la segmentation des tumeurs cérébrales. Ces méthodes peuvent être globalement classées en deux catégories : les méthodes traditionnelles et les méthodes d'apprentissage profond.

Les méthodes traditionnelles incluent la croissance régionale, la méthode de définition de niveaux et la segmentation basée sur un atlas [24]. Ces méthodes utilisent les descripteurs de l'image tels que les informations sur l'intensité, la texture et le dégradé pour séparer les tumeurs cérébrales des tissus environnants. Prenons l'exemple de la croissance régionale où les méthodes utilisent des points de départ pour développer la région tumorale en incluant des voxels qui ont des valeurs d'intensité similaires à celles des points de départ. Durant ces dernières années, les méthodes d'apprentissage, notamment les réseaux de neurones convolutifs (CNN) sont devenus majoritairement les plus utilisées pour la segmentation des tumeurs cérébrales. En effet, les réseaux de neurones convolutifs peuvent apprendre des caractéristiques complexes d'image à partir d'une grande masse de données annotées, ce qui conduit à des performances améliorées par rapport aux méthodes traditionnelles.

Récemment, les transformers témoignent de progrès remarquables et aboutissent à des performances supérieures en matière de segmentation des tumeurs cérébrales que les approches

précédentes où l'ultime but consiste à capturer des descripteurs globaux plus discriminants en obtenant des résultats impressionnants.

II.4 Problématique tirée dans le pronostic automatique du cancer du cerveau

Pour la segmentation des gliomes, l'IRM est couramment utilisée. Même si l'imagerie IRM est non invasive, elle peut être longue, difficile à interpréter et parfois de mauvaise qualité. En plus d'évaluer l'emplacement, la morphologie et la taille de la tumeur à l'aide de méthodes évaluant la différence d'intensité entre les tissus tumoraux voisins, la position, la forme et la taille de la tumeur peuvent également être déterminées. En pratique clinique, une méthode de catégorisation des tumeurs très précise et totalement automatisée est nécessaire [55]. Dans ce sens, nous allons souligner quelques points problématiques :

1. Les modèles basés sur les réseaux de neurones convolutifs (CNN) ont du mal à capturer des descripteurs sémantiques de segmentation qui peuvent caractériser les dépendances globales. Bien que des méthodes basées sur des transformers aient été développées pour la segmentation des tumeurs, le coût de calcul élevé a conduit à des difficultés dans l'application de ces modèles en clinique pratique. De plus, les architectures de réseau spécifiques à des tâches précises nécessitent souvent une conception manuelle minutieuse basée sur une expérience et une expertise substantielle [21].
2. Avec le développement continu de la technologie du deep learning, c'est devenu une tendance de recherche d'utiliser l'algorithme de CNN pour résoudre de nombreux problèmes dans les tâches de segmentation d'images médicales. Cependant, le contraste entre les images IRM de tumeurs cérébrales est faible, ce qui conduit à une segmentation inexacte des images de tumeurs cérébrales par l'algorithme de CNN [29].
3. Une méthode automatisée de segmentation des tumeurs cérébrales par IRM, rapide et précise, est essentielle au diagnostic clinique et au traitement du cancer du cerveau. La méthode de segmentation basée sur un réseau de neurones convolutifs a atteint des performances impressionnantes dans la segmentation d'images médicales avec de puissantes capacités de représentation locale. Néanmoins, ils présentent des limites dans la modélisation des interactions contextuelles globales ou à longue portée et des dépendances spatiales, qui sont essentielles à la segmentation des images médicales [30].

4. Pour l'imagerie médicale, un réseau CNN 2D ne permet pas de réaliser pleinement l'utilisation des informations de localisation spatiale de données d'images médicales 3D [29].
5. La majorité des travaux reposaient uniquement sur des images IRM à contraste amélioré T1 pour la classification des tumeurs, cependant aucune garantie sur le fait que le modèle peut résoudre les classifications appropriées pour différentes classes de tumeurs [55].
6. Les modèles proposés à ce jour basés sur DL présentent des problèmes [55] tels qu'une faible précision de classification, une complexité de calcul élevée, un nombre limité de classes étudiées, un apprentissage lent, un overfitting, une faible robustesse, un manque de convergence, la nécessité d'affiner et de définir plusieurs paramètres et difficulté d'évaluation. En outre, la question de l'incertitude et de la non-généralisation des méthodes résultante d'un training avec un ensemble limité d'images IRM a ajouté aux défis et aux limites des recherches antérieures.

II.5 Etat de l'art sur l'aide au diagnostic médical du cancer du cerveau

Pour relever les défis mentionnés ci-dessus, le développement d'un système automatique de reconnaissance des tumeurs cérébrales gagne en popularité. De nombreux chercheurs ont développé divers systèmes basés à la fois sur des approches traditionnelles d'apprentissage automatique et sur des approches d'apprentissage profond. Nous présentons sur le tableau suivant quelques approches de pronostic automatique du cancer du cerveau.

Tableau II.1 Approches du pronostic automatique du cancer du cerveau

Réf	Type	Méthode	Dataset	Résultats
[58]	Classification	LBP + KNN α LBP + KNN	Nanfang Hospital, Guangzhou, China	93.28% 90.57%
[59]	Segmentation + Classification	Segmentation basée Adaptive Regularized Kernel based Fuzzy C-Means Clustering et une combinaison des SVM et ANN pour la classification	Une base de donnée locale	91.4%
[28]	Segmentation + Classification	Des stratégies de décomposition adaptative : en mode empirique bidimensionnel (BEMD) et la décomposition en mode variationnel quasi-bivarié modifié (MQBVMD) pour l'obtention des sous-images significatives ensuite une classification basée les descripteurs de ELDP et les SVM est appliquée pour la distinction entre LGG et HGG	BRATS 2015	94.44%
[60]	Segmentation + Classification	DWT-PCA + SVM	BRATS	82%
[61]	Segmentation + Classification	Segmentation des tumeurs basée Bayesian fuzzy clustering approach qui seront classifiées par Optimization algorithm-based multi-SVNN par l'utilisation des descripteurs : the information theoretic measures, scattering transform et les Ondelettes	BRATS	93%

[55]	Classification	Un modèle parallèle à deux branches : le module de Transformer avec l'unité d'auto-attention et les CNN est proposé pour la classification des tumeur cérébrales où des descripteurs locaux et globaux dérivées de CNN et TM ont été combinées. iVGG iDensNet	Figshare	98.59% 98.94%
[62]	Classification	Une approche effectuant l'extraction des descripteurs basée siamese neural network (SNN) ensuite une classification des tumeurs cérébrales est réalisée par un K-NN	Figshare	92.6%
[63]	Classification	Une extraction et sélection des descripteurs à l'aide de l'analyse statistique sont effectuées pour la classification des tumeurs cérébrales en LGG et HGG à travers les techniques SVM et K-nn.	BRATS	94.6%
[64]	Classification	Le réseau unifié de visualisation et de classification des gliomes (UniVis-Net) a été proposé pour remédier au désalignement de l'attention en introduisant une approche de sous-régions basée sur un mécanisme d'attention. De plus, des cartes de caractéristiques multi-échelle sont fusionnées pour obtenir une résolution plus élevée, permettant la génération d'explications visuelles détaillées.	BRATS 2019	89.3%

[21]	Segmentation + Classification	Un transformer d'apprentissage multitâche avec recherche d'architecture neuronale (NAS) a été proposé pour la segmentation et la classification des tumeurs cérébrales (BTSC-TNAS). Pour la segmentation, un transformer en forme de U (NTU-NAS) avec NAS a été développé pour prédire les tumeurs cérébrales. Quant à la classification, les descripteurs multi-échelles obtenues à partir de l'encodeur de NTU-NAS ont été utilisés comme entrées du réseau de classification (MSC-NET).	BRATS 2019	94.1%
------	-------------------------------	--	------------	-------

TABLE II.1 – Approches du pronostic automatique du cancer du cerveau.

II.6 Conclusion

L'exploration des modèles d'intelligence artificielle pour le diagnostic automatique du cancer du cerveau représente un développement important dans les domaines médical et scientifique, et avec la compétition continue entre les chercheurs pour développer le meilleur modèle, nous pourrions atteindre le stade de l'application de ces modèles dans la vie quotidienne pour faciliter le travail des médecins et obtenir des résultats précis et rapides, établissant ainsi un diagnostic médical précis. Nous examinerons l'un de ces modèles dans le chapitre suivant.

CHAPITRE III

APPROCHE PROPOSÉE POUR LA
SEGMENTATION 3D DU GLIOME ET LE
PRONOSTIC AUTOMATIQUE EN LGG OU
HGG

III.1 Introduction :

Dans ce chapitre, nous aborderons l'approche proposée pour la problématique mentionnée au chapitre précédent, qui consiste à appliquer une approche de diagnostic automatique du cancer du cerveau basée sur les branches de l'intelligence artificielle que nous avons évoquées précédemment, qui sont : L'apprentissage automatique (machine learning) et l'apprentissage profond (Deep learning) ainsi que les Transformers.

III.2 Plan d'action

Suite aux points problématiques soulignés au chapitre précédent, nous allons présenter notre plan d'actions comme suit :

1. La classification des tumeurs cérébrales à l'aide d'une analyse pathologique nécessite des biopsies, ce qui en fait une méthode invasive et fastidieuse qui peut être angoissante en raison d'un échantillonnage erroné et de la variabilité inter-observateur. Par conséquent, la segmentation automatique des tumeurs générées à partir d'images IRM pourrait être extrêmement utile pour les applications médicales [65]. De plus, le niveau d'intensité des images varie selon les patients. Par conséquent, une étape préalable au traitement des images IRM est nécessaire pour les données d'entrée de la méthode de segmentation et de notation.
2. Les données IRM ont été obtenues à partir de plusieurs scanners. Par conséquent, le pré-traitement de l'image est très important pour extraire des descripteurs stables et obtenir des résultats reproductibles. Toutes les images ont été prétraitées pour standardiser l'intensité et variations géométriques des données IRM en se basant sur le filtre Gaussien et l'algorithme du skull stripping.
3. Plusieurs études ont amélioré la performance des tâches grâce à la fusion des modalités car elle permet à plusieurs modalités de se compléter [57]. Cependant, les modèles CNN doivent concevoir un certain nombre de connexions pour y parvenir. Comparé à CNN, Transformer est plus adapté pour gérer la multimodalité. D'un autre côté, les modèles de segmentation basés sur CNN ne peuvent pas tirer pleinement parti du contexte informations pour la capture de descripteurs. Tandis que l'avantage du Transformer est l'excellente capacité à capturer la perception globale et la modélisation de la dépendance à long terme par l'approche de l'attention, qui permet au Transformer d'obtenir de meilleurs résultats dans la tâche de la segmentation. En effet, nous obtenons de meilleurs résultats

grâce à une conception de fusion de modalités à plusieurs niveaux. Nous appliquons le Transformer pour chacun de modalités afin que les caractéristiques optimales extraites puissent être fusionnées pour former un meilleur ensemble de descripteurs afin de garantir l'exactitude des tâches de segmentation.

4. Notre méthode peut exploiter pleinement les caractéristiques de la multimodalité, améliorant ainsi les performances de segmentation dans une certaine mesure. Nous traitons les quatre modalités d'images IRM comme quatre canaux du fait que chaque modalité se concentre sur différentes informations. Par exemple, à l'IRM, l'œdème produit un fort signal en imagerie pondérée T2 et un signal faible en imagerie pondérée T1. De plus, chaque modalité affiche différentes textures et autres caractéristiques dues à différents protocoles d'imagerie. La modalité FLAIR est très sensible aux tumeurs et fait la différence entre le liquide céphalo-rachidien (LCR) et la tumeur [56]. Les images T1ce sont produites en améliorant le contraste des images T1 à l'aide d'un agent d'amélioration du contraste. Avec ces différences d'imagerie et diverses caractéristiques et informations distinctives affichées dans chaque modalité, les images des modalités T2 et FLAIR sont les plus adaptées à la détection des tumeurs d'œdème péritumoral. En revanche, les images de modalités T1 et T1ce sont les mieux adaptées pour classer le noyau tumoral. Par conséquent, il est évident que chaque modalité capture des caractéristiques distinctes de l'anatomie cérébrale et de la tumeur. Les données de modalité complètes fournissent des données riches et omniprésentes, cruciales pour les performances de segmentation des modèles utilisés pour ces tâches.
5. Avec le développement de la technologie IRM, les images IRM multimodales ont été largement appliqués à la segmentation des tumeurs cérébrales, ce qui peut fournir une compréhension plus complète des tumeurs et de leurs tissus environnants cérébraux [24]. Concrètement, l'IRM contient quatre modalités en termes de T1, T2, T1ce et FLAIR, qui sont des modalités d'imagerie complémentaires pour le diagnostic et le suivi des tumeurs cérébrales, c'est-à-dire les différentes modalités d'IRM peuvent fournir des informations complémentaires sur l'aspect et les propriétés de la tumeur. Par exemple, la modalité IRM de la pondération T1 est souvent utilisée pour fournir des images haute résolution du cerveau, et la modalité IRM pondérée en T2 peut fournir des informations sur la teneur en fluide des tissus, particulièrement utile pour différencier entre la tumeur et le tissu cérébral normal. De plus, la modalité d'IRM pondérée en T1 peut fournir des informations sur les vaisseaux sanguins et le modèle de rehaussement de la tumeur, qui

peut être utilisé pour aider à diagnostiquer le type de tumeur et son agressivité. La combinaison de plusieurs modalités d'IRM peut fournir une analyse plus complète et plus précise de la tumeur et du tissu cérébral environnant, qui est essentielle pour une segmentation efficace des gliomes.

6. La plupart des recherches existantes séparent la tâche de segmentation des tumeurs cérébrales et la tâche de la classification, où chaque tâche peut impliquer la conception de certains réseau coûteux en termes de calcul qui entrave son application en clinique pratique. Nous proposons un modèle qui atteint conjointement l'objectif de segmentation des lésions tumorales ainsi que leur classification. Les Transformers peuvent capturer les dépendances de caractéristiques à longue portée mais ignorent les détails des caractéristiques locales, alors que les opérations convolutives sont capable de capturer les caractéristiques locales [21].
7. U-Net et ses variantes excellent dans la fusion des descripteurs à partir de plusieurs niveaux et démontrent des performances de segmentation avancées avec une architecture en forme de U très simple. Cependant, ils ont souvent du mal à capturer le contexte global des images de tumeurs cérébrales et traiter également toutes les zones d'une image. Pour répondre à ces problématiques, les chercheurs ont intégré des modules d'attention dans l'architecture U-Net pour capturer des informations plus informatives au contexte de la segmentation cérébrale [66]. Les modules d'attention intégrée permettent au modèle de se concentrer sur des régions spécifiques et de rassembler les informations à partir des dimensions spatiales/canaux, améliorant ainsi les performances d'extractions de descripteurs locaux dans les zones de lésions. Cependant, ces modèles spatiales/canaux liés à l'attention ont des limites dans la capture des contextes globaux des images de tumeurs cérébrales pour la segmentation. Plus récemment, motivé par le succès de Vision Transformer (ViT), les chercheurs introduisent davantage les modules Transformer dans l'architecture U-Net. En capturant les relations à long terme des images de la tumeur cérébrale via l'opérateur d'auto-attention, U-Net fusionné avec un transformer démontrent des résultats de segmentation plus prometteurs sur ce domaine médical. Cependant, ces modèles améliorés intègrent principalement les structure du Transformer dans des positions fixes, mais négligent l'importance de structures hiérarchiques dans ces tâches. De plus, ils sont également confrontés aux défis liés à la diversité des caractéristiques à mesure que la profondeur du modèle augmente.
8. Les réseaux de segmentation traditionnels basés sur UNet ont du mal à établir des dé-

pendances explicites à longue portée à partir de l'espace des descripteurs en raison des limitations du champ de réception CNN. Ceci est particulièrement crucial pour les tâches de prédiction denses telles que la segmentation des tumeurs cérébrales. Des travaux récents ont intégré la puissante capacité de modélisation globale du Transformer dans UNet pour obtenir des résultats de segmentation plus précis. En effet, bien que les méthodes basées sur les CNN démontrent d'excellentes performances en segmentation, le champ réceptif limité des opérations convolutives entrave la capacité du modèle à capturer des informations contextuelles globales, cruciales pour une segmentation précise des tumeurs cérébrales puisque la taille du champ récepteur détermine directement la capacité du modèle à « voir » plus loin [23]. En revanche, la taille et la forme fixes des noyaux convolutifs empêchent l'adaptabilité à différents types d'images d'entrée, limitant l'applicabilité de la convolution. D'autre part, les mécanismes d'attention sont largement utilisés dans diverses tâches de prédiction en améliorent les caractéristiques importantes tout en supprimant ceux qui ne sont pas saillantes. Transformer a été utilisé pour capturer les corrélations entre deux séquences d'entrée quelconques grâce au mécanisme d'auto-attention. Où des résultats prometteurs ont été obtenus dans les transformers visuels en histopathologie suggérant leur universalité [64].

9. Compte tenu de l'application réussie de Vision Transformer (ViT) dans les tâches de classification d'images, des recherches récentes ont tenté d'appliquer Transformer dans la segmentation d'images, en les combinant avec des méthodes traditionnelles basées sur CNN, donnant des résultats remarquables. Malgré les performances exceptionnelles de Transformer dans diverses tâches médicales, les coûts élevés d'apprentissage et de déploiement limitent leur application ultérieure. Il s'agit d'un défi essentiel qui réduit la complexité informatique du Transformer dans la segmentation 3D des tumeurs cérébrales. Donc, CNN et Transformer sont deux paradigmes distincts, chacun présentant des avantages et des limites où il a été démontré que les résultats des approches hybrides de segmentation des tumeurs cérébrales ont de meilleures performances [24]. Dans ce sens, nous nous sommes concentrés sur la réduction du coût de calcul de Transformer par l'utilisation de SwinUNETR s'appuyant sur une attention basée sur la répartition en fenêtres.
10. Un processus d'augmentation des données est utilisé dans ce travail pour augmenter le nombre d'images et atténuer les problèmes potentiels d'overfitting associés à la taille limitée de la base de données. Des transformations affines des pixels avec mise à l'échelle et des transformations de décalage ont été appliquées [67][65]. Par conséquent, la transforma-

tion affine génère trois images de données augmentées où chaque image source devient trois images après le processus d'augmentation des données. Par rapport aux autres réseaux de neurones artificiels, la méthode SVM est préférée en raison de certains avantages tels qu'un taux de précision plus élevé, une interprétation géométrique directe et elle ne nécessite pas un énorme échantillon pour éviter le overfitting [63].

III.3 Méthodologie proposée pour le diagnostic automatique du cancer du cerveau

L'architecture de l'approche suggérée est présentée sur la figure III.1 révélant que la méthode suggérée consiste en un prétraitement utilisant le filtre Gaussien pour réduire le bruit dans les images de l'IRM cérébrale. Après suppression du bruit, le dépouillement du crâne ou skull stripping était nécessaire pour localiser avec précision la tumeur en phase de segmentation. Ensuite, la tumeur est localisée à l'aide d'une technique de segmentation basée SwinUNETR, dont nous utilisons quatre modalités (T1w, CE_T1w, T2w et FLAIR) comme entrées pour extraire des caractéristiques à différentes échelles relatives à chaque modalité. En effet, la technologie IRM peut exploiter diverses séquences d'imagerie pour discerner des caractéristiques de tissus distincts au sein de la région tumorale.

Nous avons systématiquement affiné davantage la segmentation en présentant la tumeur segmentée en 3D. Après la segmentation, l'extraction des caractéristiques est effectuée à l'aide de diverses techniques : les ondelettes, les filtres de Gabor, les lois de Zipf et Zipf inverse ainsi que le ResNet50. Par la fin, l'algorithme SVM est appliqué pour reconnaître et classer les tumeurs cérébrales en deux types : le gliome à haut grade et celui à bas grade. Cette approche permet aux praticiens d'acquérir une compréhension plus holistique des attributs de lésions où la fusion des 4 modalités de l'IRM a considérablement augmenté la capacité du système proposer pour la classification des tumeurs cérébrales. De plus, nous soulignons l'importance du prétraitement des images et le post-traitement dans la segmentation d'images médicales où nous avons présenté la tumeur segmentée en 3D.

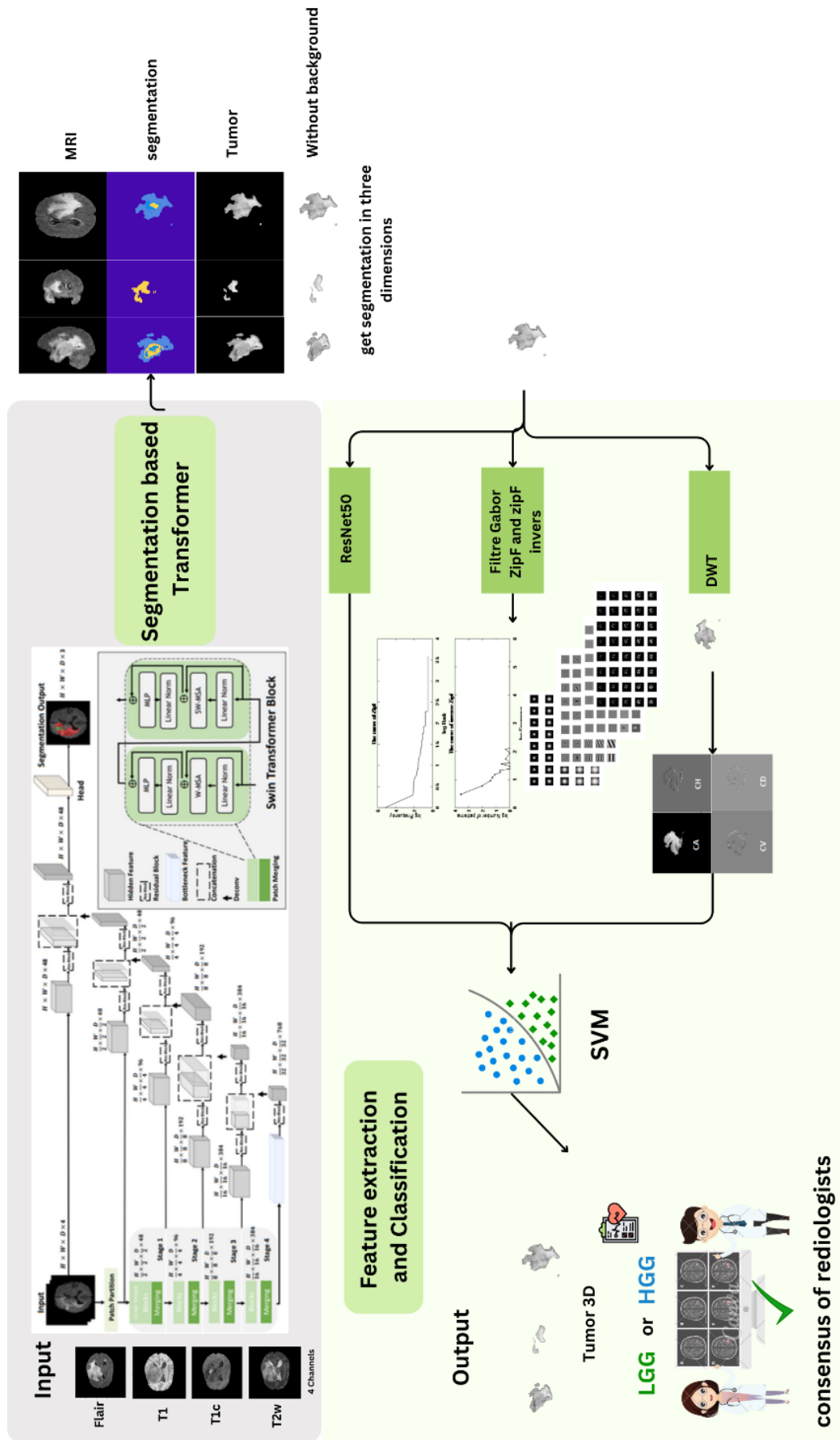


FIGURE III.1 – Architecture générale de l'approche proposée de segmentation et classification automatique du gliome

Nous allons détailler, à travers les sections suivantes, les étapes de segmentation ainsi que de classification du gliome.

III.4 Segmentation des tumeurs cérébrales dans les IRM

Le modèle Swin UnetR sera implémenté pour la segmentation des tumeurs cérébrales. En effet, nous distinguons deux phases :

III.4.1 Base de données utilisée

Le défi international de segmentation des tumeurs cérébrales est survenu depuis l'année 2012 via la base de données BraTS. En effet, BraTS désigne un environnement de référence et une base de données pour identifier les tumeurs cérébrales chez les adultes. Elle, contient des images IRM multimodales de personnes atteintes de glioblastome (HGG et LGG).

Chaque patient dispose de 4 volumes où chacun désigne un tableau 3D avec une dimension de $240 \times 240 \times 155$ où 155 est le nombre de tranches à travers le cerveau couvrant l'intégralité de l'anatomie du cerveau. Chaque cas de la base de données dispose de quatre modalités (T1c, T1n, T2f, T2w) acquises à l'aide de différents protocoles cliniques et de différents scanners provenant de plusieurs institutions fournissant des données. Tous les scans BraTS mpMRI sont disponibles sous forme de fichiers NIfTI (.nii.gz) et décrivent a) natif (T1), b) pondéré en T1 post-contraste (T1Gd), c) pondéré en T2 (T2) et d) T2 lisse, Inversion atténuée (T2-FLAIR) [68].

Tous les ensembles de données d'imagerie ont été annotés manuellement, par un à quatre examinateurs, suivant le même protocole d'annotation, et leurs annotations ont été approuvées par des neuroradiologues expérimentés. Les annotations incluent la GD renforçant la tumeur (ET - étiquette 3), le tissu œdémateux/invasif entourant la tumeur (ED - étiquette 2) et le noyau tumoral nécrosant (NCR - étiquette 1). Nous présentons sur la figure III.2 un sondage sur les bases de données les plus couramment utilisées.

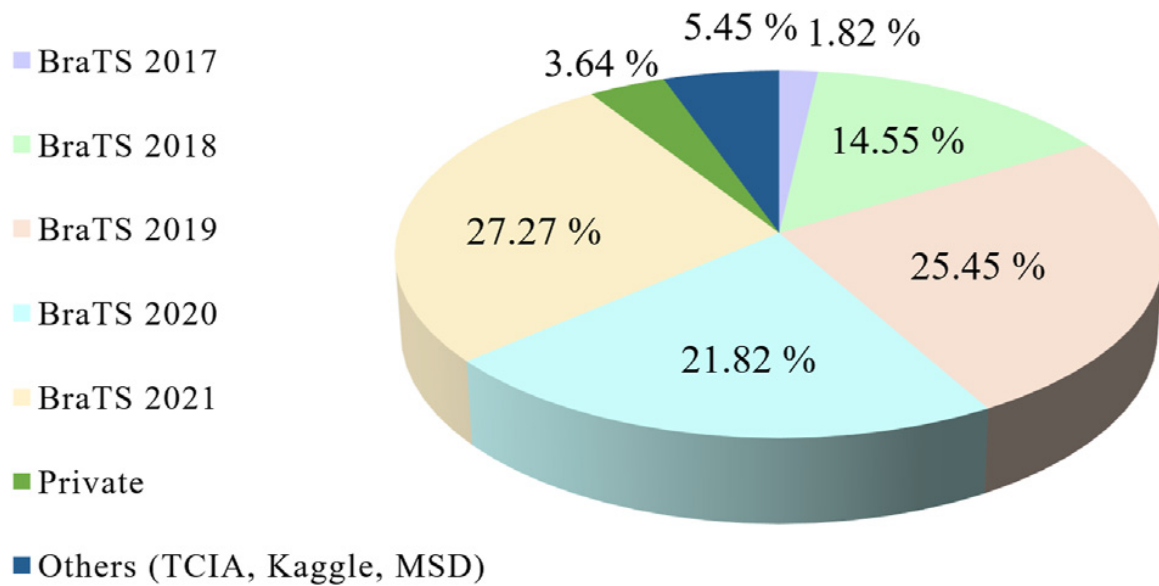


FIGURE III.2 – Résultats statistiques des bases de données les plus couramment utilisées.
Source [24]

Dans l'approche de segmentation présentée dans ce travail, nous avons utilisé la base de données BraTS 2023 Adult Glioma, qui comprend respectivement 1251 et 219 cas dans les ensembles d'apprentissage et de validation. Nous présentons sur la figure III.3 quelques IRM cérébrales de cette base de données.

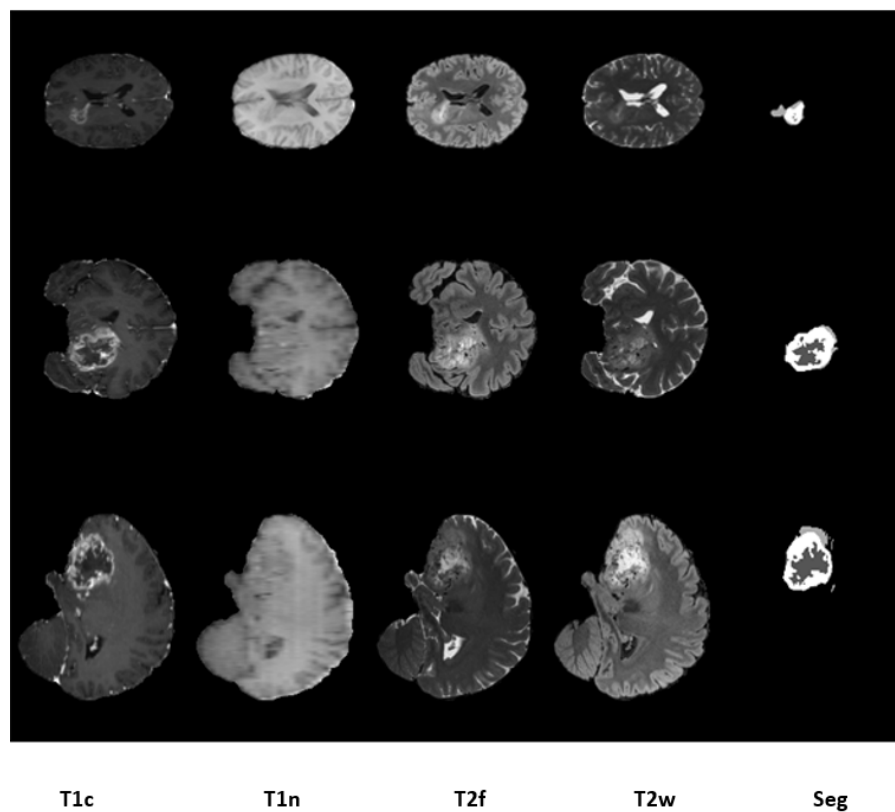


FIGURE III.3 – IRM cérébrales de la base de données BraTS 2023 avec leurs masques respectifs.

III.4.2 Prétraitement des IRM cérébrales

Nous utiliserons un filtre gaussien pour supprimer le bruit d'acquisition surgissant au sein des images IRM.

Le filtre gaussien est très efficace pour lisser les images. Il est utilisé pour rendre l'image lisse et également pour supprimer les distorsions indésirables (bruit) avant la détection des contours. Le filtrage peut être effectué en convoluant des points dans le tableau d'entrée et en les additionnant pour produire un tableau de sortie avec un noyau gaussien. Un filtrage est appliqué à l'image, « noyau » qui détermine en outre le coefficient du filtre [69].

Du moment que le filtre gaussien présente un filtre linéaire passe-bas, donc, il laisse passer les basses fréquences et atténue les hautes fréquences (il influence sur les détails de l'image consistants en les contours). Par conséquent, nous avons par la suite procédé à la technique du Skull stripping pour une meilleure précision des contours.

Quant au Skull stripping, nous utilisons des techniques de traitement d'image telles que la transformation des couleurs, l'identification des composants connectés et l'utilisation d'un masque pour isoler la région du cerveau.

Nous présentons sur la figure III.4, un exemple du prétraitement appliqué dans ce travail sur une IRM cérébrale avec ses 4 modalités.

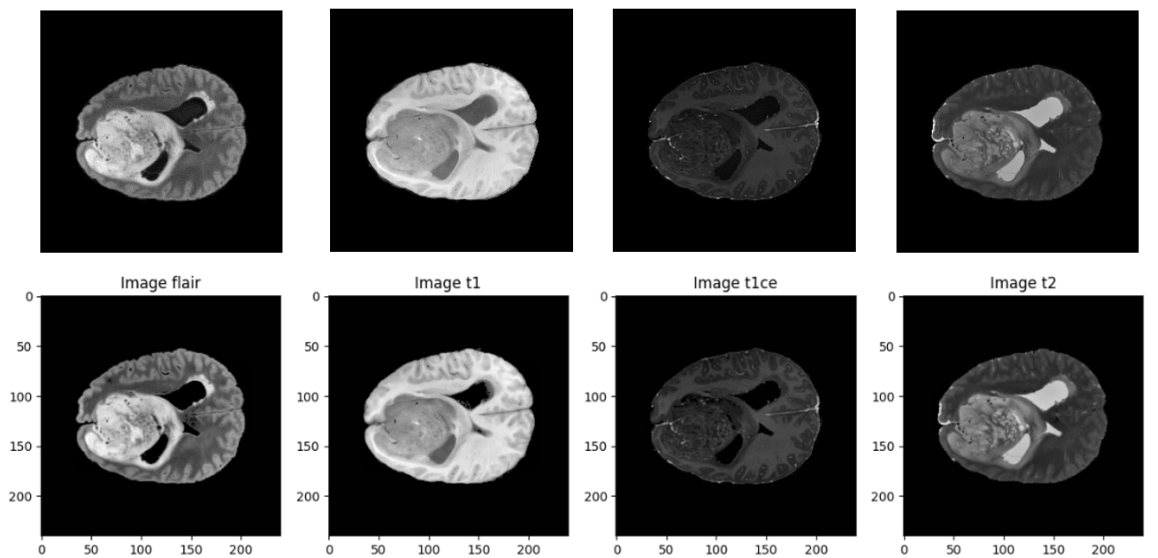


FIGURE III.4 – Prétraitement d'une IRM cérébrale de la base de données BraTS 2023 où en haut nous présentons les modalités originales et en bas suite à leur prétraitement par un filtre Gaussien suivi du skull stripping.

III.4.3 Segmentation des tumeurs cérébrales basée Swin UnetR

Nous avons implémenté le modèle proposé dans [14], comme nous pouvons le constater sur la Figure III.5, une image IRM multimodale 3D avec quatre canaux est entrée dans le modèle Swin UnetR. Le modèle crée des correctifs qui ne se chevauchent pas (creates non-overlapping patches) pour les données et utilise des partitions de correctifs (patch partition layer) pour créer des fenêtres de la taille requise pour calculer l'intérêt personnel. Nous pouvons voir ce processus sur la Figure III.6 .

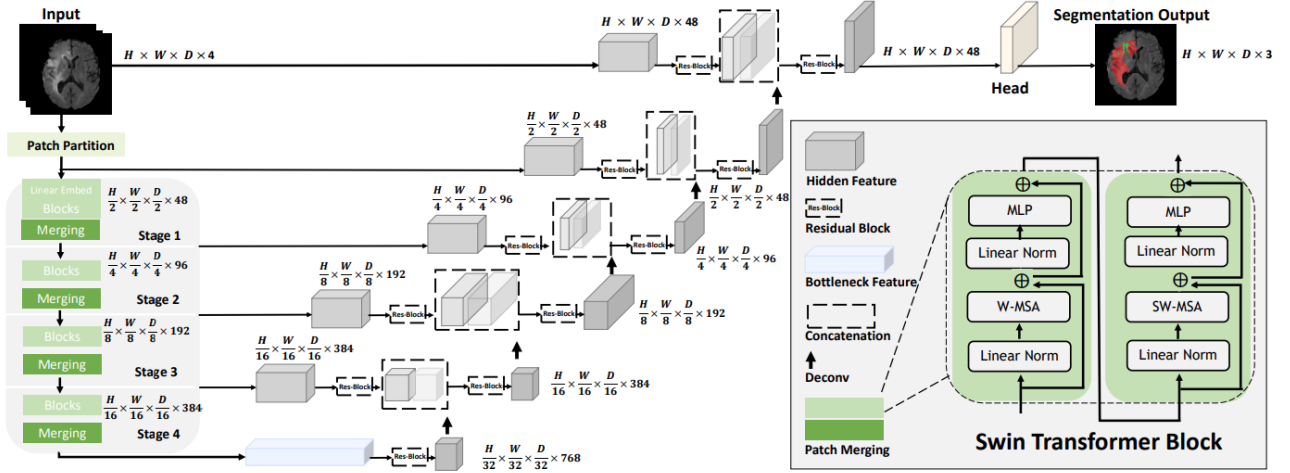


FIGURE III.5 – Architecture de Swin UnetR. Source [14]

• L'encodeur

Les entrées sont saisies dans un modèle Swin UnetR, $X \in \mathbb{R}^{H \times W \times D \times S}$ est un symbole unique avec une précision correcte de $(H' \times W' \times D')$ et des dimensions de $(H' \times W' \times D' \times S)$ (S nombre de canaux dans notre cas 4). La couche de partition de patch est d'abord appliquée pour générer une séquence de symboles 3D $\left[\frac{H}{H'}\right] \left[\frac{W}{W'}\right] \left[\frac{D}{D'}\right]$, puis les projeter dans l'espace d'intégration avec la dimension C . Suivi d'un calcul d'intérêt personnel dans des fenêtres non superposées créées lors de l'étape de partitionnement pour modéliser l'interaction discriminante efficace. La figure III.6 représente le changement de fenêtres dans les couches suivantes. Plus précisément, des fenêtres de taille $M \times M \times M$ sont utilisées pour diviser uniformément le jeton 3D en régions $\left[\frac{H'}{M}\right] \left[\frac{W'}{M}\right] \left[\frac{D'}{M}\right]$ dans une couche donnée dans l'encodeur du transformateur. Ensuite, dans la couche $l + 1$, les régions de fenêtre divisées sont décalées de $\left(\left[\frac{M}{2}\right], \left[\frac{M}{2}\right], \left[\frac{M}{2}\right]\right)$ voxels. Dans les couches suivantes de l et $l + 1$ dans l'encodeur, les sorties sont calculées comme suit :

$$\hat{z}^l = W\text{-MSA}(\text{LN}(z^{(l-1)})) + z^{(l-1)} \quad (\text{III.1})$$

$$z^l = \text{MLP}(\text{LN}(\hat{z}^l)) + \hat{z}^l \quad (\text{III.2})$$

$$\hat{z}^{l+1} = \text{SW-MSA}(\text{LN}(\hat{z}^l)) + \hat{z}^l \quad (\text{III.3})$$

$$z^{l+1} = \text{MLP}(\text{LN}(\hat{z}^{l+1})) + \hat{z}^{l+1} \quad (\text{III.4})$$

Ici, W-MSA et SW-MSA sont des modules d'auto-attention à têtes multiples réguliers et partitionnés par fenêtres respectivement ; $\hat{z}^l$ et $\hat{z}^{l+1}$ représentent les sorties de W-MSA et SW-MSA ; MLP et LN désignent respectivement la normalisation de couche et le perceptron multicouche. Pour un calcul efficace du mécanisme de décalage de fenêtre, nous utilisons un décalage cyclique en 3D et calculons l'auto-attention en :

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (\text{III.5})$$

Ceci dans lequel Q, K et V représentent respectivement les requêtes, les clés et les valeurs ; d représente la taille de la requête et de la clé. L'encodeur Swin UnetR a une taille de patch de $2 \times 2 \times 2$ et une dimension de caractéristique de $2 \times 2 \times 2 \times 4 = 32$, en tenant compte des images IRM multimodales avec 4 canaux. La taille de l'espace d'encodage C est fixée à 48 dans notre encodeur.

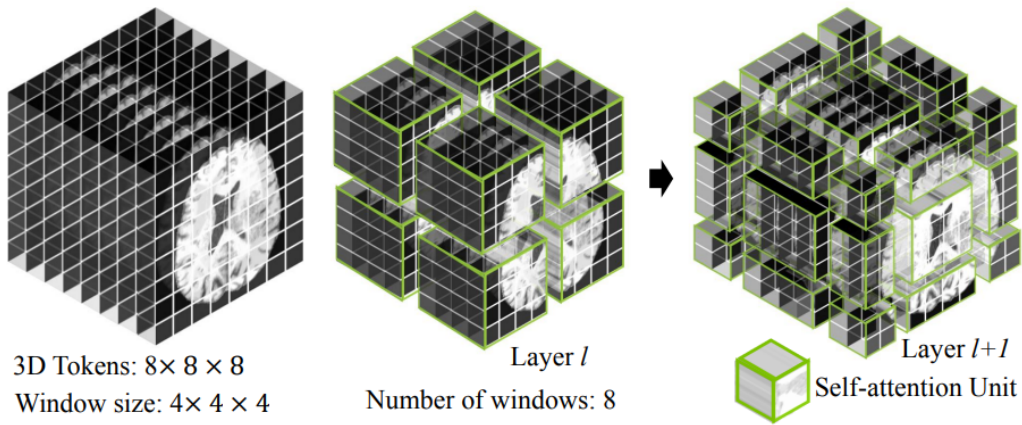


FIGURE III.6 – Vue d'ensemble du mécanisme de fenêtrage décalé. Source [14]

De plus, l'encodeur Swin UnetR a 4 étapes, chacune composée de 2 blocs transformateurs. Par conséquent, le nombre total de couches dans l'encodeur est de $L = 8$. À l'étape 1, une couche d'encodage linéaire est utilisée pour créer des jetons $\frac{H}{2} \times \frac{W}{2} \times \frac{D}{2}$ 3D. Pour maintenir la structure hiérarchique de l'encodeur, une couche de fusion de patches est utilisée pour diminuer la résolution des représentations de caractéristiques d'un facteur de 2 à la fin de chaque étape. De plus, une couche de fusion de patches groupe des patches avec une résolution de $2 \times 2 \times 2$ et les concatène, ce qui donne une représentation de caractéristique de dimension $4C$. La taille des représentations de caractéristiques est ensuite réduite à $2C$ avec une couche linéaire. Les

étapes 2, 3 et 4, avec des résolutions respectives de $\frac{H}{4} \times \frac{W}{4} \times \frac{D}{4}$, $\frac{H}{8} \times \frac{W}{8} \times \frac{D}{8}$ et $\frac{H}{16} \times \frac{W}{16} \times \frac{D}{16}$ suivent la même conception de réseau.

- **Le décodeur**

Swin UNET Transformers (Swin UnetR), dont l'architecture est connectée à un décodeur basé sur CNN à différentes résolutions via des connexions sautées, le Swin UNETR a une conception de réseau en forme de U dans laquelle les représentations de caractéristiques extraites de l'encodeur sont utilisées dans le décodeur via des connexions de saut à chaque résolution. À chaque étape i (i appartenant à l'ensemble 0,1,2,3,4) dans l'encodeur et dans la zone de rétrécissement ($i=5$), les représentations de caractéristiques de sortie sont remodelées en taille $\frac{H}{2^i} \times \frac{W}{2^i} \times \frac{D}{2^i}$ et introduites dans un bloc résiduel composé de deux couches de convolution $3 \times 3 \times 3$ normalisées par des couches de normalisation par instance. Par la suite, la résolution des cartes de caractéristiques est augmentée d'un facteur de 2 à l'aide d'une couche de déconvolution, et les sorties sont concaténées avec les sorties de l'étape précédente. Les caractéristiques concaténées sont ensuite introduites dans un autre bloc résiduel comme décrit précédemment. Les sorties de segmentation finales sont calculées en utilisant une couche de convolution $1 \times 1 \times 1$ et une fonction d'activation sigmoïde. Nous utilisons la fonction de perte douce de Dice qui est calculée par voxel comme suit :

$$L(G, Y) = 1 - \frac{2}{J} \sum_{j=1}^J \frac{\sum_{i=1}^I G_{i,j} Y_{i,j}}{\sqrt{\sum_{i=1}^I G_{i,j}^2 \sum_{i=1}^I Y_{i,j}^2}} \quad (\text{III.6})$$

Où I représente le nombre de voxels ; J le nombre de classes ; $Y_{i,j}$ et $G_{i,j}$ représentent respectivement la probabilité de sortie et la vérité de terrain codée à une touche pour la classe j au niveau du voxel i [14]. Le résultat final est la sortie de segmentation finale qui consiste en 3 canaux de sortie correspondant aux sous-régions ET, WT et TC, où : TC est le cœur de la tumeur, WT est la tumeur entière et ET est la tumeur d'amélioration.

III.4.4 Segmentation des tumeurs cérébrales basée CNN (SegResNet)

SegResNet (Segmentation-ResNet) est un type de modèle profond utilisé dans le domaine de la segmentation d'images comme mentionné sur la figure III.7, en particulier d'images médicales. Cette approche suit une architecture CNN basée sur un décodeur avec un encodeur asymétrique plus grand pour extraire les caractéristiques de l'image et un décodeur plus petit pour reconstruire le masque de segmentation. L'entrée est un recadrage IRM 3D à quatre

canaux, suivi d'une convolution 3D initiale $3 \times 3 \times 3$ avec 32 filtres. Chaque bloc vert est un bloc de type ResNet avec la normalisation GroupNorm. La sortie du décodeur de segmentation comporte trois canaux (avec la même taille spatiale que l'entrée) suivis d'un sigmoïde pour les cartes de segmentation des trois sous-régions tumorales (WT, TC, ET). La branche VAE reconstruit l'image d'entrée en elle-même et n'est utilisée que lors de l'apprentissage pour régulariser l'encodeur partagé.

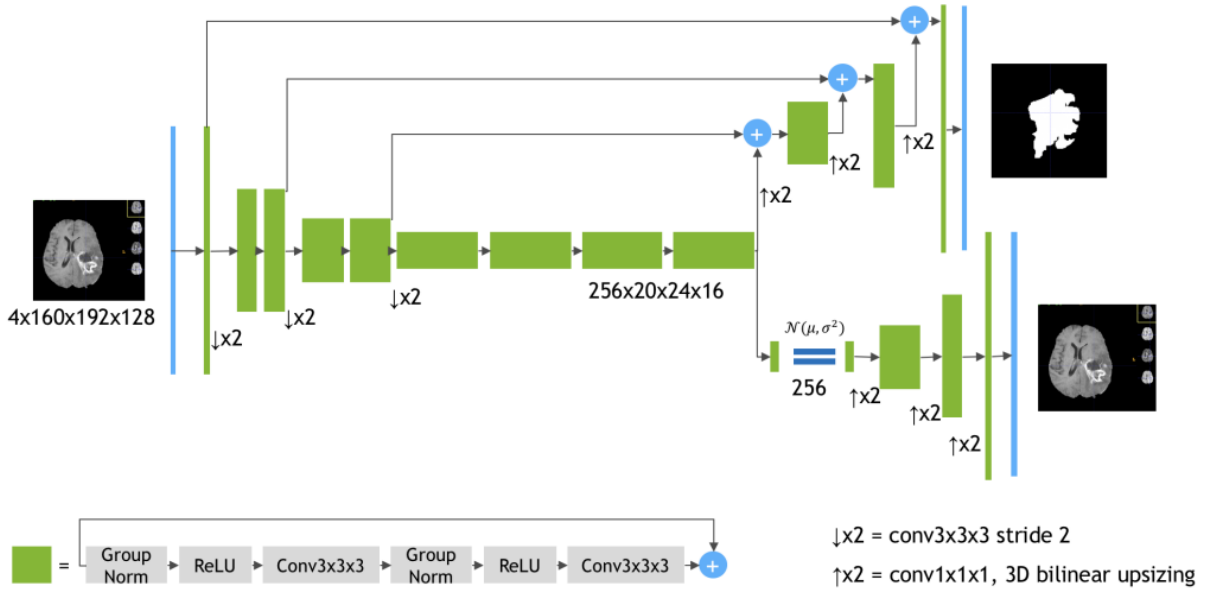


FIGURE III.7 – Architecture de SegResNet. Source [26]

• Partie encodeur

La partie encodeur utilise des blocs ResNet, où chaque bloc se compose de deux bobines avec normalisation et ReLU, suivies d'une connexion additive de saut d'identité. Pour la normalisation, on utilise la normalisation de groupe (GN), qui présente de meilleures performances que la normalisation par lots (BatchNorm) lorsque la taille des lots est faible. Elle suit l'approche CNN commune pour réduire progressivement la dimension de l'image de 2 et augmenter la taille des caractéristiques de 2 en même temps. Pour réduire la taille, il utilise des convolutions. Toutes les convolutions sont de $3 \times 3 \times 3$ avec un nombre initial de filtres égal à 32. La taille du point final de cryptage est de $256 \times 20 \times 24 \times 16$, ce qui est spatialement 8 fois plus petit que l'image d'entrée.

- **Partie décodeur**

La structure du décodeur est similaire à celle du codeur, mais avec un bloc par niveau spatial. Chaque niveau du décodeur commence par augmenter la taille : en réduisant le nombre de caractéristiques par un facteur de 2 (à l'aide de convolutions $1 \times 1 \times 1$) et en doublant la dimension spatiale (à l'aide d'une approximation linéaire 3D), puis en ajoutant la sortie de chiffrement du niveau spatial équivalent. L'extrémité du décodeur a la même dimension spatiale que l'image originale, et le nombre de caractéristiques est égal à la taille de la caractéristique d'entrée initiale, suivie d'une convolution $1 \times 1 \times 1$ en 3 canaux et d'une fonction σ .

- **Perte**

Notre fonction de perte se compose de 3 termes :

$$L = L_{\text{dice}} + 0.1 \times L_{L2} + 0.1 \times L_{KL} \quad (\text{III.7})$$

L_{dice} est une perte de dé douce qui est appliquée à la sortie du décodeur pour correspondre au masque de hachage correct. Le masque de hachage correct :

$$L_{\text{dice}} = \frac{2 \times \sum P_{\text{true}} \times P_{\text{pred}}}{\sum P_{\text{true}}^2 + \sum P_{\text{pred}}^2 + \epsilon} \quad (\text{III.8})$$

L'addition se fait en termes de pixels, ce qui est une petite constante pour éviter de diviser par zéro. Étant donné que la sortie du décodeur de hachage comporte trois canaux (ET, WT et TC, où : TC est le cœur de la tumeur, WT est la tumeur entière et ET est la tumeur améliorée), les trois fonctions de perte de dés sont simplement additionnées. L_{L2} est la perte $L2$ à la sortie de la branche VAE I_{pred} pour correspondre à l'image d'entrée. L'image d'entrée :

$$L_{L2} = \|I_{\text{input}} - I_{\text{pred}}\|_2^2 \quad (\text{III.9})$$

L_{KL} est un terme de pénalité standard de la VAE, qui est la différence KL entre la distribution normale estimée $N(\mu^2 + \sigma^2)$ et la distribution préalable $N(0,1)$ qui a une représentation fermée :

$$L_{KL} = \frac{1}{N} \sum (\mu^2 + \sigma^2 - \log \sigma^2 - 1) \quad (\text{III.10})$$

N est le nombre total de voxels de l'image. Il a été constaté empiriquement qu'un paramètre excessif de 0,1 pour obtenir un bon équilibre entre les conditions de dés et la perte VAE. Équation (III.7)[26].

III.5 Classification des gliomes en LGG ou HGG

Nous présentons sur la figure III.8 le pipeline suivi pour la classification des gliomes.

• Training phase

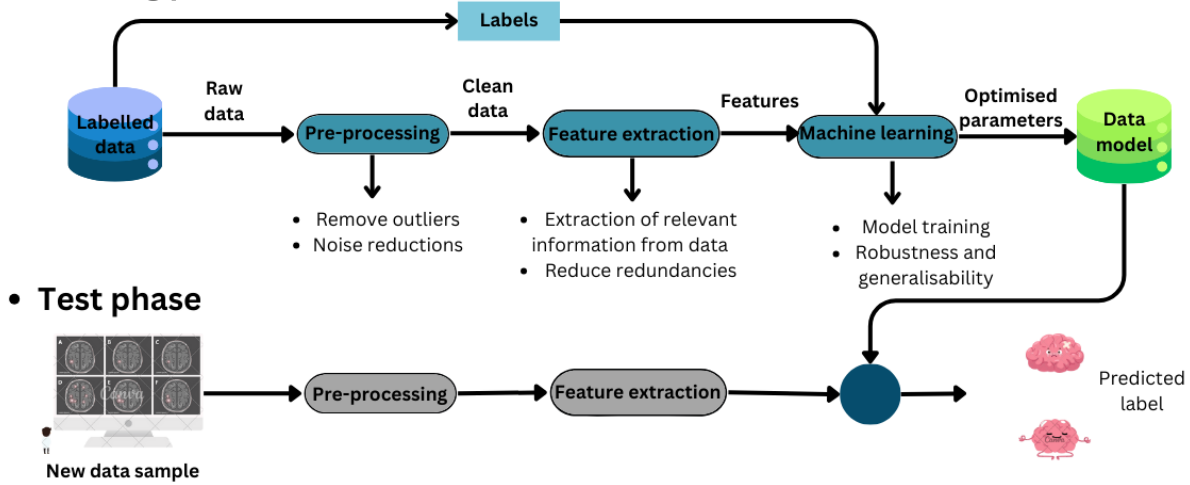


FIGURE III.8 – Pipeline adopté pour la classification des gliomes en LGG ou HGG.

Nous allons détailler les étapes à travers les sous-sections suivantes.

III.5.1 Base de données utilisée pour la classification

A cause du fait que le labeling (tumeur LGG ou tumeur HGG) n'est pas disponible pour la base de données BraTS 2023, nous avons utilisé la base de données BraTS 2020, qui contient 369 IRM, chaque cas contenant 155 diapositives.

Nous avons converti toutes les IRM au format d'image PNG en vue du processus de classification.

Une phase d'augmentation de données a été adoptée pour palier à l'inconvénient du overfitting engendrant une augmentation du nombre des IRM de 369 à 1107. En effet, ceci consistait à sélectionner les diapositives numérotées 80, 85 et 90, qui sont les diapositives dans lesquelles la tumeur apparaît clairement puisqu'il s'agit d'une séquence d'images (qui forment une vidéo), chaque diapositive était différente comme nous l'observons sur la figure III.9.

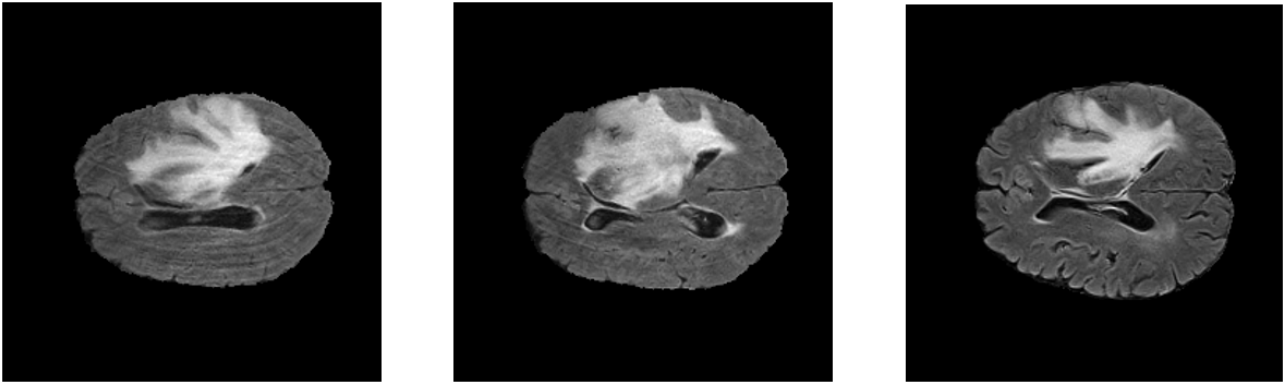


FIGURE III.9 – Méthode proposée d’augmentation des données par la considération d’autres slices 80 et 85 et 90 respectivement.

III.5.2 Cropping des gliomes

À ce stade, nous avons traité les données en réalisant avec précision le cropping des gliomes de BraTS 2020 afin d’en extraire facilement les caractéristiques ultérieurement. Nous présentons sur la figure III.10 le processus du cropping.



FIGURE III.10 – Résultat du cropping d’une tumeur.

III.5.3 Extraction et sélection des descripteurs

La phase de classification des gliomes consiste à extraire les descripteurs les plus discriminants des tumeurs, car l'identification des descripteurs peut aider à réduire les erreurs de classification. Il existe de nombreuses techniques utilisées pour extraire des caractéristiques des images médicales, cependant il est difficile de déterminer la meilleure et la plus appropriée dans ce domaine, nous avons donc dû essayer de nombreuses techniques comme suit :

- **Fusion des descripteurs des lois de Zipf et de Zipf inverse avec les filtres de Gabor**

Il s'agit d'une technique basée sur l'utilisation des lois de puissance de Zipf et de Zipf inverse pour analyser les tumeurs résultantes après avoir identifié avec précision la tumeur à partir de l'image originale du cerveau affecté. La technique repose sur la distinction du contenu structurel des tumeurs pour la quantification des tissus texturaux. En effet, la texture désigne un descripteur pertinent de la vision par ordinateur et spécialement dans le domaine médical. Cette fusion se base sur le principe de fusionner les descripteurs issus des lois de Zipf et Zipf inverse agissant dans le domaine spatial avec ceux des filtres de Gabor agissant dans le domaine spatial ainsi que fréquentiel pour bénéficier de leurs apports complémentaires. Les figures [III.11](#), [III.12](#) représentent respectivement les courbes de Zipf et Zipf inverses d'un gliome à haut grade et un gliome à bas grade.

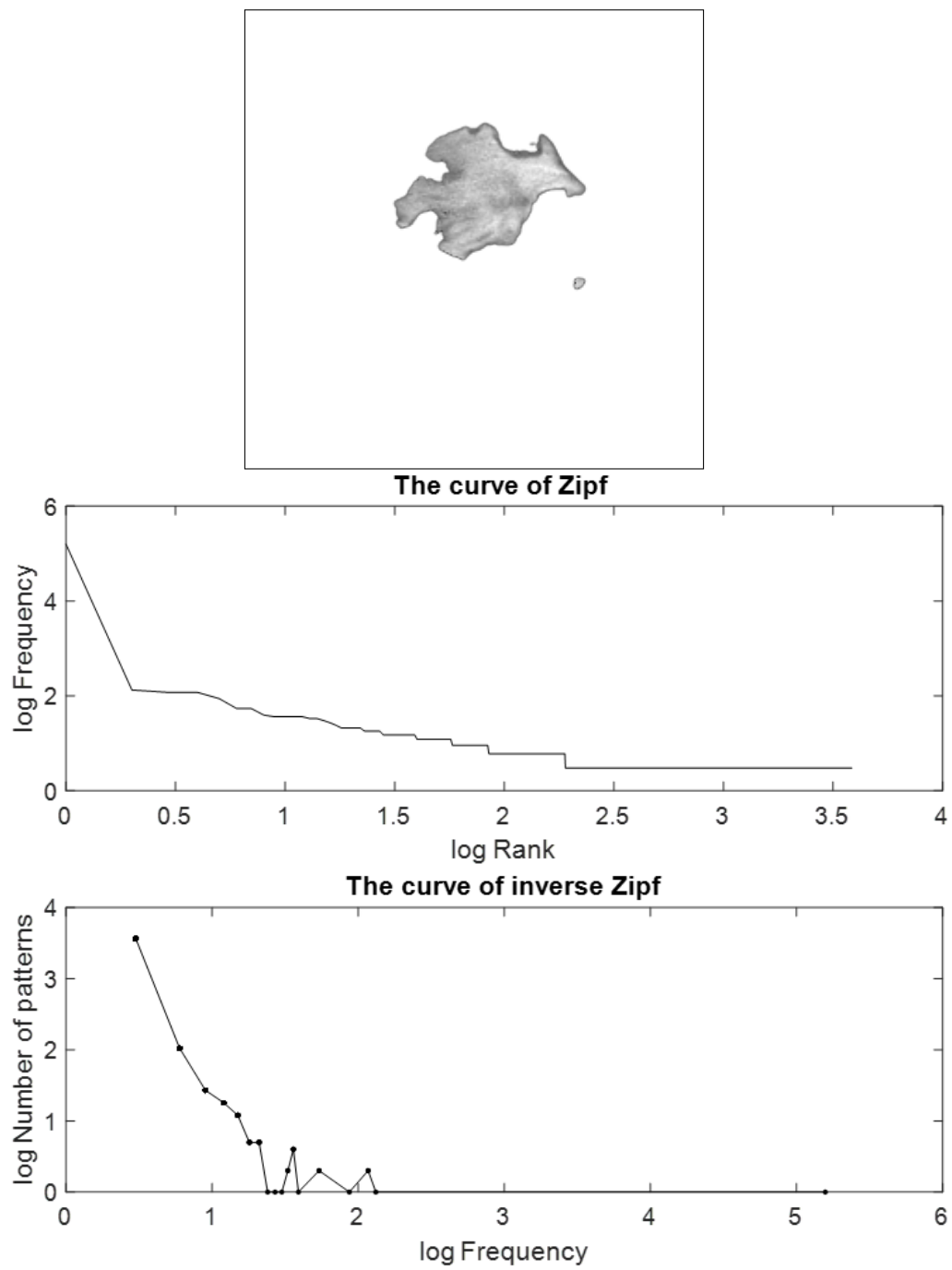


FIGURE III.11 – Courbes de Zipf et de Zipf inverse d'un gliome à haut grade (HGG).

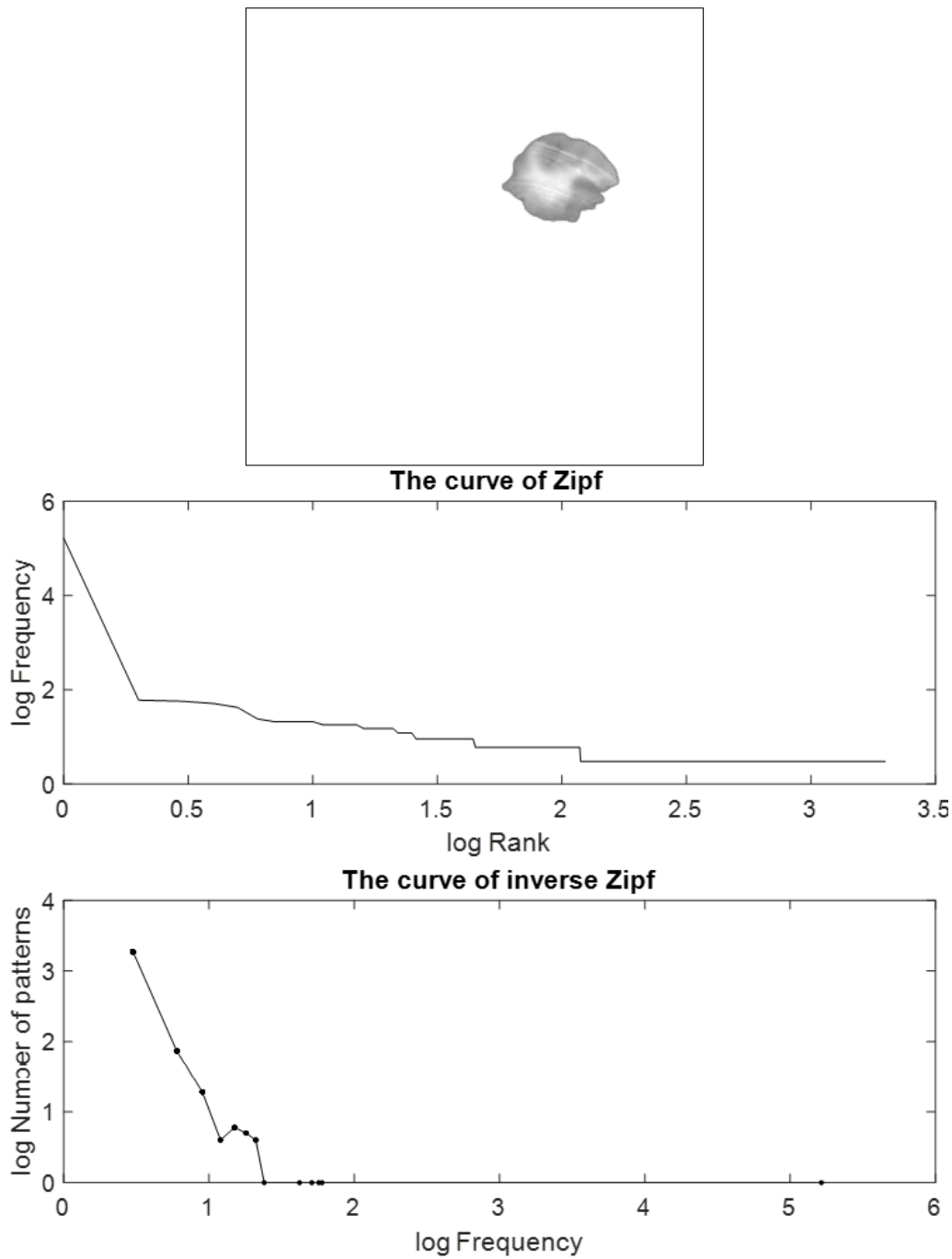


FIGURE III.12 – Courbe de Zipf et de Zipf inverse d'un gliome à bas grade (LGG).

Nous extrayons les descripteurs suivant à partir des courbes de Zipf et de Zipf inverses [70] :

Les pentes des courbes de Zipf et de Zipf inverse La pente moyenne d'une courbe est le coefficient directeur de la droite des moindres carrés. Elle est donnée par la formule III.11 :

$$p = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (\text{III.11})$$

L'aire délimitée par la courbe de Zipf On calcule l'aire spécifique de la courbe Zipf à l'aide des

courbes Zipf via la formule suivante :

$$A = \sum_{i=1}^{n-1} \frac{(f_i + f_{i+1})(r_{i+1} - r_i)}{2} \quad (\text{III.12})$$

n est le nombre de motifs dans la courbe, f_i est la fréquence et r_i est le rang du motif. Entropie1 de la courbe de Zipf L'entropie des modèles d'images de mammographie est déterminée par la formule III.13 :

$$Hw = - \sum_{r=1}^R \left(\frac{f(r)}{T} \log \frac{f(r)}{T} \right) \quad (\text{III.13})$$

Où $f(r)$ est la fréquence du motif pour la ligne r , T est le nombre total de motifs différents et nous utilisons un logarithme de base R . Entropie2 de la courbe de Zipf L'entropie relative à la fréquence d'apparition des motifs est définie par la formule suivante :

$$Hf = - \sum_{f=1}^F \left(\frac{I(f)}{R} \log F \frac{I(f)}{R} \right) \quad (\text{III.14})$$

Où $I(f)$ représente le nombre de motifs distincts qui ont une fréquence d'apparition égale à f et F est le nombre entier d'occurrences de motifs dans l'image. Les ordonnées à l'origine des courbes de Zipf et Zipf inverse La constante alpha Pour tout motif appartenant à l'image, la fréquence d'apparition de ce motif * son rang dans la liste ordonnée décroissante des fréquences des motifs de l'image représente une constante. Quant aux filtres de Gabor traite des images de gliomes similairement à la perception du système visuel humain. Le filtre de Gabor 2D est une fonction de noyau gaussien modifiée par une fonction sinusoïdale complexe, où la formule du filtre de Gabor 2D est la suivante :

$$h(x, y) = \frac{1}{\sigma_x \sigma_y 2\pi} \exp \left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} \right) \exp (j2\pi u_0 (x \cos \theta + y \sin \theta)) \quad (\text{III.15})$$

- u_0 : la fréquence spatiale.
- θ : la direction de modulation du filtre.
- σ_x et σ_y : Ce sont les écarts-types (ou déviations standard) qui contrôlent la largeur de la distribution gaussienne dans les directions x et y respectivement.
- $\exp \left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} \right)$: C'est le terme gaussien qui décrit la distribution spatiale de la fonction. Il représente une gaussienne bidimensionnelle centrée en $(0,0)$ avec des écartstypes σ_x et σ_y dans les directions x et y respectivement.
- $\exp (j2\pi u_0 (x \cos \theta + y \sin \theta))$: C'est le terme de modulation sinusoïdal qui introduit une variation de phase spatiale en fonction de la direction θ et de la fréquence spatiale u_0 . Ce terme peut être interprété comme une onde sinusoïdale dépendante de la position (x,y) dans l'espace.

Les filtres de Gabor consistent en un type de filtrage passe-bande conservant l'information extraite au niveau du spectre relatif à une image à travers la bande de fréquence sélectionnée. Par conséquent, ces filtres peuvent analyser des caractéristiques locales des tumeurs. La configuration la plus utilisée d'extraction de caractéristiques spatio-fréquentielles dans les textures est obtenue par la convolution de la région tumorale avec une banque de filtres [71] où chacun se centrera sur une fréquence ainsi qu'une orientation pour couvrir tout le domaine fréquentiel. Effectivement, chaque pixel appartenant à une tumeur donnera une réponse pour chaque filtre. Au niveau de la classification, nous analysons la relation qui existe entre les réponses issues des filtres de Gabor de chacune des tumeurs. Nous avons appliqué une banque des filtres de Gabor avec une taille de 39 x 39 selon 5 fréquences et 8 orientations comme mentionné sur la figure III.13 [72] :

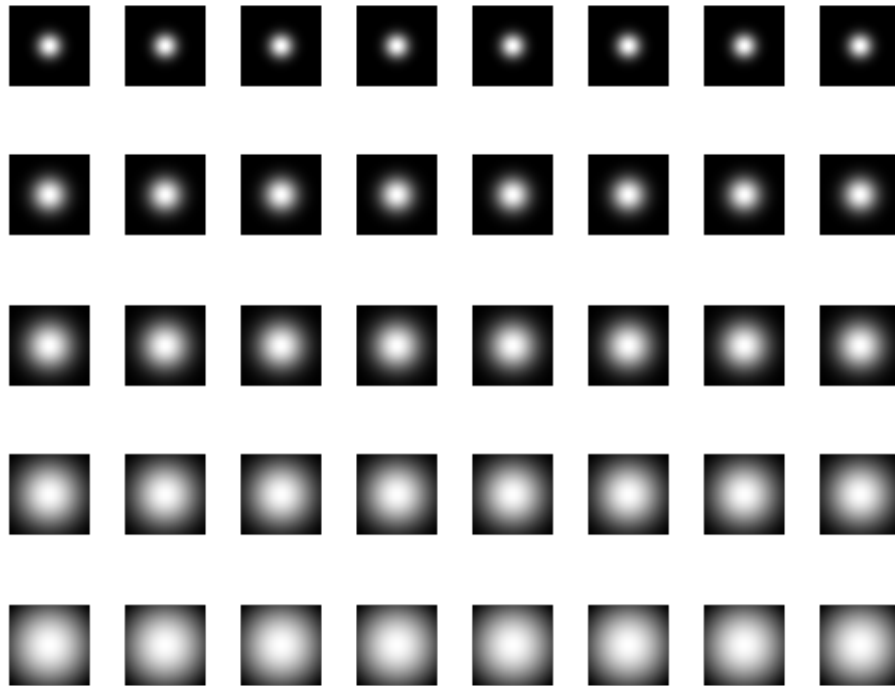


FIGURE III.13 – Application des filtres de Gabor dans le domaine fréquentiel.

Suite à l'utilisation d'un ensemble riche en variation de filtres de Gabor, ceci permettra une couverture plus étendue de l'espace fréquentiel en détectant davantage d'orientations. La figure III.14, démontre la partie réelle de ces filtres.

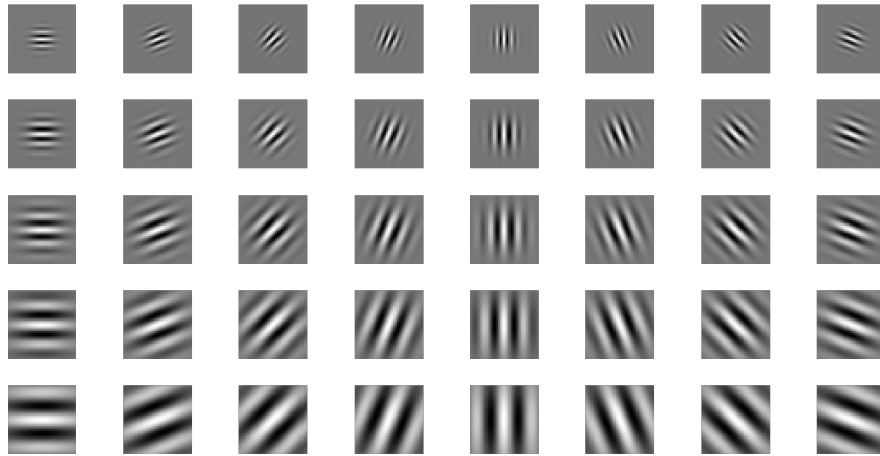


FIGURE III.14 – Parties réelles des filtres de Gabor utilisés.

Nous allons démontrer sur la figure [III.15](#), la partie réelle des tumeurs filtrées :

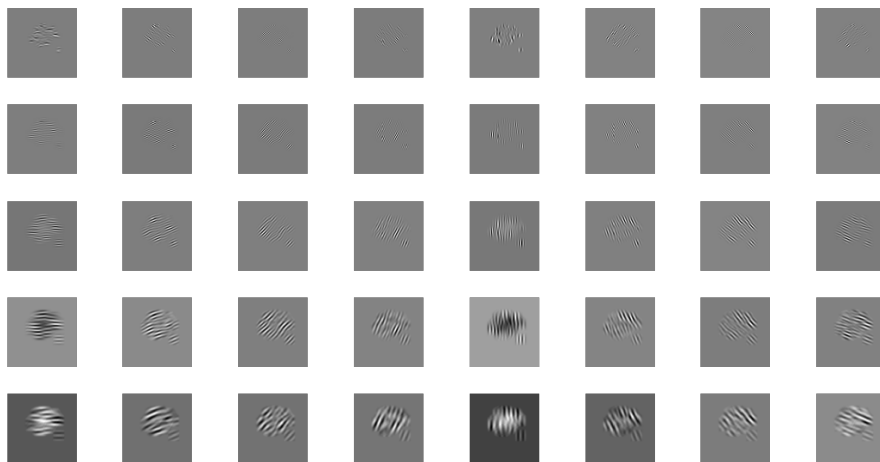


FIGURE III.15 – Parties réelles des tumeurs filtrées.

La figure [III.16](#) démontre la magnitude de la réponse des tumeurs filtrées.

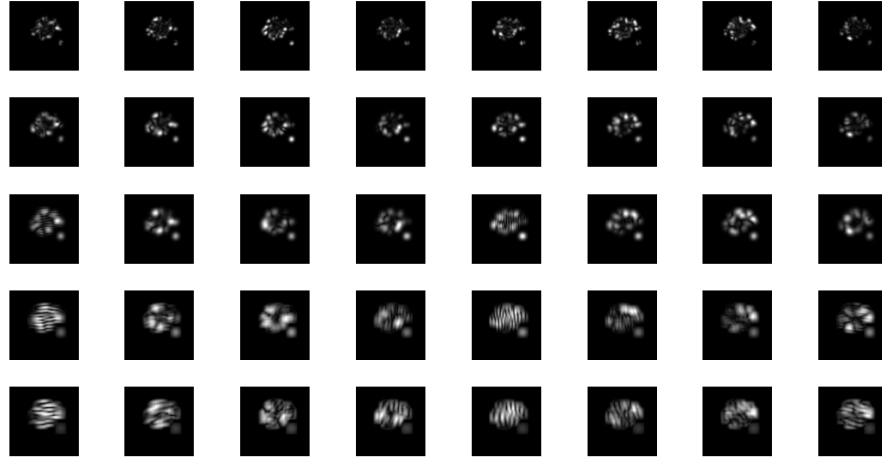


FIGURE III.16 – La magnitude de la réponse relative au filtre de Gabor suite à la convolution de la tumeur avec une banque de filtres de Gabor.

Une fois toutes les sorties du filtre sont disponibles, nous calculons des statistiques locales pour générer le vecteur des descripteurs texturaux comme indiqué dans [72].

- **RESNet50**

Nous avons utilisé le modèle de deep learning ResNet50 pour l'extraction des descripteurs. Il s'agit d'un modèle CNN qui est formé sur diverses combinaisons, notamment des couches convolutionnelles, de pooling et entièrement connectées [73]. Nous l'utiliserons comme extracteur de caractéristiques pour les données représentées par de fines images de gliomes, car ResNet50 a une profondeur élevée avec 50 couches, ce qui le rend performant pour l'extraction des caractéristiques complexes et profondes des images.

Cette approche est idéale pour les applications médicales nécessitant une analyse précise des images, telles que la détection de tumeurs cérébrales. Les couches inférieures du modèle sont utilisées pour extraire les caractéristiques des images.

Nous présentons sur la figure III.17 la structure de ResNet50.

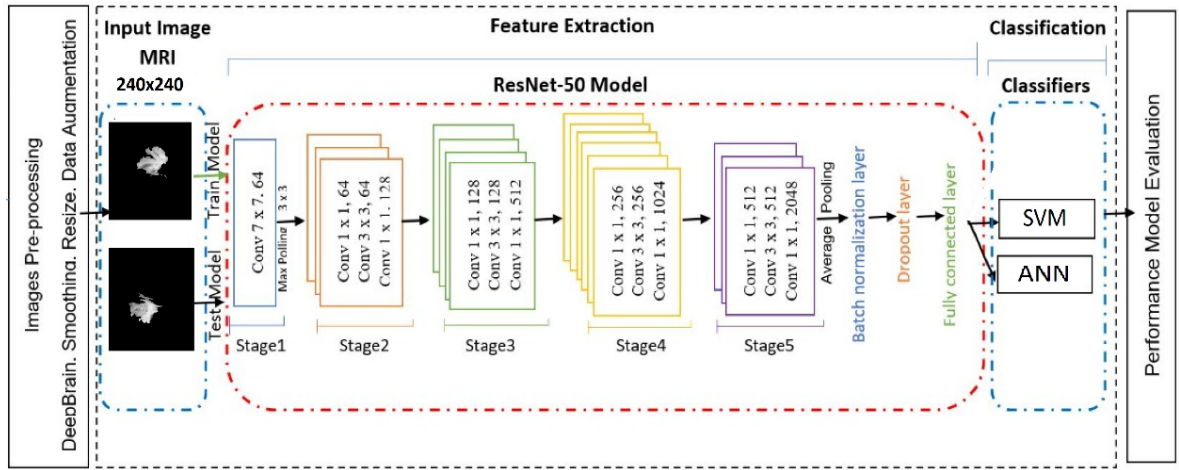


FIGURE III.17 – Structure générale de ResNet50. Source[27]

• Les Ondelettes (DWT)

La détection des tumeurs cérébrales est l'un des domaines médicaux les plus importants, où la détection et le diagnostic précoces des tumeurs cérébrales aident à déterminer un plan de traitement précis, ce qui à son tour augmente les chances de succès du traitement, les méthodes de transformée en ondelettes sont des méthodes efficaces pour extraire les caractéristiques des images médicales, y compris la transformée en ondelettes discrète (DWT) [74], qui est une technique puissante de traitement des signaux et a été utilisée dans le domaine de l'imagerie médicale, y compris la détection et la classification des tumeurs cérébrales, nous parlerons en général de la transformée en ondelettes et ensuite nous nous plongerons dans la transformée en ondelettes discrète (DWT).

La transformée en Ondelettes est une technique qui permet d'analyser une image en la décomposant en une série d'images en sous-bandes. Les images en sous-bandes sont considérées comme des coefficients d'ondelettes, puis les caractéristiques sont calculées ; la transformée en ondelettes est exprimée par la formule suivante :

$$\omega(t) = \int_{-\infty}^{+\infty} |\phi(t)|^2 dt \quad (\text{III.16})$$

La transformée en ondelettes discrète (DWT) est une technique mathématique qui décompose un signal en différentes composantes de fréquence à l'aide de fonctions d'ondelettes. Contrairement à la transformée de Fourier, qui représente un signal à l'aide de fonctions sinusoïdales à différentes fréquences, la transformée en ondelettes discrète permet une analyse multi-résolution des images médicales, où les images sont décomposées en composantes approximatives (basse fréquence) et détaillées (haute fréquence) à différents niveaux. Dans le cas de la détection des tumeurs cérébrales, cette analyse multi-résolution permet de capturer des détails

importants liés au tissu tumoral, aux bords et à d'autres caractéristiques à différentes échelles, la figure III.18 illustre l'application de la DWT 2D à une tumeur du gliome.

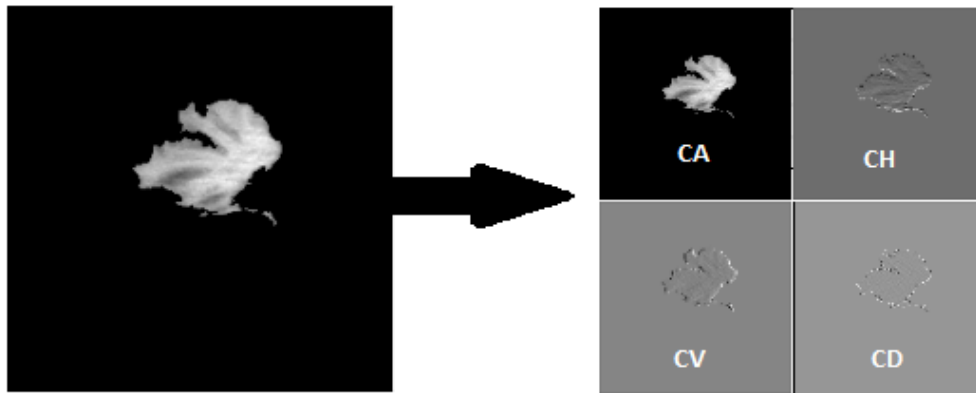


FIGURE III.18 – Application de la DWT 2D à une tumeur du gliome.

Où :

- cA : Section d'approximation, qui est une approximation de l'image à basse résolution.
- cH : Détail horizontal, un signal contenant des détails horizontaux à haute fréquence.
- cV : Détail vertical, signal contenant des détails verticaux à haute fréquence.
- cD : Diagonal Détail, signal contenant des détails diagonaux à haute fréquence.

III.5.4 Classification des gliomes à bas grade et à haut grade par les machines à vecteurs de support (SVM)

Après avoir obtenu les caractéristiques de la tumeur, l'étape suivante consiste à utiliser l'un des modèles de classification les plus puissants, le classificateur SVM (Support Vector Machine), pour entraîner et tester les données. Le SVM est un modèle d'apprentissage automatique utilisé pour classer les données en deux classes ou plus. Le modèle utilise les données d'entrée, c'est-à-dire l'ensemble des caractéristiques, et les classe en deux classes. Il divise ces données en données d'apprentissage et de test, puis nous définissons des hyperparamètres : Le noyau utilisé (linéaire, RBF kernel, ...), la valeur du paramètre de régularisation C (entre 0,1 et 100) et gamma (scale, auto). En effet, nous avons utilisé GridSearchCV, une fonctionnalité de la bibliothèque Scikit-learn (sklearn) qui nous permet d'effectuer une recherche globale des meilleurs hyperparamètres pour un modèle.

III.6 Conclusion

Au long de ce chapitre, nous avons examiné en détail l'approche proposée pour diagnostiquer les gliomes, ainsi que les techniques et les étapes impliquées dans ce processus : traitement initial des données, segmentation des tumeurs cérébrales, identification, extraction des caractéristiques et classification.

CHAPITRE IV

EVALUATION DES PERFORMANCES DE L'APPROCHE PROPOSÉE

IV.1 Introduction :

Après avoir présenté l'approche adoptée dans le chapitre précédent, nous présentons dans ce chapitre les résultats obtenus après avoir implémenté l'approche, ainsi que les outils et les méthodes utilisés pour mettre en œuvre l'approche. Bien évidemment, une évaluation des résultats obtenus est en ordre.

IV.2 Environnement De Développement Et Outils Utilisés

IV.2.1 Langage Python

Le langage Python est un langage de programmation de haut niveau, développé à l'Institut néerlandais de mathématiques et d'informatique (CWI) d'Amsterdam par Guido van Rossum à la fin des années 1980, et annoncé pour la première fois en 1991[75], il était moins répandu à l'époque, bien que ces dernières années il ait connu une large diffusion jusqu'à se classer premier dans l'index Tiobe[76], qui est mis à jour mensuellement, grâce à sa capacité à l'apprentissage automatique, à la science des données et à l'intelligence artificielle. Il est utilisé pour créer des applications web, développer des algorithmes et analyser des données, et c'est un langage de programmation facile à apprendre.



- **Raison d'utilisation de Python**

Parce qu'il est facile à utiliser et à apprendre, et parce qu'il procède de nombreuses bibliothèques contenant un large éventail de fonctions et d'outils qui peuvent être utilisés directement sans avoir à les écrire de A à Z. Cela économise beaucoup de temps et d'efforts pour les développeurs en économisant beaucoup de temps et d'efforts.

Par exemple, la bibliothèque Scikit-learn fournit un certain nombre d'implémentations d'algorithmes d'apprentissage automatique tels que la régression linéaire, la régression logistique, les machines à vecteurs de support, etc. D'autres bibliothèques incluent spaCy, Natural Language Toolkit (NLTK), etc. Les bibliothèques TensorFlow, PyTorch et Keras sont très populaires au sein de la communauté de l'intelligence artificielle. De nombreuses autres bibliothèques, telles que NumPy, Pandas et Seaborn, facilitent le traitement des données. Ces bibliothèques, qui ne sont disponibles qu'en Python, réduisent le temps et la complexité de l'écriture du code [77].

- **Anaconda**

Anaconda est une plateforme de distribution et d'installation de paquets Python et Conda provenant de diverses sources. Elle prend en charge la science des données, l'apprentissage automatique et d'autres domaines [78].

- **Jupyter**

Jupyter Notebook est un environnement de développement open source utilisé pour créer des documents interactifs contenant des graphiques, du code, du texte et des images, conçu spécifiquement pour travailler avec Python et d'autres langages de programmation tels que R et Julia, raison pour laquelle nous utilisons cet environnement. Jupyter Notebook est populaire parmi les développeurs en raison de ses capacités interactives et de sa facilité d'utilisation pour l'analyse de données, le développement de modèles et la visualisation.

- **Kaggle**

Kaggle est une plateforme en ligne qui constitue la plus grande communauté mondiale de data scientists et d'ingénieurs en apprentissage automatique (machine learning) [79][80]. Kaggle offre de puissants outils et ressources pour aider les utilisateurs à atteindre leurs objectifs en science des données [79][80]. Kaggle a été fondée en 2010 par Anthony Goldbloom (en) 2 et Google a acquis la société en 2017.



Nous allons citer les principales fonctionnalités de Kaggle comme suit :

- Héberger des jeux de données publics et permettre aux utilisateurs d'en publier de nouveaux [79][80].
- Fournir un environnement Jupyter Notebook en ligne, avec accès gratuit aux GPU, pour développer et partager des modèles [79].
- Organiser des compétitions de machine learning, sponsorisées par des entreprises ou des institutions, où les participants peuvent soumettre leurs solutions [79][80].
- Offrir une communauté active d'échanges et de collaboration entre data scientists [79][80].
- Proposer des ressources d'apprentissage comme des tutoriels et des notebooks de référence [79][80].

En résumé, Kaggle est une plateforme essentielle pour le développement des compétences en science des données et en apprentissage automatique, à travers la compétition, la collaboration et l'apprentissage en communauté [79][80].

IV.2.2 Matériel de configuration

Nous présentons sur le tableau IV.1 les caractéristiques du matériel local utilisé.

	Machine 1	Machine 2
Fabricant	Hp	Hp
Type	HP EliteBook 840 G1	HP ProBook 640 G2
Processeur	Intel(R) Core(TM) i5-4300U CPU @ 1.90GHz 2.50 GHz	Intel® Core™ i5-6300U CPU @ 2.40GHz
GPU	Intel HD Graphics 4400	Intel HD Graphics 520
Disque dur	500 SSD	256 SSD
RAM	12 GB	8 GB
Système d'exploitation	Windows 10	Windows 10

TABLE IV.1 – Caractéristiques du matériel local utilisées.

- **Online**

Nous avons utilisé le GPU : **P100 de Kaggle**.

IV.3 Mesures de performance utilisées

Après avoir développé un modèle d'apprentissage automatique, l'étape suivante consiste à évaluer son efficacité à l'aide de diverses métriques de performance, en utilisant un ensemble de données de test distinct des données d'entraînement. Le choix des bonnes métriques de performance est crucial, car elles influencent grandement les résultats et les décisions prises sur la base du modèle.

IV.3.1 Matrice de Confusion

La matrice de confusion est un outil essentiel en apprentissage automatique supervisé qui permet d'évaluer la qualité d'un système de classification en comparant les prédictions du modèle aux valeurs réelles des données [81]. La matrice de confusion présente de manière structurée les prévisions du modèle, en mettant en évidence les prédictions correctes (vrais positifs et vrais négatifs) et les erreurs de classification (faux positifs et faux négatifs). Nous allons présenter sur la figure IV.1 un exemple de matrice de confusion pour un modèle de classification binaire :

		Classe réelle	
		-	+
Classe prédite	-	True Negatives <i>(vrais négatifs)</i>	False Negatives <i>(faux négatifs)</i>
	+	False Positives <i>(faux positifs)</i>	True Positives <i>(vrais positifs)</i>

FIGURE IV.1 – Matrice de Confusion.

Les quatre éléments composant la matrice de confusion sont [81] :

- Vrais positifs (TP) : les prédictions positives correctes.
- Faux positifs (FP) : les prédictions positives incorrectes.
- Vrais négatifs (TN) : les prédictions négatives correctes.
- Faux négatifs (FN) : les prédictions négatives incorrectes.

A cet effet, la matrice de confusion permet d'analyser en détail les performances du modèle, d'identifier les types d'erreurs les plus fréquents et d'ajuster les paramètres pour améliorer les résultats. Elle est largement utilisée dans les statistiques, le data mining, les modèles d'apprentissage automatique et d'autres applications de l'intelligence artificielle [81].

IV.3.2 Précision de la classification (Accuracy)

L'Accuracy (précision globale) est une métrique couramment utilisée pour évaluer les performances d'un modèle de classification. Elle représente le taux de prédictions correctes par rapport au nombre total d'échantillons dans l'ensemble de données. L'accuracy mesure la proportion de toutes les prédictions correctes (vrais positifs et vrais négatifs) par rapport à l'ensemble de données. L'accuracy peut être utilisée pour des ensembles de données équilibrés, où le nombre de classes positives et négatives est similaire. Cependant, elle peut ne pas être appropriée pour des ensembles de données déséquilibrés, où la classe minoritaire peut être ignorée par le modèle. Dans ce cas, il est souvent recommandé d'utiliser d'autres métriques, telles que la précision, le rappel ou le score F1, qui prennent en compte les performances par classe [82].

L'accuracy est une métrique utile pour évaluer les performances globales d'un modèle de classification, mais elle doit être utilisée avec précaution sur des données déséquilibrées [83].

• Précision et Rappel

La précision (precision) et le rappel (recall) sont deux métriques complémentaires utilisées pour évaluer les performances des systèmes de classification et de recherche d'information. La précision mesure la proportion d'éléments pertinents parmi ceux qui ont été sélectionnés par le système. Elle indique la fiabilité des prédictions positives du modèle. Le rappel mesure la proportion d'éléments pertinents qui ont été correctement identifiés par le système. Il reflète la capacité du modèle à détecter tous les vrais positifs [84]. Formules :

— Précision = Vrais positifs / (Vrais positifs + Faux positifs) [85]

— Rappel = Vrais positifs / (Vrais positifs + Faux négatifs) [85]

Ces métriques sont généralement calculées à partir de la matrice de confusion, qui compare les prédictions du modèle aux valeurs réelles [84].

IV.4 Les codes sources utilisés

Nous avons commencé ce projet par l'apprentissage du modèle SwinUnetR en vue d'une segmentation 3D complète du gliome. La figure IV.2 expose le code du modèle utilisé :

```
json_list = "/kaggle/working/dataset.json"
roi = (128, 128, 128)
batch_size = 1
sw_batch_size = 1
fold = 1
infer_overlap = 0.5
max_epochs = 50
val_every = 2
train_loader, val_loader = get_loader(batch_size, json_list, roi)

os.environ["CUDA_DEVICE_ORDER"] = "PCI_BUS_ID"
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")

model = SwinUNETR(
    img_size=roi,
    in_channels=4,
    out_channels=3,
    feature_size=48,
    drop_rate=0.0,
    attn_drop_rate=0.0,
    dropout_path_rate=0.0,
    use_checkpoint=True,
).to(device)
```

FIGURE IV.2 – Code source du modèle Transformer (SwinUnetR).

Nous allons détailler le modèle SwinUnetR comme suit :

- Le nombre de canaux d'entrée est de 4, représentant l'IRM multimédia avec (t2f, t2w, t1c, t1n)

- Le nombre de sorties est de 3 canaux (TC, WT, ET).
- Le GPU a été utilisé pour accélérer le processus.

• **ConvertToMultiChannelBasedOnBratsClassesd** : Cette transformation est conçue pour convertir les étiquettes en représentations multicanaux basées sur les classes de segmentation des tumeurs cérébrales BRATS 2021. Nous avons modifié le code pour l'adapter aux données BraTS 2023.

• Cette modification a été effectuée via le code source indiqué sur la figure IV.3 vu le changement apporté aux données BraTS 2023 en changeant le nombre d'étiquettes où la valeur de classification pour ET sera « 3 » (par rapport à « 4 » dans les défis BraTS précédents) [86].

```
class ConvertToMultiChannelBasedOnBratsClassesd(MapTransform):
    """
    Convert labels to multi channels based on brats 2023 classes:
    label 1 necrotic tumor core (NCR)
    label 2 peritumoral edematous/invaded tissue
    label 3 GD-enhancing tumor
    label 4 is not used in the new dataset version
    The possible classes are:
    TC (Tumor core): merge labels 1 and 3
    WT (Whole tumor): merge labels 1,2 and 3
    ET (Enhancing tumor): label 3
    """
    def __call__(self, data):
        d = dict(data)
        for key in self.keys:
            result = []
            # merge label 1 and label 3 to construct TC
            result.append(torch.logical_or(d[key] == 1, d[key] == 3))
            # merge labels 1, 2 and 3 to construct WT
            result.append(
                torch.logical_or(
                    torch.logical_or(d[key] == 1, d[key] == 2), d[key] == 3
                )
            )
            # Label 3 is ET
            result.append(d[key] == 3)
            d[key] = torch.stack(result, axis=0).astype(np.float32)
        return d
```

FIGURE IV.3 – Code source du ConvertToMultiChannelBasedOnBratsClassesd

• Nous avons utilisé l'optimiseur **AdamW** qui se distingue des autres optimiseurs car il produit une meilleure perte d'entraînement et les modèles se généralisent beaucoup mieux que les modèles formés avec Adam [87] comme indiqué sur la figure IV.4.

```
optimizer = torch.optim.AdamW(model.parameters(), lr=1e-4, weight_decay=1e-5)
```

FIGURE IV.4 – Utilisation de l'optimiseur AdamW.

• Quant au processus du **Skull stripping**, nous avons utilisé le programme indiqué sur la figure IV.5.

```
for i in range(img_data.shape[2]): # Assuming the slice dimension is the third dimension
    filtered_slices[:, :, i] = gaussian_filter(img_data[:, :, i], 0.2)
    slice_data = filtered_slices[:, :, i]
    threshold_value=1
    slice_data_uint8 = cv2.normalize(slice_data, None, 0, 255, cv2.NORM_MINMAX, cv2.CV_8U)
    # Skull stripping
    gray = slice_data_uint8
    ret, thresh = cv2.threshold(gray, 0, 255, cv2.THRESH_OTSU)
    if ret > 0:
        colormask = np.zeros_like(slice_data_uint8, dtype=np.uint8)
        colormask[thresh != 0] = 255
        colormask_bgr = cv2.cvtColor(colormask, cv2.COLOR_GRAY2BGR)
        colormask_bgr[thresh == 0] = [0, 0, 0]
        blended = cv2.addWeighted(cv2.cvtColor(slice_data_uint8, cv2.COLOR_GRAY2BGR), 0.7, colormask_bgr, 0.1, 0)
        ret, markers = cv2.connectedComponents(thresh)
        marker_area = [np.sum(markers == m) for m in range(1, np.max(markers) + 1)]
        if not marker_area:
            skull_stripped_slices[:, :, i] = filtered_slices[:, :, i]
            clahe_slices[:, :, i] = filtered_slices[:, :, i]
        else:
            largest_component = np.argmax(marker_area) + 1
            brain_mask = markers == largest_component
            brain_out = slice_data.copy()
            brain_out[brain_mask == False] = 0
            center_x, center_y = np.array(brain_out.shape) // 2
            half_roi_size = roi_size // 2
            roi_start_x = max(0, center_x - half_roi_size)
            roi_end_x = min(brain_out.shape[0], center_x + half_roi_size)
            roi_start_y = max(0, center_y - half_roi_size)
            roi_end_y = min(brain_out.shape[1], center_y + half_roi_size)
            skull_stripped_roi = brain_out[roi_start_x:roi_end_x, roi_start_y:roi_end_y]
            # Append the skull-stripped ROI to the list
            skull_stripped_slices[:, :, i] = skull_stripped_roi
            slice_data_norm = cv2.normalize(skull_stripped_roi, None, 0, 255, cv2.NORM_MINMAX, dtype=cv2.CV_8U)
```

FIGURE IV.5 – Code source du Skull stripping.

Nous définissons le seuil à l'aide de `cv2.THRESH_OTSU` pour déterminer les pixels blancs et noirs de l'image. Il n'en résulte que deux valeurs différentes et le masque est créé. Nous appliquons ensuite `cv2.connectedComponents` pour déterminer les composantes connectées dans le masque et rejetons l'étiquette 0 désignant l'arrière-plan.

Par la suite, la région d'intérêt centrale est extraite du cerveau et ajouté à `skull_stripped_slices`. Nous présentons sur la figure IV.6 un exemple d'exécution du programme sur une IRM cérébrale pour l'effectuation du skull striping.

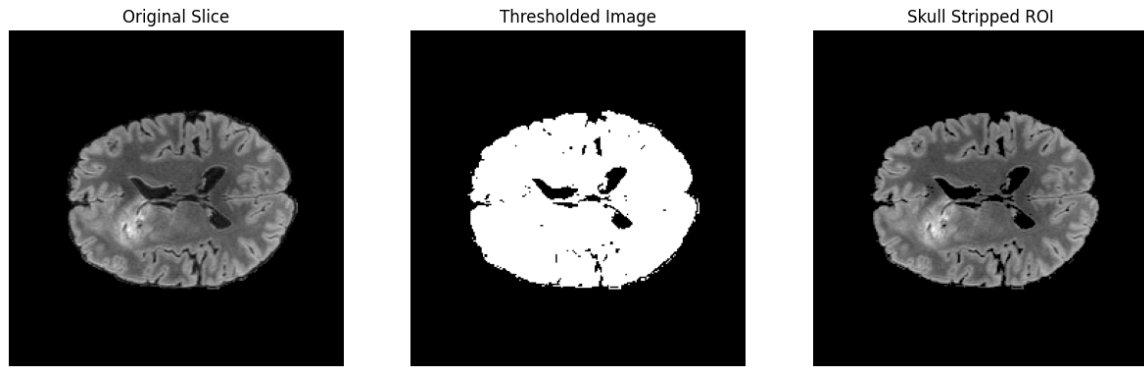


FIGURE IV.6 – Exemple du résultat généré par skull stripping sur une image IRM cérébrale.

IV.5 Evaluation des performances de l'approche de segmentation et de classification du gliome basée SwinUnetR

IV.5.1 Evaluation des performances de la segmentation

La figure IV.7 illustre la minimisation de la fonction d'Average Loss d'apprentissage au fil des époques :

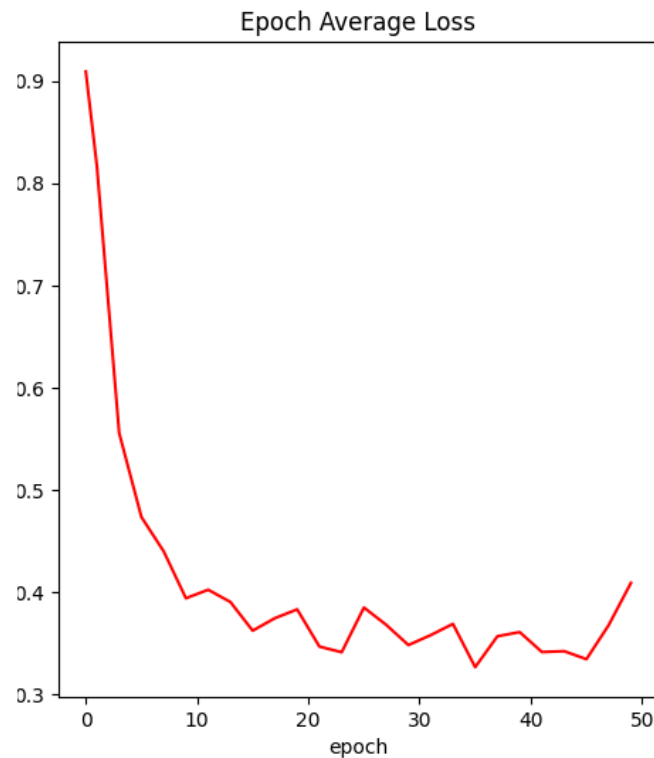


FIGURE IV.7 – Evolution de la fonction d'Average Loss au long des époques par SwinUnetR.

La figure IV.8 illustre l'évolution de la fonction de Val Mean Dice d'apprentissage au fil des époques.

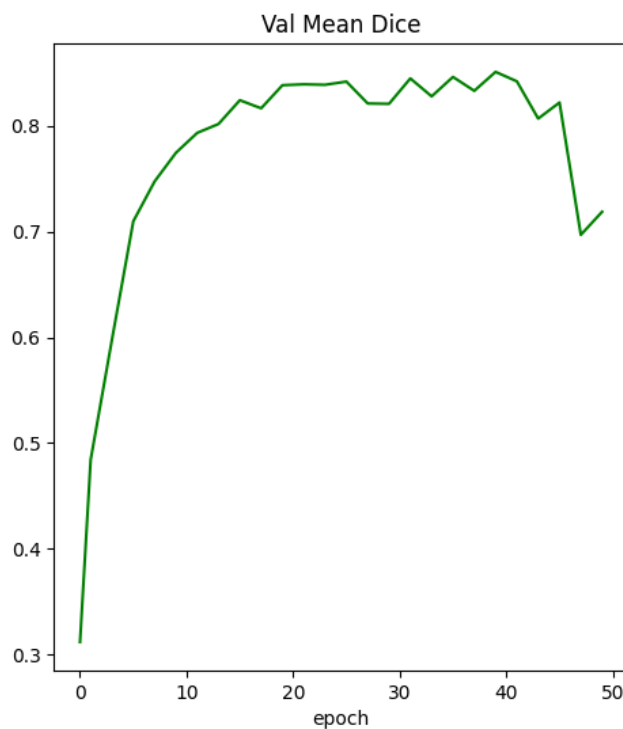


FIGURE IV.8 – Evolution de la fonction de Val Mean Dice au long des époques.

La figure IV.9 illustre l'évolution de la fonction de Val Mean Dice TC d'apprentissage au fil des époques :

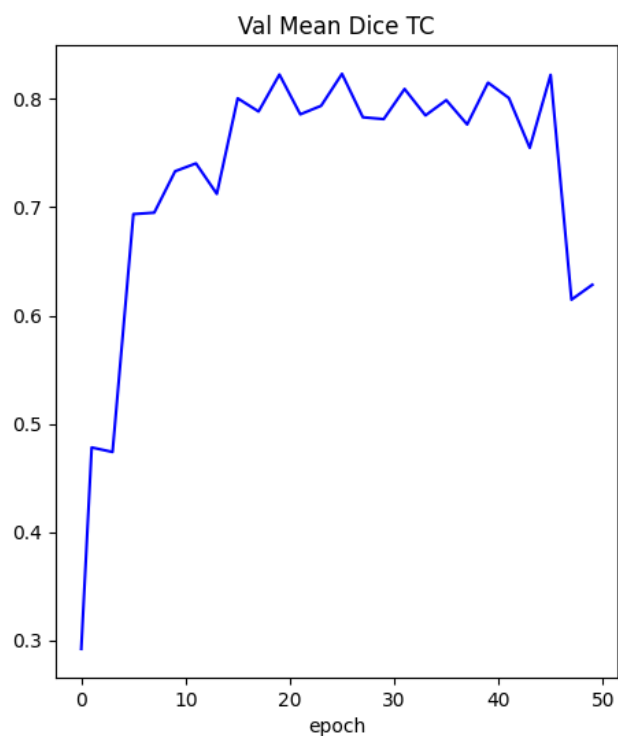


FIGURE IV.9 – Evolution de la fonction de Val Mean Dice TC au long des époques.

La figure IV.10 illustre l'évolution de la fonction de Val Mean Dice WT d'apprentissage au fil des époques :

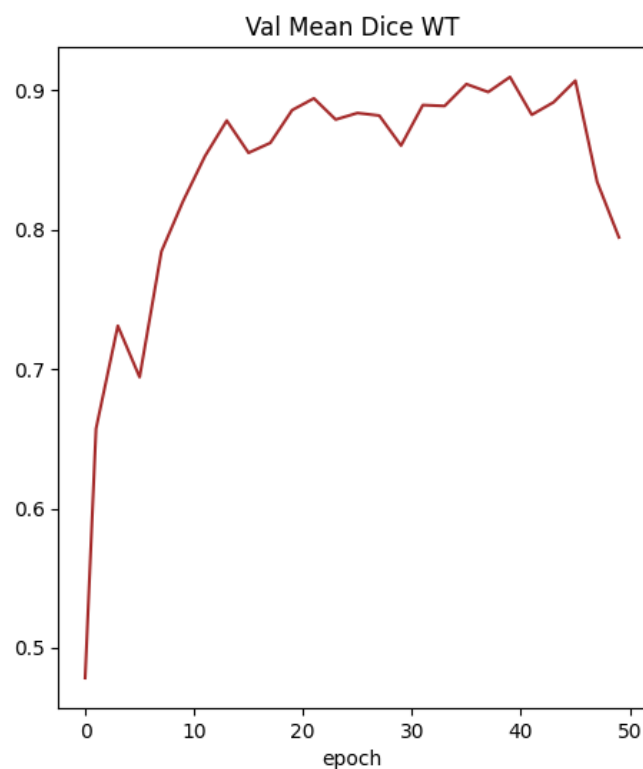


FIGURE IV.10 – Evolution de la fonction de Val Mean Dice WT au long des époques.

La figure IV.11 illustre l'évolution de la fonction de Val Mean Dice ET d'apprentissage au fil des époques :

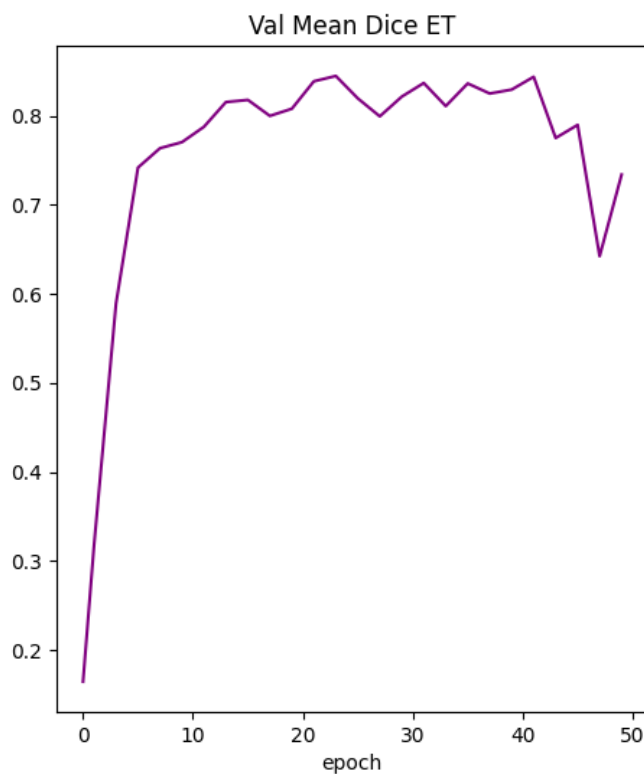
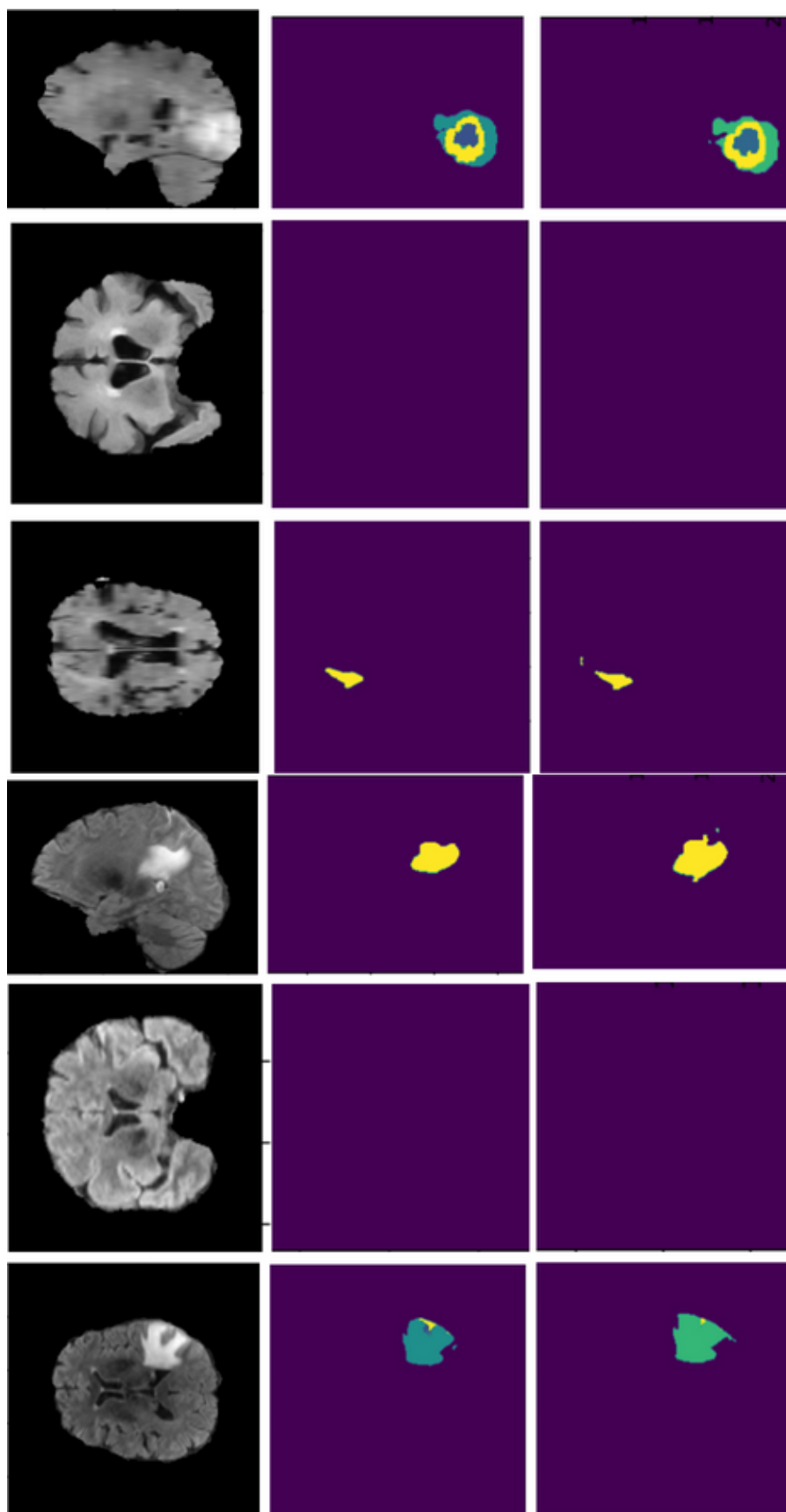
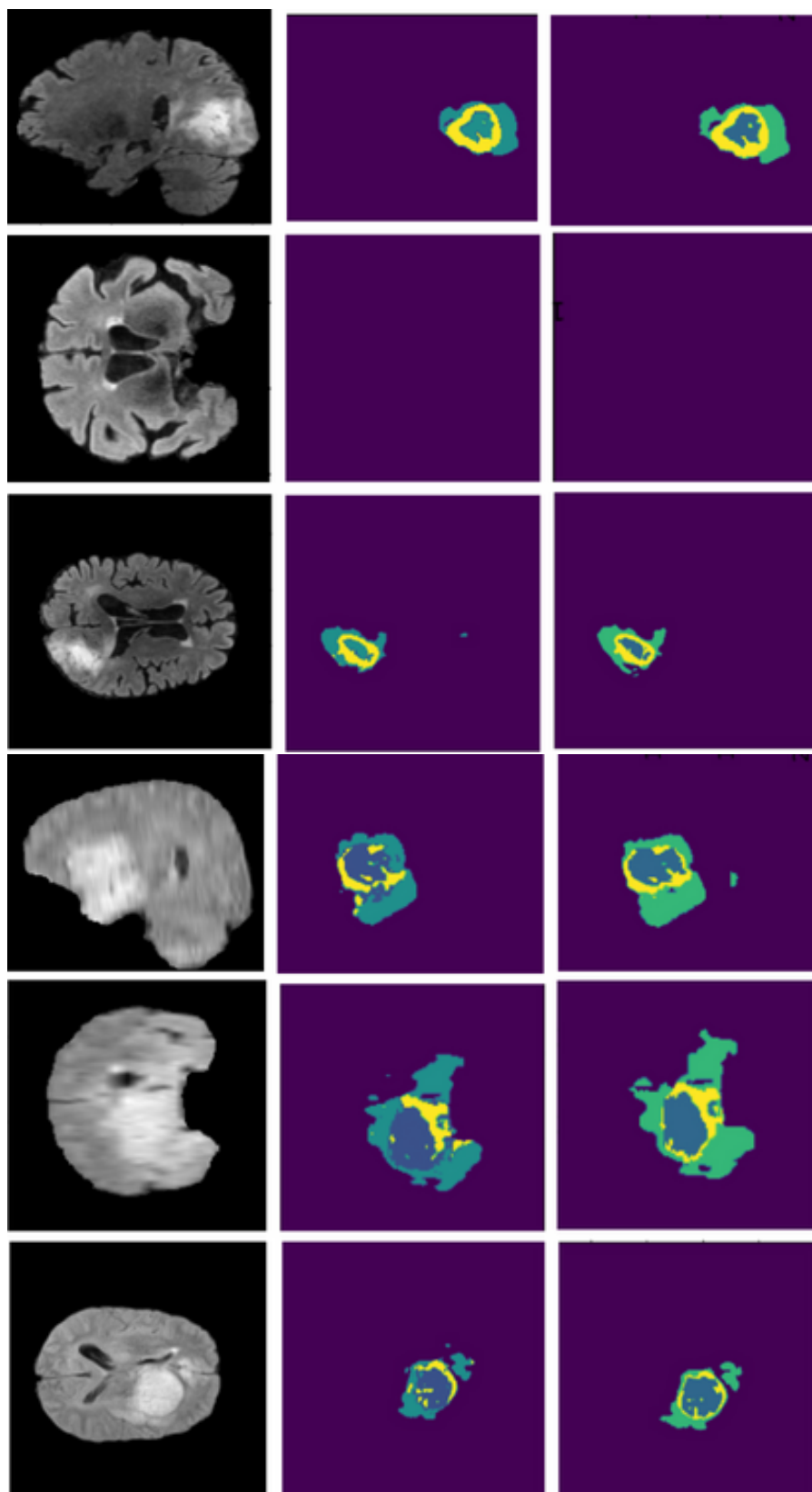


FIGURE IV.11 – Evolution de la fonction de Val Mean Dice ET au long des époques.

Nous présentons sur la figure IV.12 quelques exemples de résultats de la segmentation 3D basées SwinUnetR du gliome comme suit :





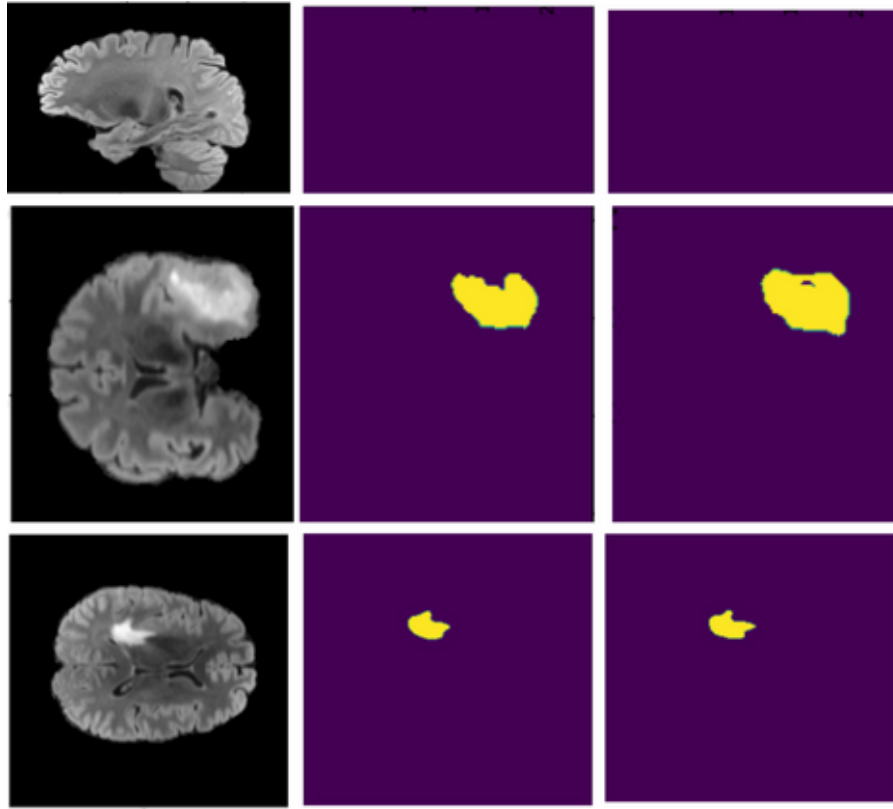


FIGURE IV.12 – Exemples des résultats de la segmentation 3D du gliome basée Swin UnetR.

Comme mentionné sur les sections de la problématique ainsi que du plan d'action, l'ultime but du choix de SwinUNETR était porté sur la réduction du coût de calcul des Transformers par l'utilisation de SwinUNETR qui s'appuie sur une attention basée sur la répartition en fenêtres. De plus, comme mentionné sur le tableau IV.2 le SwinUNETR surpasse diverses architectures de Transformers en matière de segmentation du gliome.

Model	Dice score %			Dataset
	TC	WT	ET	
TransBTS [88]	81.73	90.09	78.73	BraTS 2020
MMCFormer [17]	87.4	89	80.1	BraTS 2018
SwinBTS [89]	80.30	89.06	77.36	BraTS 2020
Swin UnetR	81.51	90.96	82.93	BraTS 2023

TABLE IV.2 – Comparaison entre diverses architectures basées Transformer pour la segmentation du gliome

Nous avons réalisé le cropping des tumeurs comme mentionné sur la figure IV.13 grâce

au modèle obtenu ainsi qu'aux masques générés par ce modèle. De plus, l'arrière-plan est épargné à l'aide de Rembg désignant une bibliothèque Python à code source ouvert spécialement conçue pour supprimer les arrière-plans des images. Elle utilise un réseau neuronal profond pour identifier et supprimer avec précision l'arrière-plan des images [90].

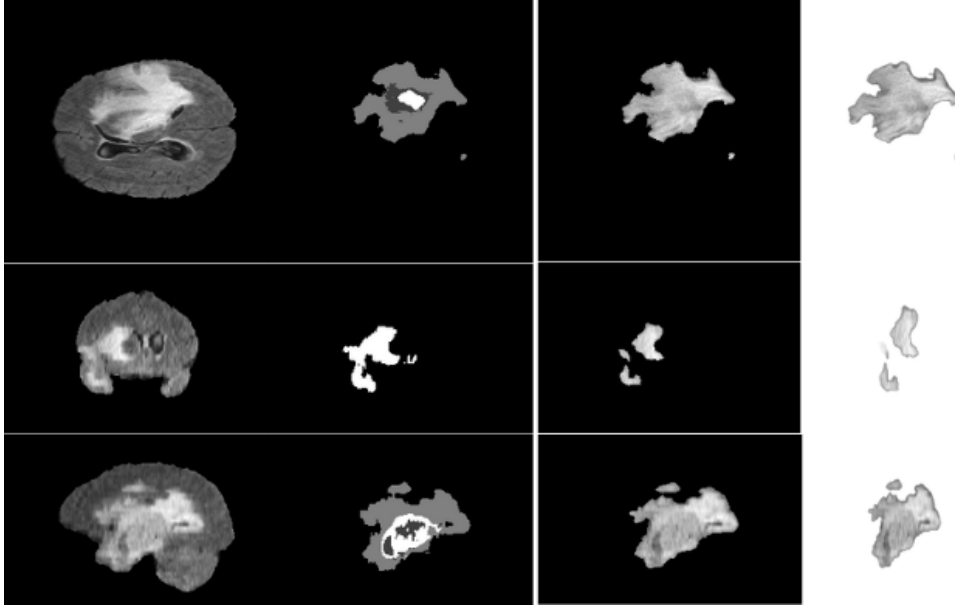


FIGURE IV.13 – Exemple du processus de cropping.

IV.5.2 Evaluation des performances de la classification

Nous allons maintenant appliquer les techniques d'extraction de caractéristiques que nous avons évoqués dans le chapitre précédent, à savoir ResNet50, Gabor, Zipf, Zipf inverse et DWT. En raison des données BraTS 2023 non étiquetées pour le processus de classification, il nous a été possible de l'utiliser uniquement durant la segmentation vu la disponibilité des masques. Nous avons eu recours à l'extraction de caractéristiques et la classification de la base de données BraTS 2020 étiquetée.

Une fois le processus d'extraction de descripteur est réalisé, nous passons à l'étape de l'apprentissage du modèle de classification SVM. Nous définissons ensuite les hyperparamètres suivants : Kernel (linéaire ou fonction de base radiale), valeur du paramètre de régularisation C (0.1, 1, 10, 100) et gamma (scale, auto), puis nous avons utilisé GridSearchCV et entraîné le modèle sur les meilleurs paramètres pour discriminer les classes de gliomes LGG et HGG. Nous présentons sur le tableau IV.3 les résultats de classification obtenus :

Approches	C	Gamma	Kernel	Accuracy
ResNet50	10	scale	Rbf	0.90
G_Z_ZI	10	scale	Rbf	0.8649
DWT	100	scale	Rbf	0.93

TABLE IV.3 – Précision de classification du gliome en HGG et LGG basée SwinUnetR et SVM.

Nous avons également utilisé l'analyse en composantes principales (ACP), qui est l'une des techniques permettant d'améliorer les performances et de réduire la dimensionnalité de l'ensemble de données lors de l'extraction de caractéristiques. Les résultats obtenus sont mentionnés dans le tableau IV.4.

Approches	C	Gamma	Kernel	Accuracy
ResNet50	100	scale	Rbf	0.91
G_Z_ZI	10	scale	Rbf	0.8604
DWT	100	scale	Rbf	0.92

TABLE IV.4 – Précision de classification du gliome en HGG et LGG basée SwinUnetR et SVM ainsi que l'incorporation de l'ACP.

Nous avons également utilisé un modèle de réseaux de neurones artificiels (ANN) pour la classification, dont les résultats obtenus sont mentionnés sur le tableau IV.5.

Approches	Accuracy	Accuracy
	(sans ACP)	(avec ACP)
ResNet50	0.90	0.91
G_Z_ZI	0.88	0.88
DWT	0.90	0.9414

TABLE IV.5 – Précision de classification du gliome en HGG et LGG basée SwinUnetR et ANN.

Pour davantage de détails sur la meilleure précision de classification atteinte ...

	precision	recall	f1-score	support
0	0.95	0.84	0.89	63
1	0.94	0.98	0.96	159
accuracy			0.94	222
macro avg	0.94	0.91	0.93	222
weighted avg	0.94	0.94	0.94	222

Pour finir l'évaluation des performances de la classification il est pivot de tracer les matrices de confusion générées comme mentionné sur la tableau IV.6 et IV.7 :

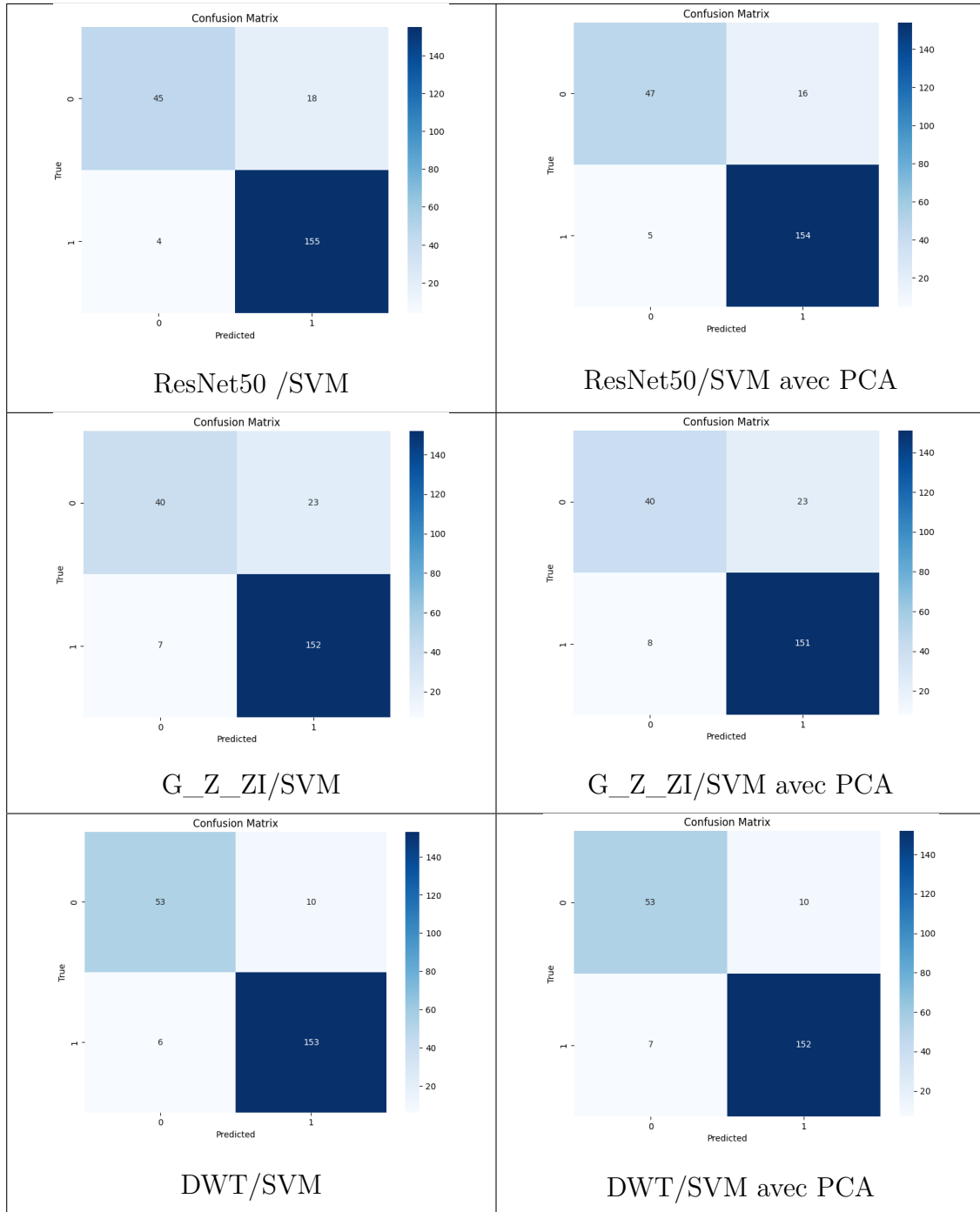


TABLE IV.6 – les matrices de confusion relatives à la segmentation basée SwinUnetR et la classification du gliome en HGG et LGG basée SVM.

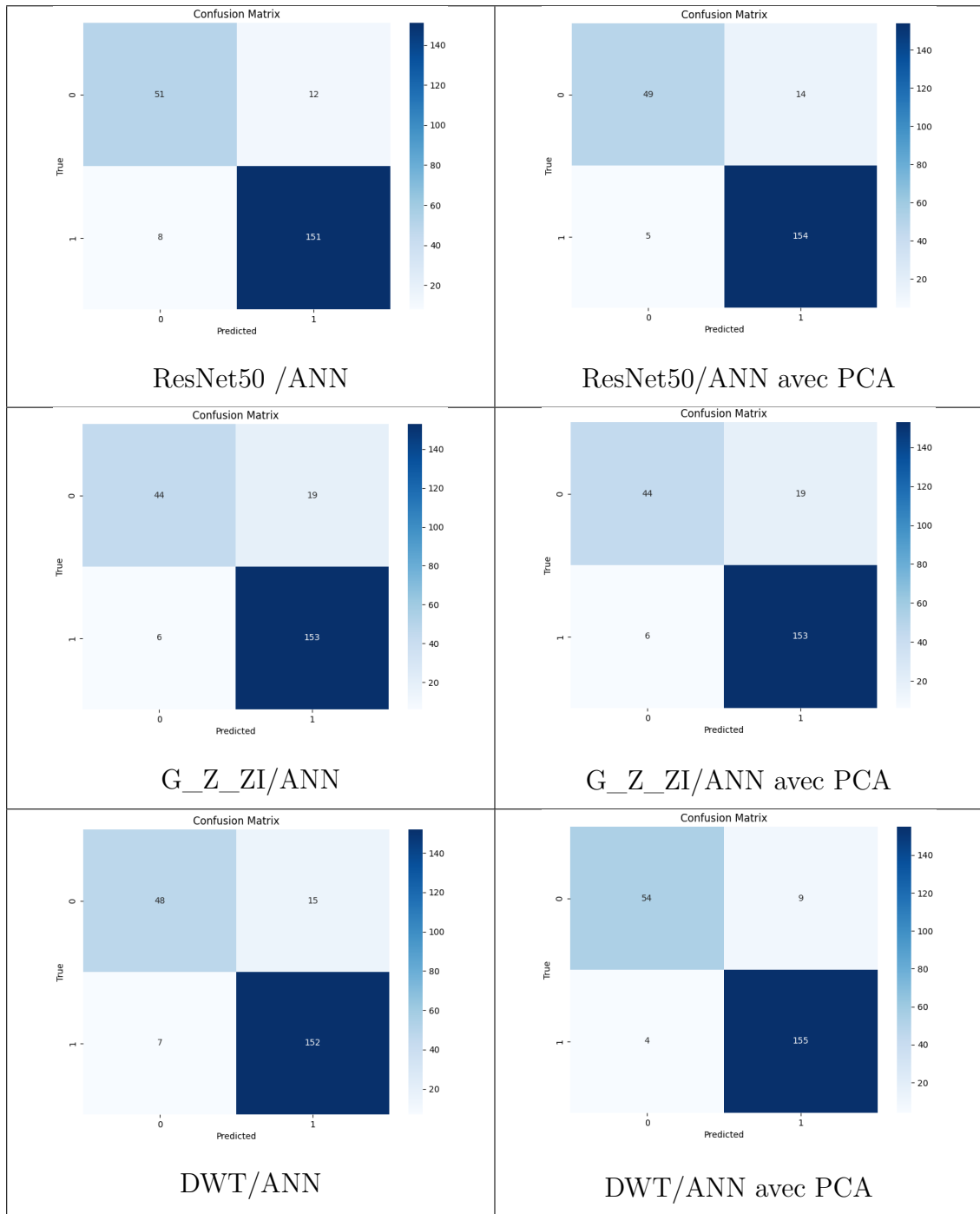


TABLE IV.7 – les matrices de confusion relatives à la segmentation basée SwinUnetR et la classification du gliome en HGG et LGG basée ANN.

IV.6 Evaluation des performances de l'approche de segmentation et de classification du gliome basée SegResNet

IV.6.1 Evaluation des performances de la segmentation

La figure IV.14 illustre la minimisation de loss function au fil des époques par le modèle SegResNet.

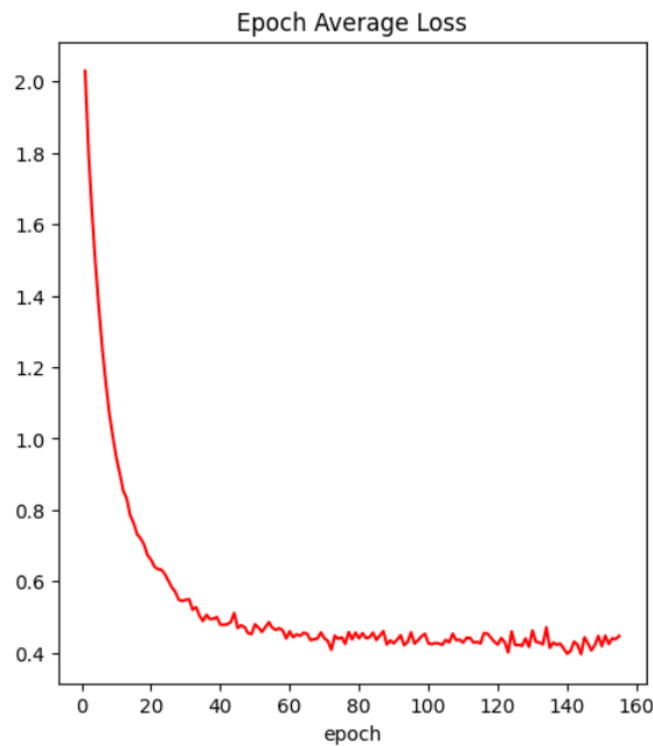


FIGURE IV.14 – Minimisation de loss function par le modèle SegResNet.

Nous exposons sur la figure IV.15 l'évolution de la fonction de Val Mean Dice au long des époques en utilisant le modèle SegResNet.

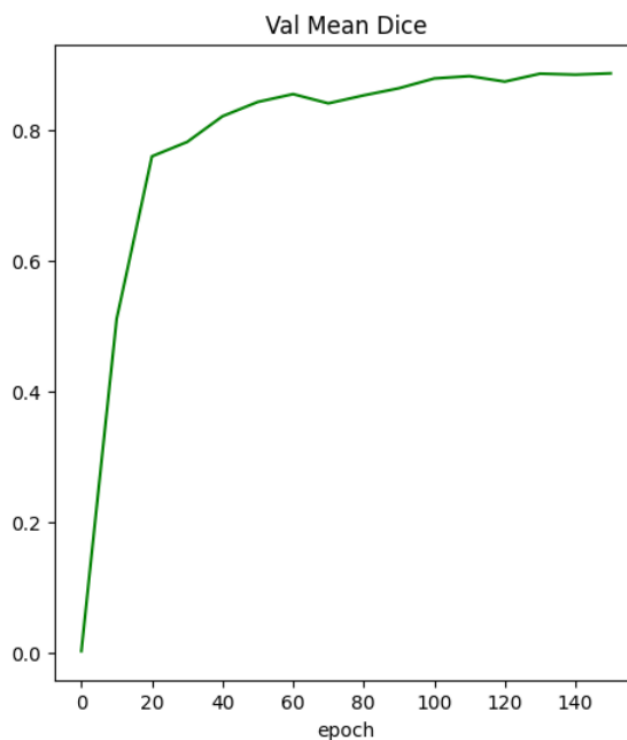


FIGURE IV.15 – Evolution de la fonction de Val Mean Dice au long des époques en utilisant le modèle SegResNet.

La figure IV.16 illustre l'évolution de la fonction de Val Mean Dice TC, WT et ET d'apprentissage au fil des époques.

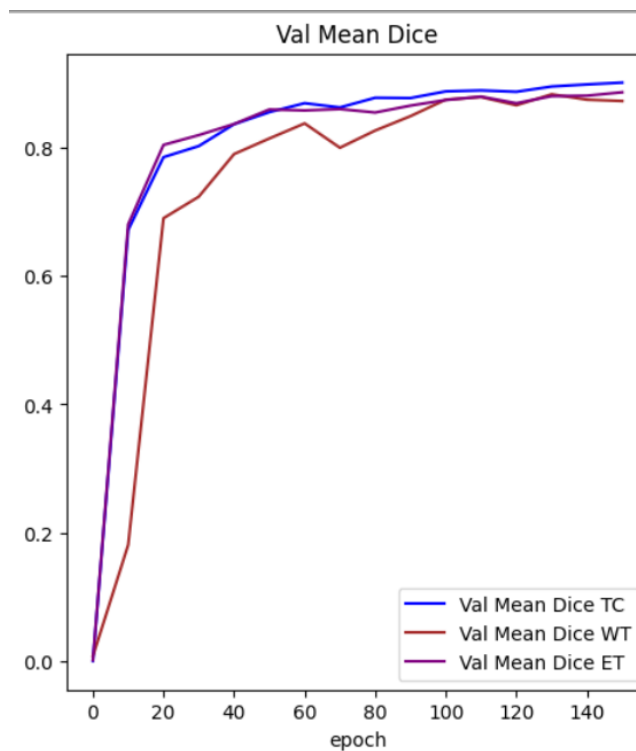
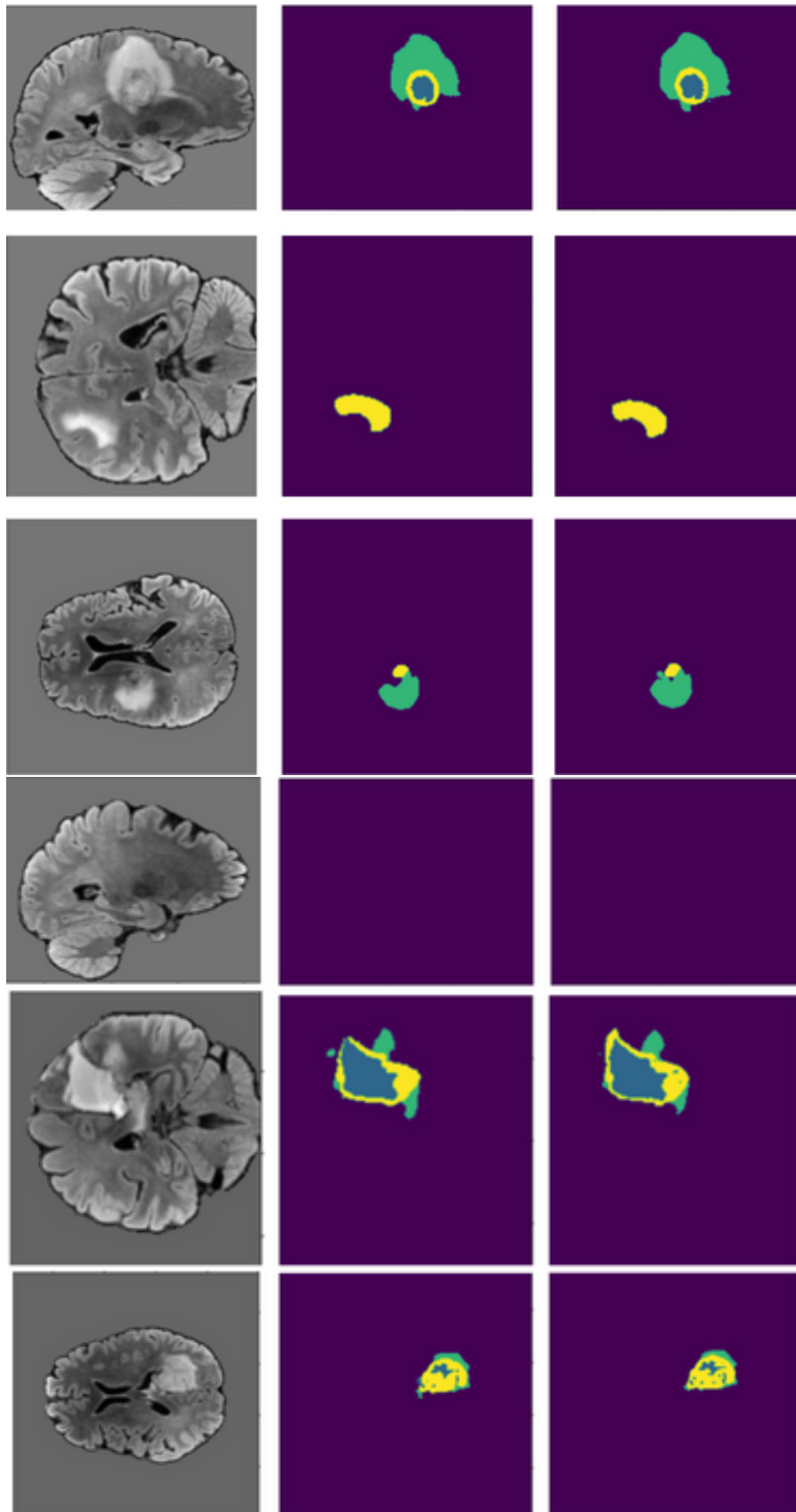


FIGURE IV.16 – Evolution de la fonction de Val Mean Dice TC, WT et ET au long des époques.

Nous présentons sur la figure IV.17 quelques exemples de résultats de la segmentation 3D basées SegResNet du gliome comme suit :



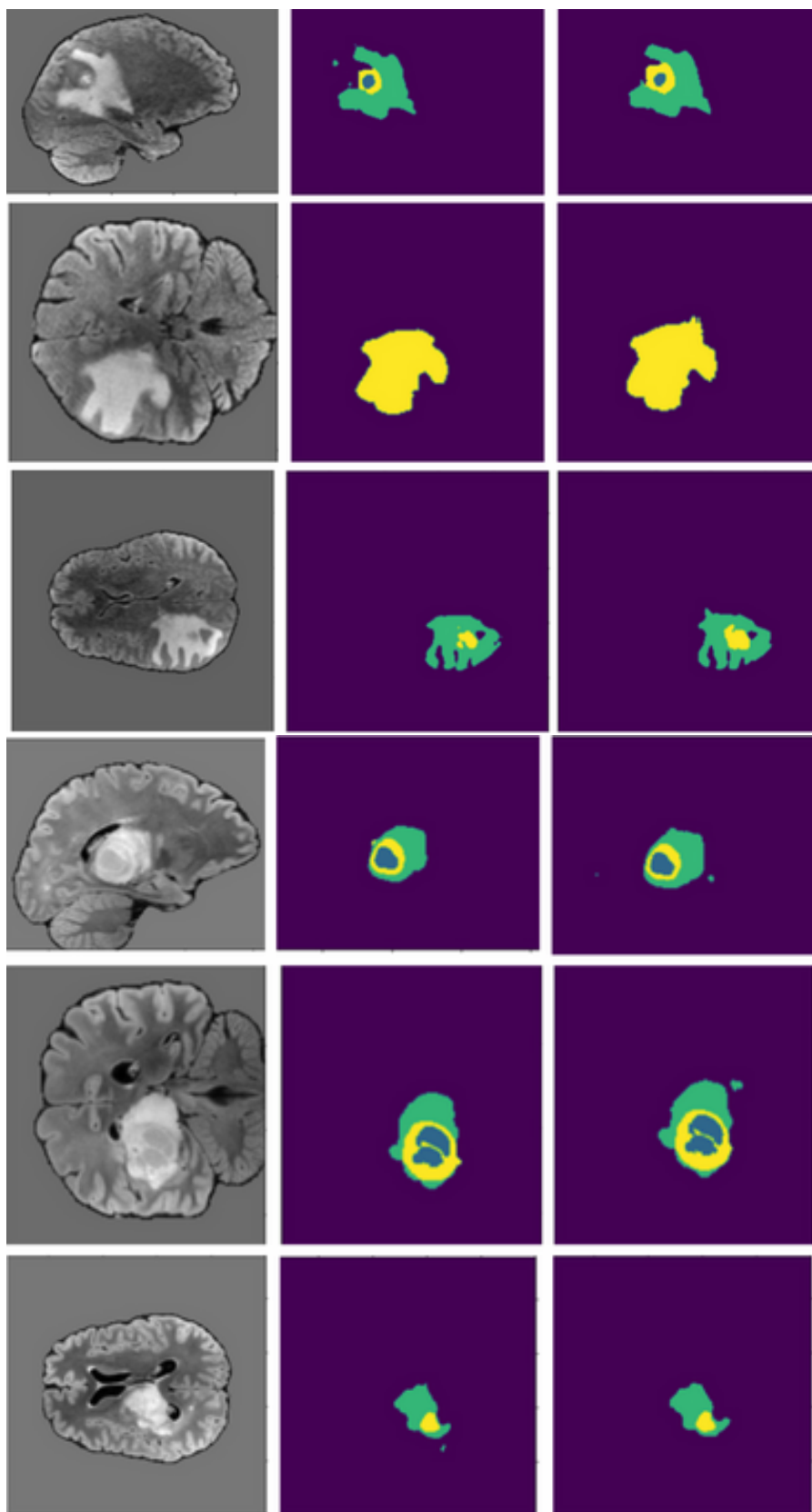


FIGURE IV.17 – Exemple de la segmentation du gliome basée SegResNet.

IV.6.2 Evaluation des performances de la classification.

Nous présentons sur le tableau [IV.8](#) les résultats obtenus de classification.

Approches	C	Gamma	Kernel	Accuracy
ResNet50	100	scale	Rbf	0.90
G_Z_ZI	10	scale	Rbf	0.8694
DWT	10	scale	Rbf	0.91

TABLE IV.8 – Précision de classification du gliome en HGG et LGG par les SVM et issue d'une segmentation basée SegResNet.

Nous présentons sur le tableau IV.9 les résultats de la classification par les SVM avec l'ACP.

Approches	C	Gamma	Kernel	Accuracy
ResNet50	10	scale	Rbf	0.91
G_Z_ZI	100	auto	Rbf	0.81
DWT	10	scale	Rbf	0.91

TABLE IV.9 – Précision de classification du gliome en HGG et LGG par les SVM et issue d'une segmentation basée SegResNet avec l'incorporation de l'ACP.

Nous avons également utilisé un modèle de réseaux de neurones artificiels (ANN) pour la classification, dont les résultats obtenus sont mentionnés sur le tableau IV.10.

Approches	Accuracy	Accuracy
	(sans ACP)	(avec ACP)
ResNet50	0.83	0.92
G_Z_ZI	0.87	0.88
DWT	0.89	0.91

TABLE IV.10 – Précision de classification du gliome en HGG et LGG basée SwinUnetR et ANN.

Pour davantage de détails concernant les performances de la classification ...

	precision	recall	f1-score	support
0	0.88	0.84	0.86	63
1	0.94	0.96	0.95	159
accuracy			0.92	222
macro avg	0.91	0.90	0.90	222
weighted avg	0.92	0.92	0.92	222

Quant aux matrices de confusion obtenues, nous les exposons sur les figures IV.11 et IV.12.

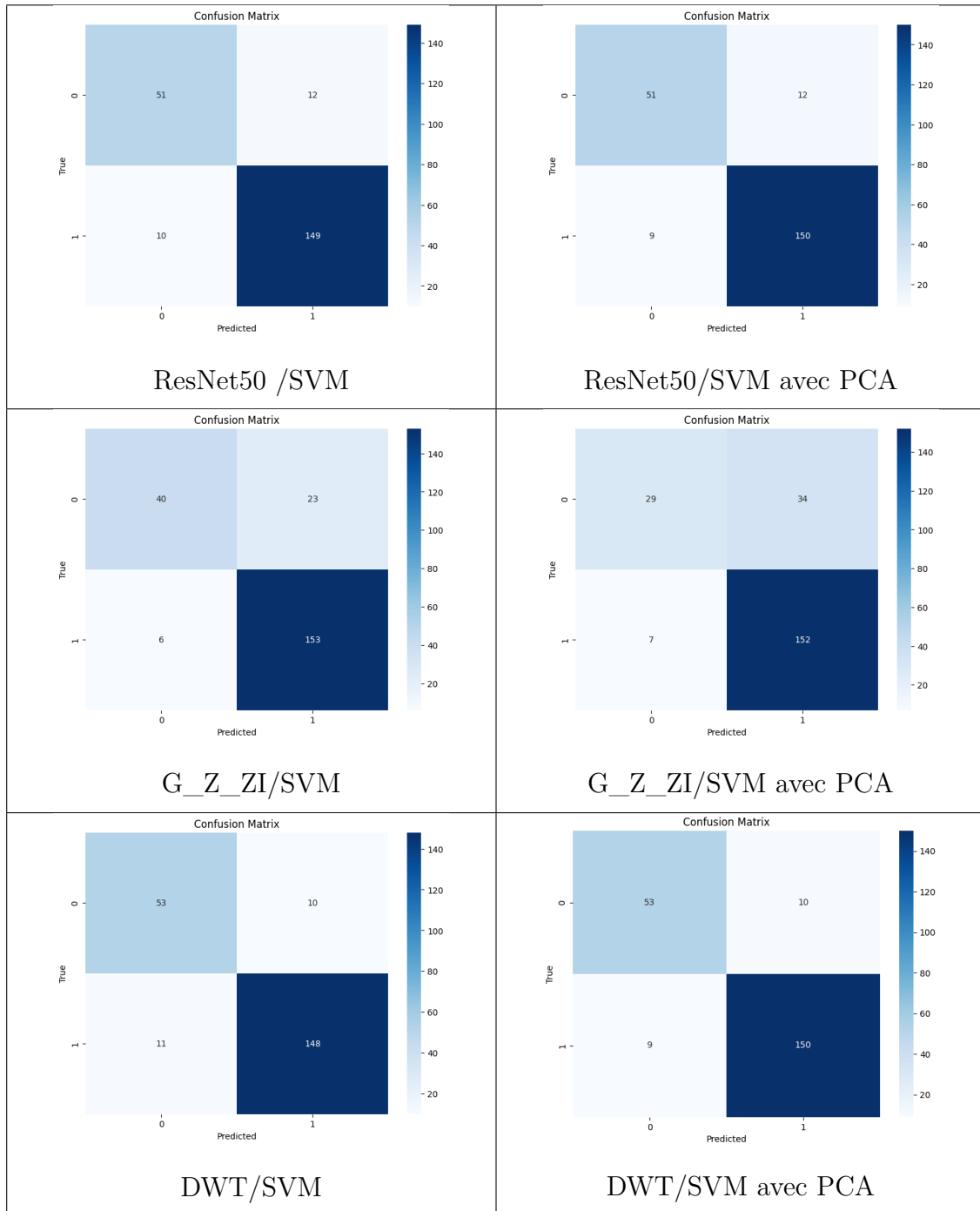


TABLE IV.11 – Les matrices de confusion relatives à la classification du gliome en HGG et LGG basée SVM et issue d'une segmentation basée SegResNet.

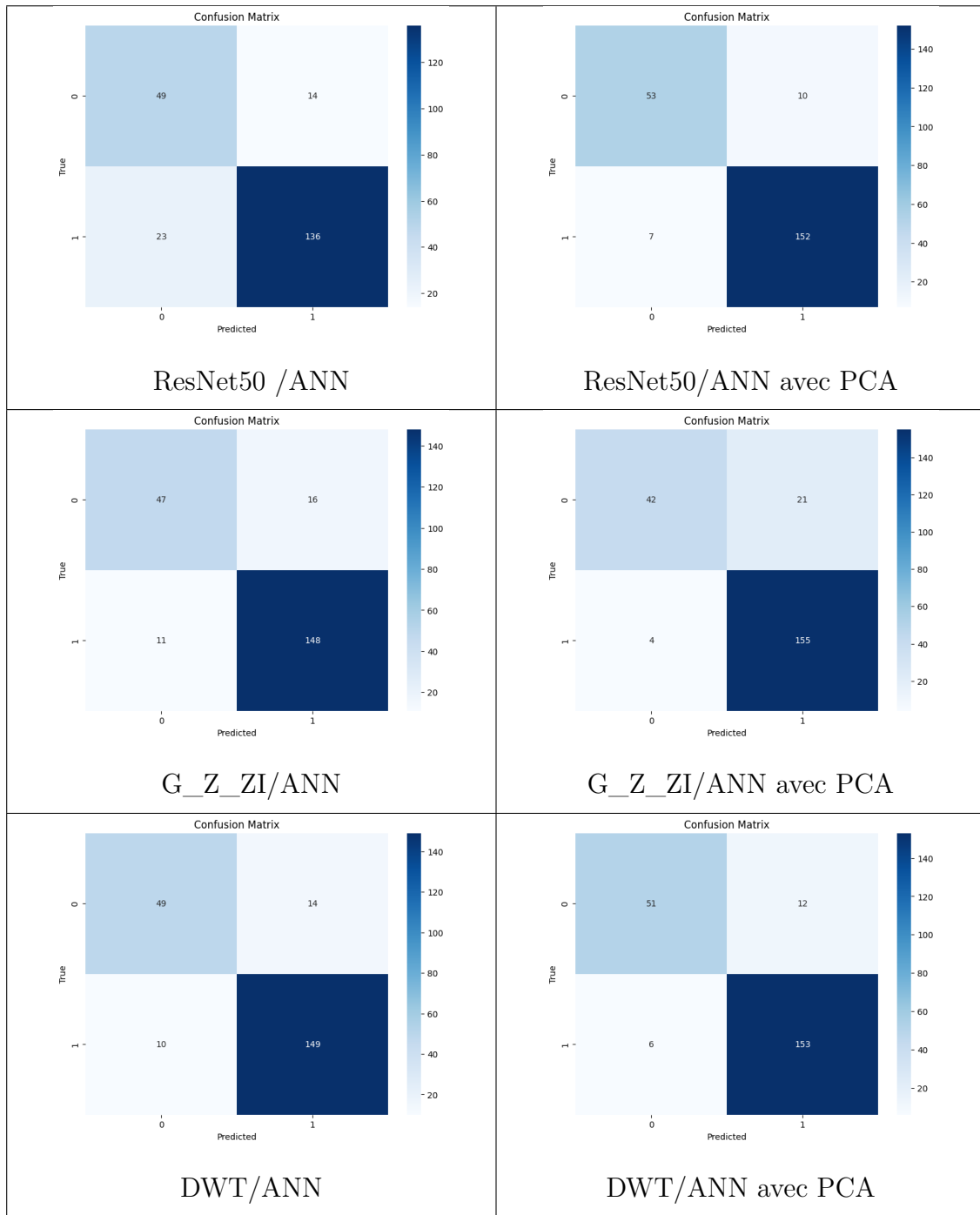


TABLE IV.12 – Les matrices de confusion relatives à la classification du gliome en HGG et LGG basée ANN et issues d'une segmentation basée SegResNet.

IV.7 Comparaison des performances de notre approches vs d'autres approches

Nous allons comparer (tableau IV.13) l'approche proposée basée SwinUnetR implémentant les quatre modalités de l'IRM cérébrale et effectuant une segmentation en 3D avec une approches basée SwinUnetR implémentant une seule modalité de l'IRM cérébrale (FLAIR) et effectuant

une segmentation en 2D. En outre, nous allons aussi comparer les résultats obtenus avec un CNN (SegResNet).

Approches	Accuracy	Dataset
SwinUnetR (IRM multimodales + 3D)	94.14%	BraTs 2023
SwinUnetR [91] (IRM FLAIR + 2D)	91%	BraTs 2021
SegResNet (IRM multimodales + 3D)	92%	BraTs 2023

TABLE IV.13 – Comparaison des performances de notre approches vs d'autres approches

Nous pouvons affirmer que l'approche proposée basée SwinUnetR surpasse les autres approches. En effet, il est très important de considérer les quatre modalités pour la segmentation de l'IRM cérébrale ainsi que les descripteurs des Ondelettes qui a abouties aux meilleures précisions de classification. Il est utile de mentionner aussi que les CNN restent puissants aussi pour la segmentation des IRM cérébrales comme l'a affirmé le modèle SegResNet implémenté.

La technique d'auto-attention adoptée par les Transformer les distingue des autres modèles d'IA. Les modèles basés sur le CNN ne peuvent pas tirer parti des informations contextuelles à longue portée de la tumeur en raison du champ réceptif limité, dans ce sens, l'un des avantages des Transformers est que leur poids dynamique peut mieux s'adapter aux caractéristiques d'entrée, ce qui est l'une des raisons pour lesquelles fusionner les Transformers avec les modèles CNN abouti aux meilleures performances. En outre, l'analyse locale des CNN fusionnée avec l'analyse d'aspect global des Transformer est sans doute bénéfique comme il a été prouvé dans ce travail.

IV.8 Présentation de l'interface de l'approche proposée

IV.8.1 L'interface du système de pronostic automatique du gliome basé SegResNet

Nous présentons sur la figure IV.18 une vue globale de l'interface.

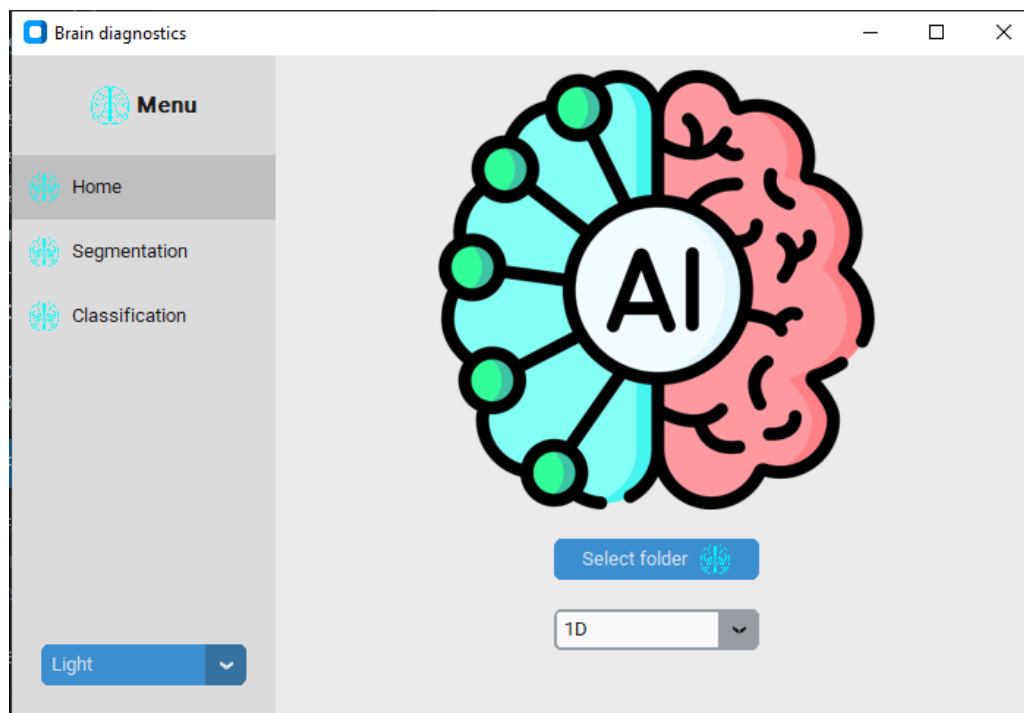


FIGURE IV.18 – Vue globale de l'interface du système de pronostic automatique du gliome basé SegResNet.

Nous présentons sur la figure IV.19 Le choix d'une IRM cérébrale pour la réalisation du pronostic automatique.

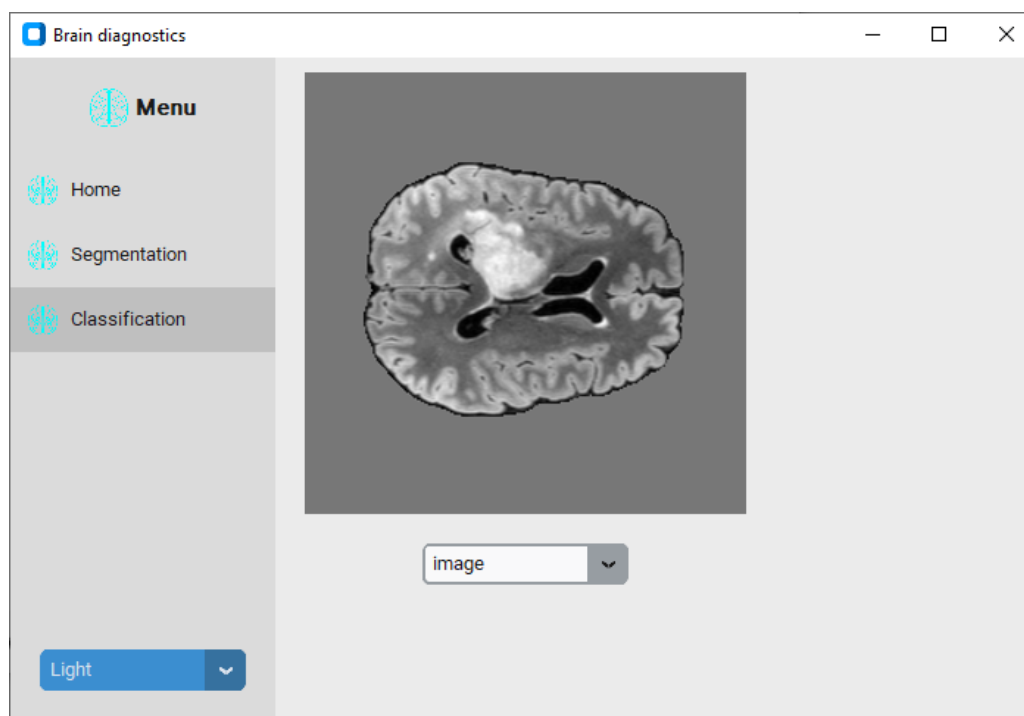


FIGURE IV.19 – Le choix d'une IRM cérébrale pour la réalisation du pronostic automatique.

La figure IV.20 expose un exemple de segmentation d'une IRM cérébrale.

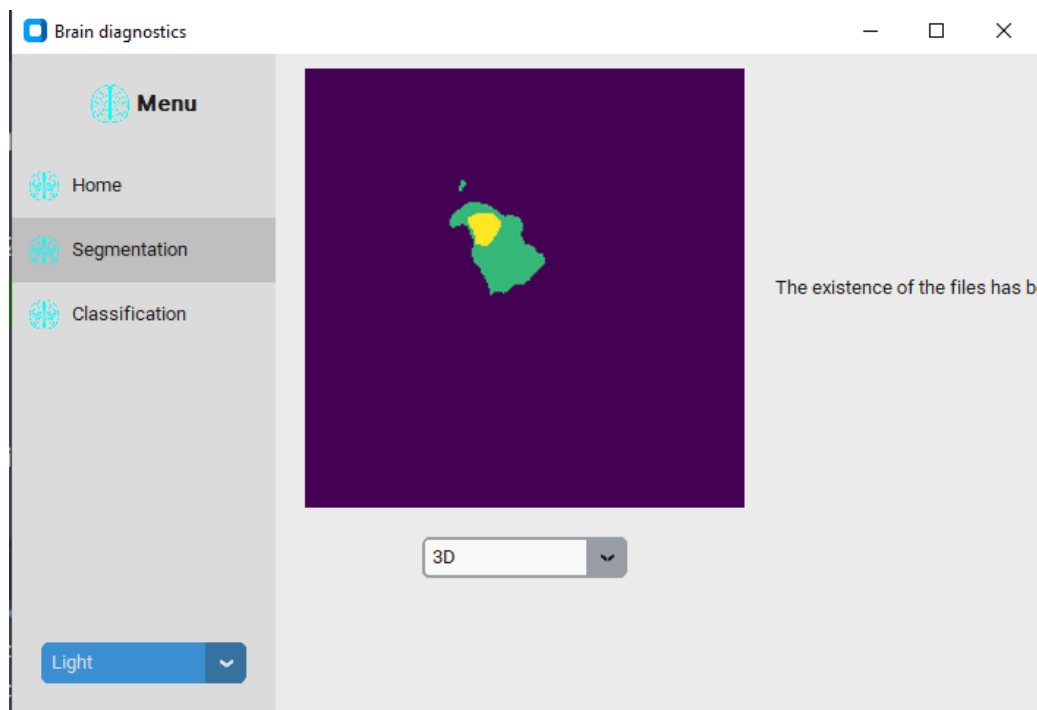


FIGURE IV.20 – Un exemple de segmentation d'une IRM cérébrale.

La figure IV.21 montre un exemple de cropping de tumeur d'une IRM cérébrale.

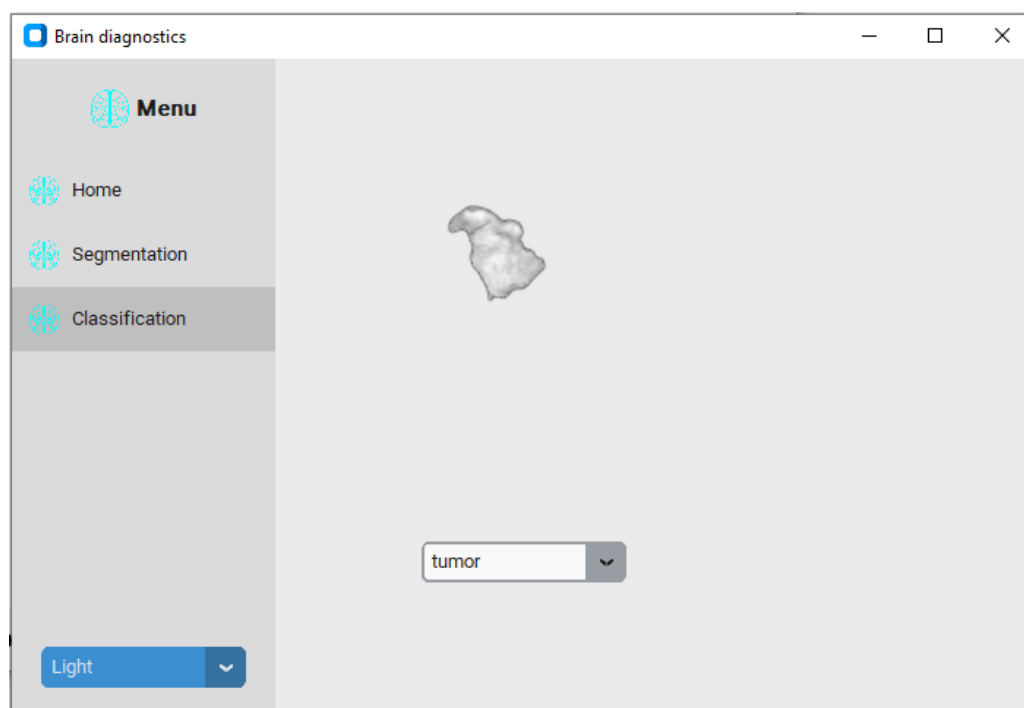


FIGURE IV.21 – Un exemple de cropping de tumeur d'une IRM cérébrale.

Quant à la figure IV.22, elle démontre le résultat de la classification de la tumeur segmentée.

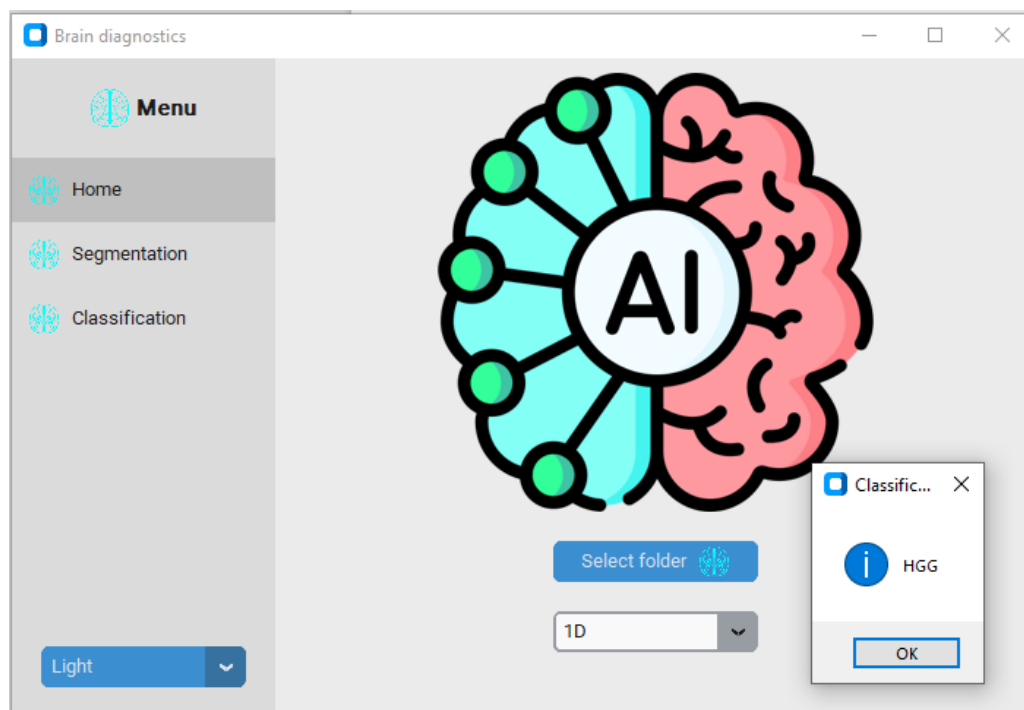


FIGURE IV.22 – Pronostic de la tumeur gliome segmentée qui s'est avéré un gliome HGG.

Nous présentons sur la figure IV.23 l'affichage de la tumeur gliome segmentée en 3D qui permettra à un radiologue d'observer soigneusement la tumeur.

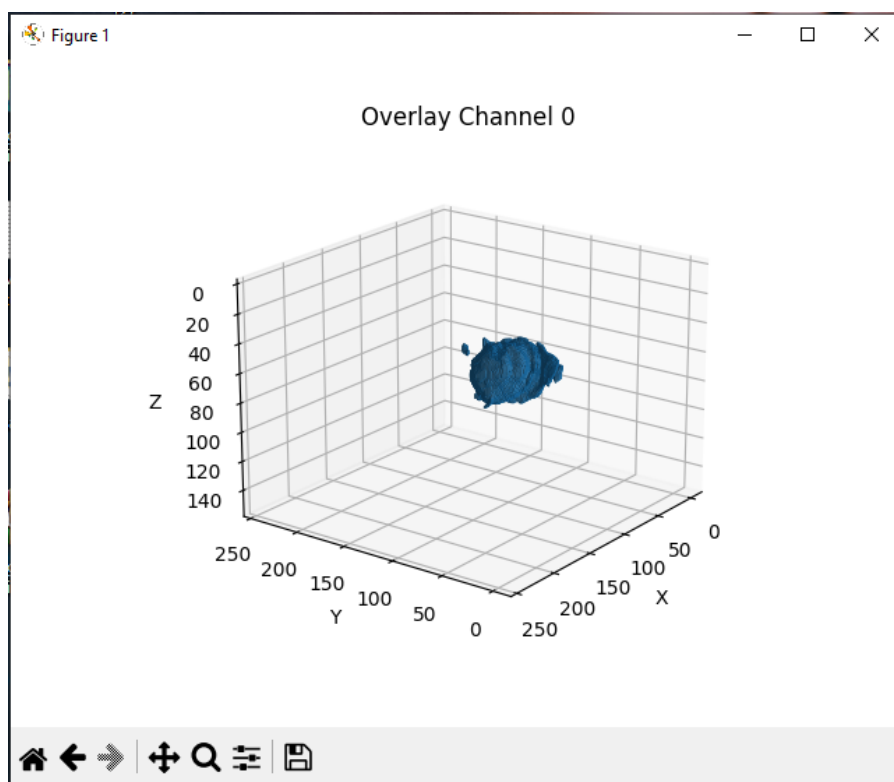


FIGURE IV.23 – Affichage de la tumeur gliome segmentée en 3D.

Sur la figure IV.24, nous présentons une autre vue 3D de la tumeur segmentée.

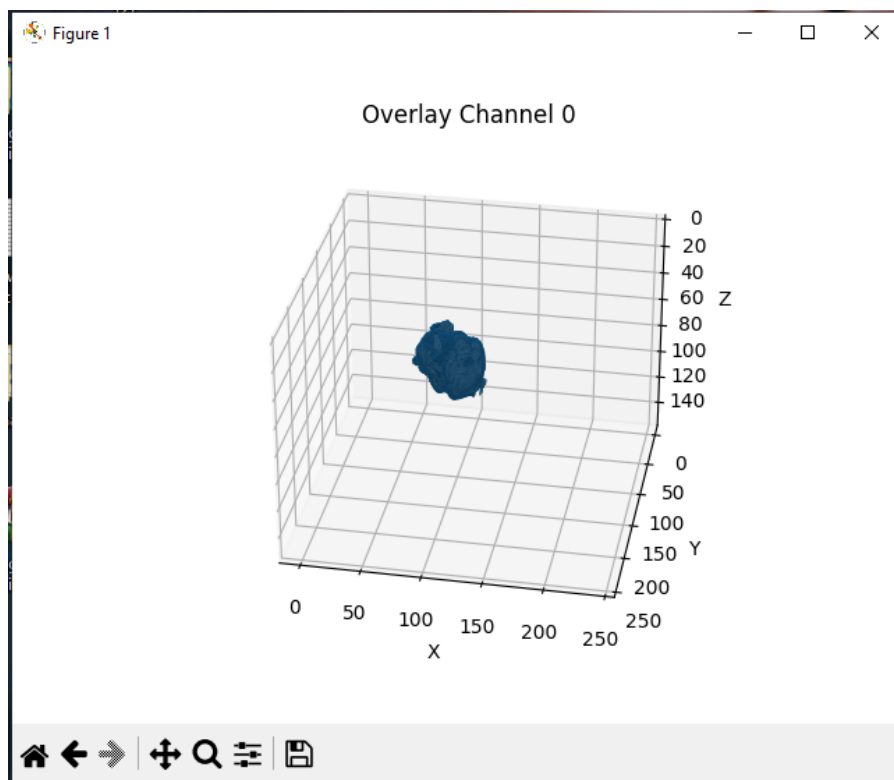


FIGURE IV.24 – Une autre vue 3D de la tumeur segmentée.

Sur la figure [IV.25](#), nous présentons une troisième vue 3D de la tumeur segmentée.

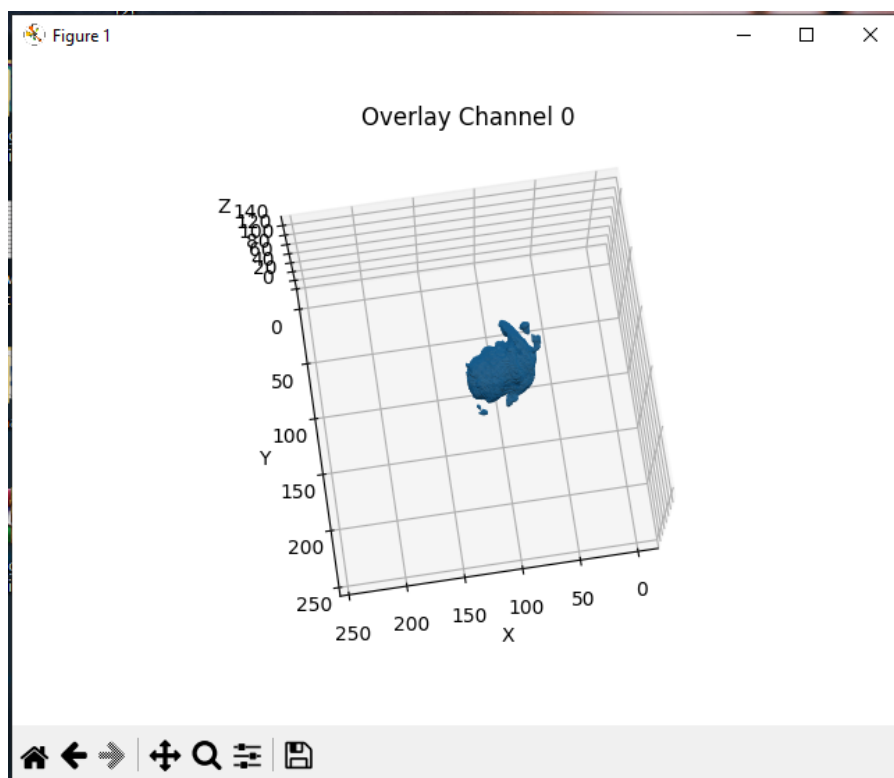
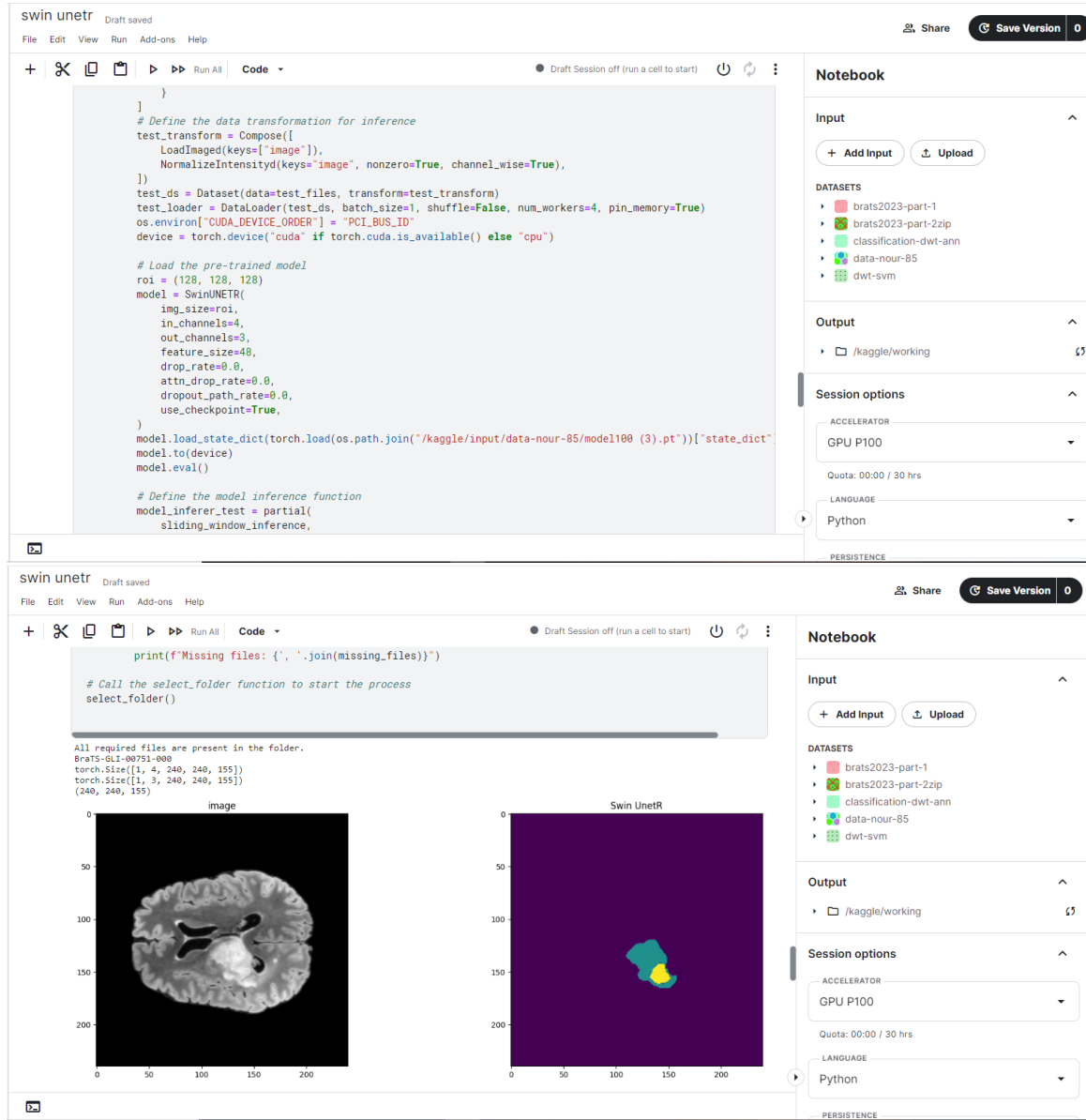
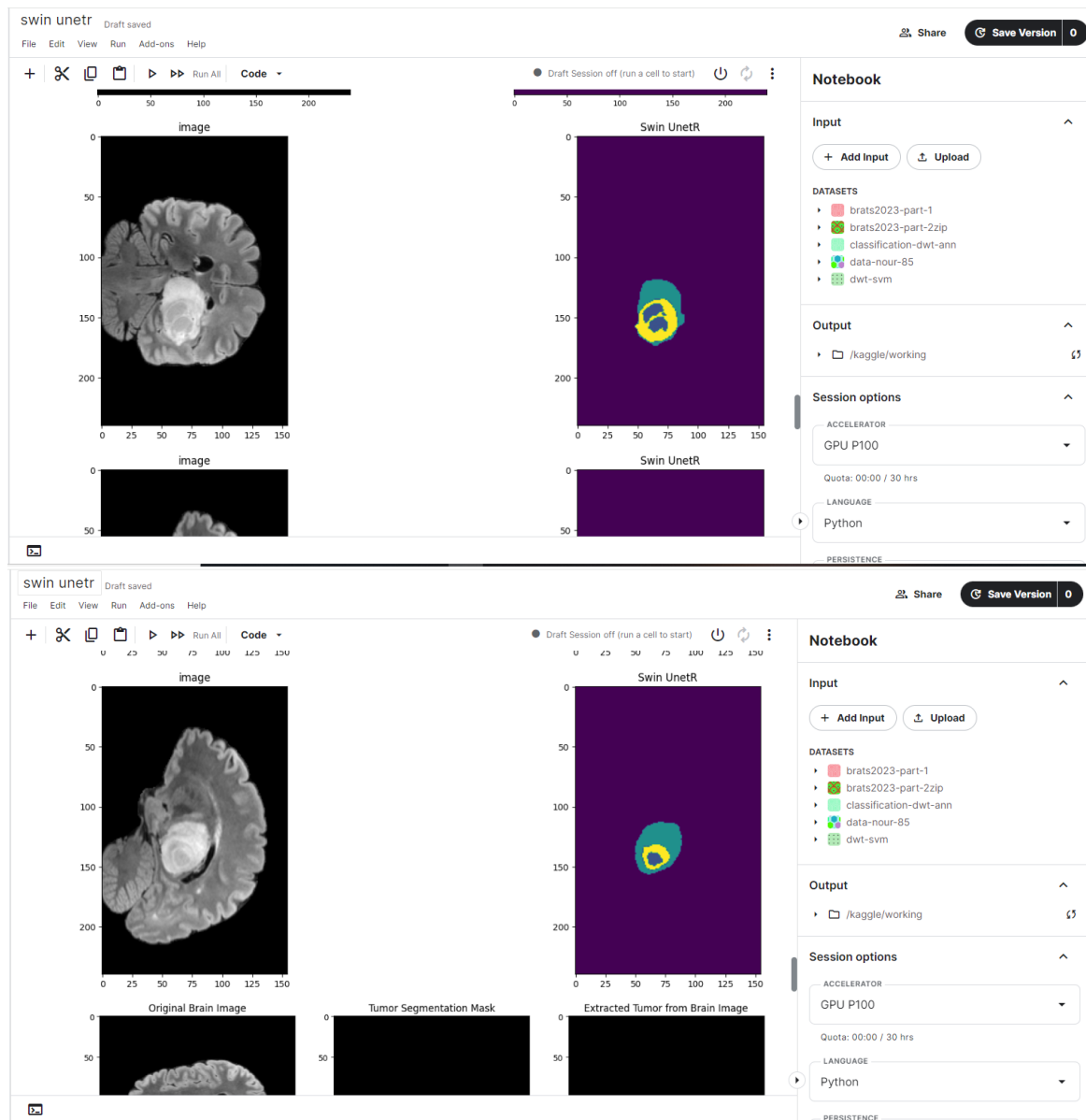


FIGURE IV.25 – Troisième vue 3D de la tumeur segmentée.

IV.8.2 Notebook du système de pronostic automatique du gliome basé SwinUnetR

Les modèles Transformers requièrent de puissantes ressources informatiques. C'est un obstacle majeur qui nous a poussé à réaliser un Notebook pour le système basé SwinUnetR. Nous présentons sur la figure IV.26 l'ensemble des figures du Notebook.





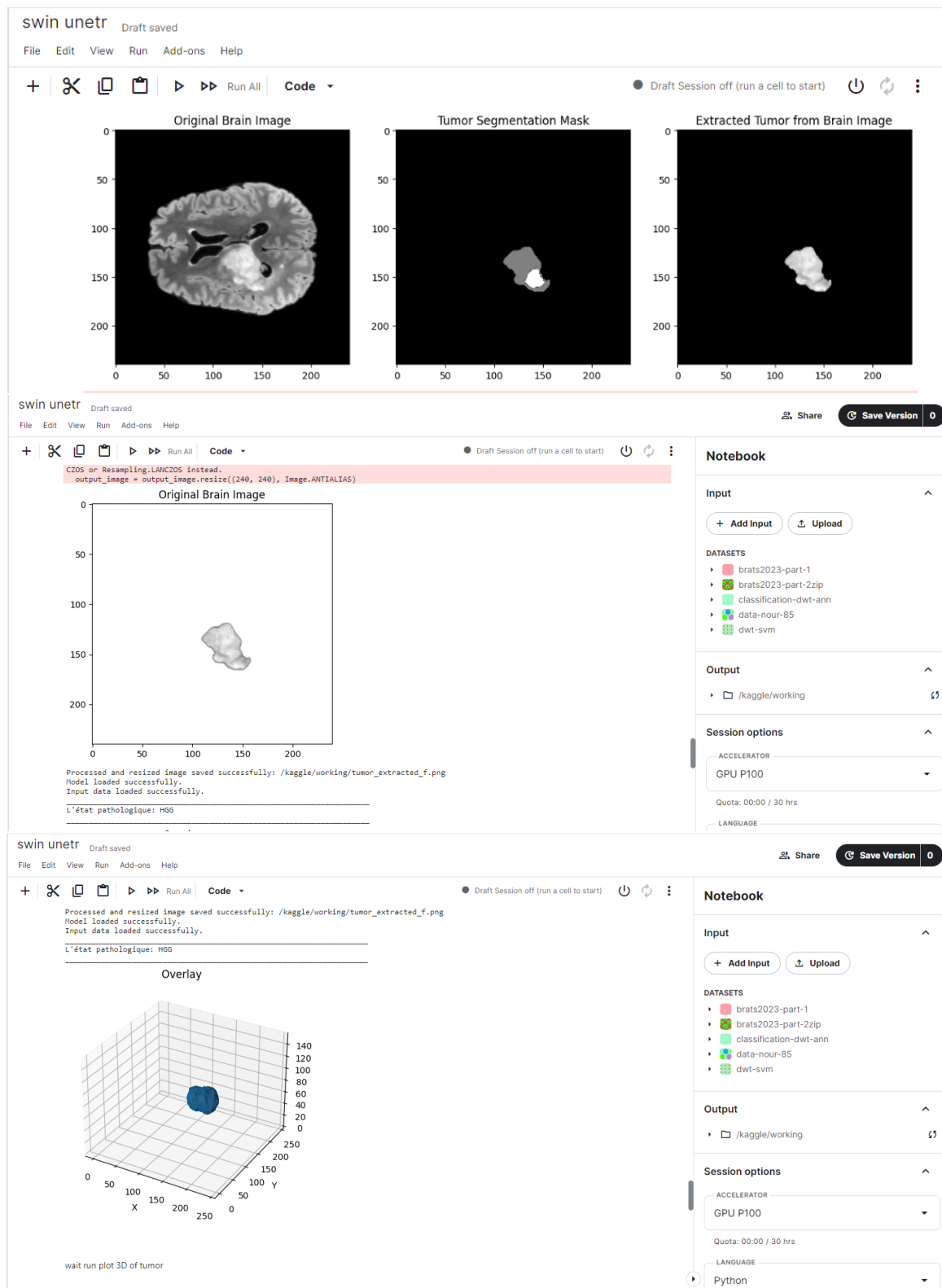


FIGURE IV.26 – Fonctionnalités du Notebook relatif au pronostic automatique du gliome basé SwinUnetR .

IV.9 Conclusion

Dans ce chapitre, nous avons présenté le langage de programmation et les outils qui nous ont semblé les plus utiles pour notre modèle. Enfin, nous avons discuté des résultats de classification que nous avons obtenus et avons constaté que le modèle de Transformer SwinUnetR est meilleur et plus précis que les autres modèles.

CONCLUSION GÉNÉRALE

De nombreuses études montrent que le cerveau est l'élément le plus important et l'organe le plus complexe du corps humain, contenant plus de cent milliards de neurones et chacun avec jusqu'à dix mille synapses [24].

Selon les statistiques de l'organisation mondiale de la santé (OMS), les tumeurs cérébrales présentent l'un des cancers les plus mortels à l'heure actuelle, et leur morbidité augmente d'année en année.

Des quatre classes de la tumeur cérébrale, à savoir le gliome, le méningiome, l'adénome hypophysaire et la tumeur de la gaine nerveuse; les gliomes désignent les tumeurs cérébrales primitives les plus courantes et les plus dangereuses, provenant de cellules gliales.

En pratique, la modalité d'imagerie par résonance magnétique (IRM) offre différents contrastes pour explorer les propriétés des tissus, fournissent des informations complètes sur la structure du cerveau et aussi des indices potentiels pour détecter des tumeurs. Cependant, la segmentation manuelle des régions cibles de l'IRM nécessite beaucoup de temps et efforts des médecins et est sensible à des facteurs subjectifs. Dans ce sens, un pronostic automatique et précis est en ordre pour la segmentation et la classification des régions tumorales cérébrales à partir de l'IRM multimodale afin d'améliorer l'efficacité du diagnostic clinique.

De toute évidence, les méthodes de diagnostic sont essentielles et dépendent fortement des connaissances de base où les tissus cérébraux doivent pouvoir être automatiquement classifiés en deux ou plusieurs lésions tumorales pour que la technique suggérée d'identification et de catégorisation des IRM fonctionne. De ce fait, le médecin spécialiste peut continuer à surveiller le patient tout en établissant un diagnostic rapide et précis.

Les méthodes précédentes de détection automatique des cancers du cerveau par IRM étaient entravées par leur incapacité à se généraliser à d'autres cas et à se reproduire. Dans ce sens, nous avons proposé un modèle qui atteint conjointement l'objectif de segmentation des lésions tumorales en 3D ainsi que leur classification. En effet, SwinUnetR était implémenté durant la phase de segmentation vu que les Transformers peuvent capturer les dépendances à long terme par l'approche de l'attention ainsi que l'analyse des caractéristiques globales mais ignorent les détails des caractéristiques locales, alors que les opérations convolutives sont capable de capturer les caractéristiques locales. En outre, nous avons opté pour la fusion des modalités car elle permet à plusieurs modalités de se compléter; ce qui peut fournir une compréhension plus complète des tumeurs et de leurs tissus environnants cérébraux.

Durant la phase d'évaluation des performances, nous avons démontré que l'approche proposée était pertinente en matière de segmentation en 3D et classification automatique du gliome atteignant un taux de précision de la classification de 94%.

Comme perspectives, nous envisageons de promouvoir les hôpitaux à adopter le diagnostic automatique du cancer du cerveau en appliquant l'approche puissante développée comme deuxième avis durant l'élaboration du diagnostic final tout en analysant d'autres types de maladies.

BIBLIOGRAPHIE

- [1] [Online]. Available : <https://cedric.cnam.fr/vertigo/cours/ml2/coursSVMLineaires.html>
- [2] [Online]. Available : https://www.researchgate.net/publication/343567991_Deep_learning_for_photoacoustic_imaging_a_survey
- [3] A. W. Salehi, S. Khan, G. Gupta, B. I. Alabdullah, A. Almjally, H. Alsolai, T. Siddiqui, and A. Mellit, “A study of cnn and transfer learning in medical imaging : Advantages, challenges, future scope,” *Sustainability*, vol. 15, no. 7, p. 5930, 2023.
- [4] [Online]. Available : <https://france.devoteam.com/paroles-dexperts/mieux-comprendre-le-deep-learning-applique-a-la-reconnaissance-dimages/>
- [5] [Online]. Available : <https://danau-data.medium.com/implementasi-deep-learning-sederhana-menggunakan-keras-3f5726f007e7>
- [6] [Online]. Available : <https://blent.ai/blog/a/unet-computer-vision>
- [7] I. Mecheter, A. Amira, M. Abbod, and H. Zaidi, “Brain mr imaging segmentation using convolutional auto encoder network for pet attenuation correction,” in *Intelligent Systems and Applications : Proceedings of the 2020 Intelligent Systems Conference (IntelliSys) Volume 3*. Springer, 2021, pp. 430–440.
- [8] J. Ni, J. Gao, J. Li, H. Yang, Z. Hao, and Z. Han, “E-alexnet : quality evaluation of strawberry based on machine learning,” *Journal of Food Measurement and Characterization*, vol. 15, no. 5, pp. 4530–4541, 2021.
- [9] H. Nahata and S. P. Singh, “Deep learning solutions for skin cancer detection and diagnosis,” *Machine learning with health care perspective : machine learning and healthcare*, pp. 159–182, 2020.

- [10] P. Sharma, I. Wahlang, S. Sanyal, and A. K. Maji, “Classification of brain mri using deep learning techniques,” in *Soft Computing : Theories and Applications : Proceedings of SoCTA 2018*. Springer, 2020, pp. 559–569.
- [11] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, . Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017.
- [12] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, “Swin transformer : Hierarchical vision transformer using shifted windows,” in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 10 012–10 022.
- [13] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, “Swin-unet : Unet-like pure transformer for medical image segmentation,” in *European conference on computer vision*. Springer, 2022, pp. 205–218.
- [14] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, and D. Xu, “Swin unetr : Swin transformers for semantic segmentation of brain tumors in mri images,” in *International MICCAI Brainlesion Workshop*. Springer, 2021, pp. 272–284.
- [15] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, “Transunet : Transformers make strong encoders for medical image segmentation,” *arXiv preprint arXiv :2102.04306*, 2021.
- [16] S. He, R. Bao, P. E. Grant, and Y. Ou, “U-netmer : U-net meets transformer for medical image segmentation,” *arXiv preprint arXiv :2304.01401*, 2023.
- [17] S. Karimijafarbigloo, R. Azad, A. Kazerouni, S. Ebadollahi, and D. Merhof, “Mmcformer : Missing modality compensation transformer for brain tumor segmentation,” in *Medical Imaging with Deep Learning*. PMLR, 2024, pp. 1144–1162.
- [18] Y. Xie, J. Zhang, C. Shen, and Y. Xia, “Cotr : Efficiently bridging cnn and transformer for 3d medical image segmentation,” in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021 : 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III 24*. Springer, 2021, pp. 171–180.
- [19] M. Heidari, A. Kazerouni, M. Soltany, R. Azad, E. K. Aghdam, J. Cohen-Adad, and D. Merhof, “Hiformer : Hierarchical multi-scale representations using transformers for medical image segmentation,” in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2023, pp. 6202–6212.

- [20] A. Hatamizadeh, Z. Xu, D. Yang, W. Li, H. Roth, and D. Xu, "Unetformer : A unified vision transformer model and pre-training framework for 3d medical image segmentation," *arXiv preprint arXiv :2204.00631*, 2022.
- [21] X. Liu, C. Yao, H. Chen, R. Xiang, H. Wu, P. Du, Z. Yu, W. Liu, J. Liu, and D. Geng, "Btsc-tnas : A neural architecture search-based transformer for brain tumor segmentation and classification," *Computerized Medical Imaging and Graphics*, vol. 110, p. 102307, 2023.
- [22] B. Sarala, G. Sumathy, A. Kalpana, and J. J. Hephzipah, "Glioma brain tumor detection using dual convolutional neural networks and histogram density segmentation algorithm," *Biomedical Signal Processing and Control*, vol. 85, p. 104859, 2023.
- [23] W. Zhang, S. Chen, Y. Ma, Y. Liu, and X. Cao, "Etunet : Exploring efficient transformer enhanced unet for 3d brain tumor segmentation," *Computers in Biology and Medicine*, vol. 171, p. 108005, 2024.
- [24] P. Wang, Q. Yang, Z. He, and Y. Yuan, "Vision transformers in multi-modal brain tumor mri segmentation : A review," *Meta-Radiology*, p. 100004, 2023.
- [25] M. I. Khalil, S. Tehsin, M. Humayun, N. Jhanjhi, and M. A. AlZain, "Multi-scale network for thoracic organs segmentation." *Computers, Materials & Continua*, vol. 70, no. 2, 2022.
- [26] A. Myronenko, "3d mri brain tumor segmentation using autoencoder regularization," in *Brainlesion : Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries : 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II 4*. Springer, 2019, pp. 311–320.
- [27] D. AlSaeed and S. F. Omar, "Brain mri analysis for alzheimer's disease diagnosis using cnn-based feature extraction and machine learning," *Sensors*, vol. 22, no. 8, p. 2911, 2022.
- [28] K. R. Reddy and R. Dhuli, "Segmentation and classification of brain tumors from mri images based on adaptive mechanisms and eldp feature descriptor," *Biomedical Signal Processing and Control*, vol. 76, p. 103704, 2022.
- [29] A. H. Nizamani, Z. Chen, A. A. Nizamani, and U. A. Bhatti, "Advance brain tumor segmentation using feature fusion methods with deep u-net model with cnn for mri data," *Journal of King Saud University-Computer and Information Sciences*, vol. 35, no. 9, p. 101793, 2023.
- [30] J. Liang, C. Yang, and L. Zeng, "3d pswinbts : an efficient transformer-based unet using 3d parallel shifted windows for brain tumor segmentation," *Digital Signal Processing*, vol. 131, p. 103784, 2022.

- [31] [Online]. Available : <https://www.ibm.com/fr-fr/topics/computer-vision>
- [32] S. Aoulia, "Segmentation des images médicales par apprentissage profond," Ph.D. dissertation, Université laarbi tebessi tebessa, 2021.
- [33] [Online]. Available : <https://www.datacamp.com/tutorial/seeing-like-a-machine-a-beginners-guide-to-image-analysis-in-machine-learning>
- [34] O. Regniers, "Méthodes d'analyse de texture pour la cartographie d'occupations du sol par télédétection très haute résolution : application à la forêt, la vigne et les parcs ostréicoles," Ph.D. dissertation, Bordeaux, 2014.
- [35] M. H. Bharati, J. J. Liu, and J. F. MacGregor, "Image texture analysis : methods and comparisons," *Chemometrics and intelligent laboratory systems*, vol. 72, no. 1, pp. 57–71, 2004.
- [36] [Online]. Available : <https://www.oracle.com/dz/artificial-intelligence/machine-learning/what-is-machine-learning/>
- [37] [Online]. Available : <https://fastercapital.com/fr/sujet/comment-utiliser-l-analyse-de-r%C3%A9gression-pour-ajuster-une-fonction-de-co%C3%BAt-aux-donn%C3%A9es.html/4>
- [38] [Online]. Available : <https://www.mathworks.com/discovery/support-vector-machine.html>
- [39] [Online]. Available : https://fr.wikipedia.org/wiki/Machine_%C3%A0_vecteurs_de_support
- [40] D. Dernoncourt, "Stabilité de la sélection de variables sur des données haute dimension : une application à l'expression génique," Ph.D. dissertation, Université Pierre et Marie Curie-Paris VI, 2014.
- [41] [Online]. Available : https://fr.wikipedia.org/wiki/R%C3%A9seau_de_neurones_artificiels
- [42] [Online]. Available : <https://fastercapital.com/fr/sujet/introduction-au-deep-learning-et-%C3%A0-son-potentiel.html>
- [43] [Online]. Available : <https://fastercapital.com/fr/contenu/Big-Data--relever-les-defis-du-Big-Data-avec-les-reseaux-de-neurones.html>
- [44] [Online]. Available : <https://merehead.com/fr/blog/avantages-inconvenients-larchitecture-reseau-neuronal/>
- [45] [Online]. Available : <https://www.ibm.com/fr-fr/topics/knn>

- [46] [Online]. Available : <https://fastercapital.com/fr/contenu/Apprentissage-automatique-non-parametrique---repousser-les-limites-de-la-modelisation-predictive.html>
- [47] [Online]. Available : <https://asana.com/fr/resources/decision-tree-analysis>
- [48] [Online]. Available : <https://aws.amazon.com/fr/what-is/deep-learning/>
- [49] [Online]. Available : <https://geekflare.com/fr/convolutional-neural-networks/>
- [50] [Online]. Available : <https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-donnees-visuelles/5083336-decouvrez-les-differentes-couches-dun-cnn>
- [51] [Online]. Available : <https://coinaute.com/fonction-relu/>
- [52] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [53] S. Lu, S.-H. Wang, and Y.-D. Zhang, “Detecting pathological brain via resnet and randomized neural networks,” *Heliyon*, vol. 6, no. 12, 2020.
- [54] A. M. Mostafa, M. A. El-Meligy, M. A. Alkhayyal, A. Alnuaim, and M. Sharaf, “A framework for brain tumor detection based on segmentation and features fusion using mri images,” *Brain Research*, vol. 1806, p. 148300, 2023.
- [55] S. Tabatabaei, K. Rezaee, and M. Zhu, “Attention transformer mechanism and fusion-based deep learning architecture for mri brain tumor classification system,” *Biomedical Signal Processing and Control*, vol. 86, p. 105119, 2023.
- [56] W. K. Soh, H. Y. Yuen, and J. C. Rajapakse, “Hut : Hybrid unet transformer for brain lesion and tumour segmentation,” *Heliyon*, vol. 9, no. 12, 2023.
- [57] W. Wu, J. Yan, Y. Zhao, Q. Sun, H. Zhang, J. Cheng, D. Liang, Y. Chen, Z. Zhang, and Z.-C. Li, “Multi-task learning for concurrent survival prediction and semi-supervised segmentation of gliomas in brain mri,” *Displays*, vol. 78, p. 102402, 2023.
- [58] K. Kaplan, Y. Kaya, M. Kuncan, and H. M. Ertunç, “Brain tumor classification using modified local binary patterns (lbp) feature extraction methods,” *Medical hypotheses*, vol. 139, p. 109696, 2020.
- [59] P. Thejaswini, M. B. Bhat, and M. K. Prakash, “Detection and classification of tumour in brain mri,” *Int. J. Eng. Manufact. (IJEM)*, vol. 9, no. 1, pp. 11–20, 2019.

- [60] B. Tahir, S. Iqbal, M. Usman Ghani Khan, T. Saba, Z. Mehmood, A. Anjum, and T. Mahmood, "Feature enhancement framework for brain tumor segmentation and classification," *Microscopy research and technique*, vol. 82, no. 6, pp. 803–811, 2019.
- [61] A. R. Raju, P. Suresh, and R. R. Rao, "Bayesian hcs-based multi-svnn : a classification approach for brain tumor segmentation and classification using bayesian fuzzy clustering," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 3, pp. 646–660, 2018.
- [62] S. Deepak and P. Ameer, "Brain tumour classification using siamese neural network and neighbourhood analysis in embedded feature space," *International Journal of Imaging Systems and Technology*, vol. 31, no. 3, pp. 1655–1669, 2021.
- [63] M. Sood, S. Jain, and J. Dogra, "Classification and pathologic diagnosis of gliomas in mr brain images," *Procedia Computer Science*, vol. 218, pp. 706–717, 2023.
- [64] Y. Zheng, D. Huang, X. Hao, J. Wei, H. Lu, and Y. Liu, "Univisnet : A unified visualization and classification network for accurate grading of gliomas from mri," *Computers in Biology and Medicine*, vol. 165, p. 107332, 2023.
- [65] A. B. Slama, H. Sahli, Y. Amri, and H. Trabelsi, "Res-net-vgg19 : Improved tumor segmentation using mr images based on res-net architecture and efficient vgg gliomas grading," *Applications in Engineering Science*, vol. 16, p. 100153, 2023.
- [66] M. Zhang, D. Liu, Q. Sun, Y. Han, B. Liu, J. Zhang, and M. Zhang, "Augmented transformer network for mri brain tumor segmentation," *Journal of King Saud University-Computer and Information Sciences*, vol. 36, no. 1, p. 101917, 2024.
- [67] H. Xiao, L. Li, Q. Liu, X. Zhu, and Q. Zhang, "Transformers in medical image segmentation : A review," *Biomedical Signal Processing and Control*, vol. 84, p. 104791, 2023.
- [68] C. F. Muccio, E. Tedeschi, L. Ugga, R. Cuocolo, G. Esposito, and F. Caranci, "Solitary cerebral metastases vs. high-grade gliomas : usefulness of two mri signs in the differential diagnosis," *Anticancer Research*, vol. 39, no. 9, pp. 4905–4909, 2019.
- [69] G. Parthasarathy, L. Ramanathan, K. Anitha, and Y. Justindhas, "Predicting source and age of brain tumor using canny edge detection algorithm and threshold technique," *Asian Pacific journal of cancer prevention : APJCP*, vol. 20, no. 5, p. 1409, 2019.
- [70] M. Hamoud, "Indexation et segmentation d'images basées loi de zipf et zipf inverse," Ph.D. dissertation, Université Badji Mokhtar, 2015.
- [71] A. Belkacem and M. Lyes, "Analyse de texture par les filtres de gabor et laws." Ph.D. dissertation, Université Mouloud Mammeri, 2014.

- [72] M. Haghighat, S. Zonouz, and M. Abdel-Mottaleb, "Cloudid : Trustworthy cloud-based and cross-enterprise biometric identification," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7905–7916, 2015.
- [73] A. A. Asiri, A. Shaf, T. Ali, M. Aamir, M. Irfan, S. Alqahtani, K. M. Mehdar, H. T. Halawani, A. H. Alghamdi, A. F. A. Alshamrani *et al.*, "Brain tumor detection and classification using fine-tuned cnn with resnet50 and u-net model : A study on tcga-lgg and tcia dataset for mri applications," *Life*, vol. 13, no. 7, p. 1449, 2023.
- [74] A. Kumar, S. K. Pandey, N. Varshney, K. U. Singh, T. Singh, and M. A. Shah, "Distinctive approach in brain tumor detection and feature extraction using biologically inspired dwt method and svm," *Scientific Reports*, vol. 13, no. 1, p. 22735, 2023.
- [75] [Online]. Available : [https://ar.wikipedia.org/wiki/%D8%A8%D8%A7%D9%8A%D8%AB%D9%88%D9%86_\(%D9%84%D8%BA%D8%A9_%D8%A8%D8%B1%D9%85%D8%AC%D8%A9\)](https://ar.wikipedia.org/wiki/%D8%A8%D8%A7%D9%8A%D8%AB%D9%88%D9%86_(%D9%84%D8%BA%D8%A9_%D8%A8%D8%B1%D9%85%D8%AC%D8%A9))
- [76] [Online]. Available : <https://www.tiobe.com/tiobe-index/>
- [77] [Online]. Available : <https://www.turing.com/kb/python-best-adapted-to-ai-and-machine-learning>
- [78] [Online]. Available : <https://www.anaconda.com/>
- [79] [Online]. Available : <https://datascientest.com/en/kaggle-everything-you-need-to-know-about-this-platform>
- [80] [Online]. Available : <https://www.coursera.org/articles/kaggle>
- [81] [Online]. Available : <https://www.guru99.com/fr/confusion-matrix-machine-learning-example.html>
- [82] [Online]. Available : <https://byjus.com/physics/accuracy-precision-measurement/>
- [83] [Online]. Available : https://en.wikipedia.org/wiki/Accuracy_and_precision
- [84] [Online]. Available : https://fr.wikipedia.org/wiki/Pr%C3%A9cision_et_rappel
- [85] [Online]. Available : https://docs.aws.amazon.com/fr_fr/lookout-for-vision/latest/developer-guide/improve.html
- [86] [Online]. Available : <https://www.synapse.org/#!Synapse:syn51156910/wiki/622351>
- [87] [Online]. Available : <https://towardsdatascience.com/why-adamw-matters-736223f31b5d>
- [88] W. Wenxuan, C. Chen, D. Meng, Y. Hong, Z. Sen, and L. Jiangyun, "Transbts : Multimodal brain tumor segmentation using transformer," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2021, pp. 109–119.

- [89] Y. Jiang, Y. Zhang, X. Lin, J. Dong, T. Cheng, and J. Liang, “Swinbts : A method for 3d multimodal brain tumor segmentation using swin transformer,” *Brain sciences*, vol. 12, no. 6, p. 797, 2022.
- [90] [Online]. Available : <https://www.horilla.com/blogs/how-to-remove-the-background-of-image-using-rembg-in-python/>
- [91] Hammadi and M. A. Y. Imad Eddine, “Analyse des irm cérébrales basée réseaux de neurones convolutifs et transformers en vue du diagnostic automatique du cancer du cerveau,” Ph.D. dissertation, UNIVERSITE ECHAHID HAMMA LAKHDAR EL OUED, 2023.