الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جامعة أبي بكر بلقايد – تلمسان
Université Aboubakr Belkaïd – Tlemcen –
Faculté de Technologie

Mémoire

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

En : Génie Biomédical

Spécialité : Imagerie Médicale

Par : ***KOTBI Djinane***

Thème :

Système intelligent d'aide au diagnostic Médical par IA avec visualisation en réalité augmentée

Soutenu publiquement le **27/06/ 2026**, devant le jury composé de :

M. Hadj Slimane Zine Eddine	Professeur	Université de Tlemcen	Président
M. Bengana AbdelFatih	MCA	Université de Tlemcen	Examinateur
M. Boukli Hacene Ismail	Professeur	Université de Tlemcen	Superviseur
M. Messadi Mahammed	Professeur	Université de Tlemcen	Co- Superviseur
M. Loudjdi Salim	Professeur	CHU Tlemcen	Membre socio-économique
M. Benadda Lotfi	MCA	Université de Tlemcen	Membre de I2E

Année Universitaire : 2025 / 2026

Abstract

The diagnosis of brain tumors relies on the analysis of MRI images, a complex task requiring precise segmentation of anatomical structures and reliable classification of lesions. Manual methods are time-consuming, subjective, and subject to inter-operator variability. This project aims to automate this process by developing NeuroVision, an intelligent medical decision-support system that combines artificial intelligence with an augmented reality visualization module. The platform integrates deep learning models for 2D segmentation (FPN, Dice = 0.8895) and 3D segmentation (nnU-Net, mean Dice = 0.8436), as well as for multi-class classification (EfficientNet-B0, Accuracy = 99.50%). An intuitive multi-role interface (Radiologist, Neurosurgeon, Researcher) allows the visualization of results through overlays, contours, MNI atlas, and immersive 3D reconstruction under Unity. Augmented reality is used exclusively downstream of the classification process as an interactive visualization tool for the results, and not as input data for the models. The performances achieved on the BRISC 2025 and BraTS 2020 datasets demonstrate the feasibility of a local, efficient, and accessible solution for diagnostic support and surgical planning in neuroimaging.

Keywords: MRI, brain tumors, artificial intelligence, augmented reality, classification, segmentation.

Résumé

Le diagnostic des tumeurs cérébrales repose sur l'analyse d'images IRM, une tâche complexe nécessitant une segmentation précise des structures anatomiques et une classification fiable des lésions. Les méthodes manuelles, longues et subjectives, sont sujettes à des variations inter-opérateurs. Ce projet vise à automatiser ce processus à travers NeuroVision, un système intelligent d'aide au diagnostic médical qui combine l'intelligence artificielle à un module de visualisation en réalité augmentée. La plateforme intègre des modèles de deep learning pour la segmentation 2D (FPN, Dice = 0,8895) et 3D (nnU-Net, Dice moyen = 0,8436), ainsi que pour la classification multi-classes (EfficientNet-B0, Accuracy = 99,50%). Une interface intuitive multi-rôles (Radiologue, Neurochirurgien, Chercheur) permet la visualisation des résultats sous forme d'overlay, de contours, d'atlas MNI et de reconstruction 3D immersive sous Unity. La réalité augmentée intervient exclusivement en aval du processus de classification, en tant qu'outil de visualisation interactive des résultats, et non comme donnée d'entrée des modèles. Les performances obtenues sur les bases BRISC 2025 et BraTS 2020 démontrent la faisabilité d'une solution locale, performante et accessible pour l'aide au diagnostic et la planification chirurgicale en neuro-imagerie.

Mots-clés : IRM, tumeurs cérébrales, intelligence artificielle, réalité augmentée, classification, segmentation.

Remerciements

C'est avec une profonde gratitude que je remercie Allah Tout-Puissant de m'avoir accordé la santé, la force et la persévérance nécessaires pour accomplir ce travail.

À ma petite famille, mes parents Mohammed et Nassima, mon frère Nidal et ma sœur Kholoud. Vous êtes mon trésor, ma raison d'avancer, mon refuge dans les tempêtes. Merci pour votre amour pur, vos sacrifices silencieux et d'avoir toujours cru en moi, même dans mes moments de doute. Votre présence à chaque étape de ma vie a été mon plus grand soutien. Sans vous, je n'aurais jamais eu la force d'aller au bout de mes rêves.

Je remercie mes encadrants, ***M. BOUKLI Hacene Ismail et M. MESSADI Mahammed,*** pour leur accompagnement, leur patience et leurs précieux conseils tout au long de ce travail. Leur disponibilité, leur rigueur et leur bienveillance m'ont permis de mener à bien ce projet et d'apprendre énormément.

J'adresse également mes vifs remerciements à **Pr. Hadj Slimane Z.E**. pour l'honneur qu'il m'a fait en présidant ma soutenance, ainsi qu'à **M. Bengana AbdelFatih** pour avoir accepté d'examiner mon travail.

Je tiens aussi à remercier **M. Benadda Lotfi,** membre de l'I2E, pour l'examen de mon projet, et le Pr. **Loudjdi Salim,** membre socio-économique du CHU de Tlemcen, pour sa précieuse contribution.

Je remercie également tout ***le département de génie biomédical*** et ***l'ensemble de ses enseignant***s pour la qualité de leur formation, leur écoute et leur bienveillance tout au long de mon parcours universitaire.

Un merci tout particulier à ***Nadjet, Fatima Zahra et Asmae.*** Vous êtes bien plus que des amies, vous êtes une famille de cœur. Merci pour votre présence réconfortante, votre soutien indéfectible et tous ces moments de joie, de rires et de complicité gravés à jamais en moi. Vous avez illuminé mon chemin.

Djinane Kotbi

Dédicace

Je rends grâce à **Allah** Tout-Puissant pour m'avoir accordé la force, la patience et la persévérance nécessaires pour mener à bien ce travail. Sa guidance et Sa bienveillance m'ont accompagnée à chaque étape de ce parcours.

À moi-même, à la petite Djnane que j'étais hier. À celle qui a traversé les nuits blanches, les doutes et les moments de fatigue sans jamais abandonner. Merci pour ton courage silencieux, ta détermination et ta foi en tes rêves. Je suis fière de la personne que tu es devenue.

À mon cher père Mohammed. Tu as été mon roc, mon modèle et ma plus grande source d'inspiration. Ton courage, ton travail discret et ton soutien inconditionnel m'ont portée tout au long de ce parcours. Tu m'as appris que rien n'est impossible quand on croit en soi. Ce mémoire est le tien, papa.

À ma chère mère Nassima. Toi qui n'as jamais cessé d'être à mes côtés, dans les joies comme dans les peines. Tu as prié pour moi pendant mes nuits de travail, tu m'as écoutée, consolée et encouragée sans jamais te lasser. Ta présence et ton soutien constant ont été mon refuge. Ce travail est le fruit de tout ton amour et de tes sacrifices.

À mon frère Nidal. Merci pour ta protection discrète mais toujours présente, pour ta force tranquille et ta présence rassurante. Tu as toujours veillé sur moi sans rien demander en retour.

À ma sœur, Dr Kholoud. Toi qui as partagé mes nuits, mes rêves et mes doutes. Merci pour ta complicité, ton écoute et ta douceur. Sans toi, ce chemin aurait été bien plus difficile.

À Nadjet, Fatima Zahra et Asmae. Mes sœurs de cœur. Vous avez été ma bouffée d'air frais, ma joie dans les jours sombres. Merci pour votre amitié sincère, pour tous ces fous rires et ces moments gravés à jamais dans mon cœur.

À mes oncles, Mohamed et Noureddine, pour leur bienveillance, leurs encouragements et leur soutien précieux.

À toute ma famille, proche et lointaine, pour leur amour, leur présence et leurs prières.

À tous ceux qui ont croisé ma route et laissé une empreinte dans mon cœur. Merci.

Djinane Kotbi

Table des matières

Abstract **4**

Résumé **4**

Remerciements **4**

Dédicace **4**

Liste des figures **10**

Liste des tableaux **14**

Liste des acronymes **15**

Introduction générale **16**

Chapitre I : Contexe médical et l'état de l'art **18**

1.1 ***Introduction*** 19

1.2 ***Contexte médical*** 19

1.2.1 ***L'imagerie médicale*** 19

1.2.1.1 ***IRM*** 19

1.2.2 ***Anatomie cérébrale*** 20

1.2.3 ***Les tumeurs cérébrales*** 20

1.2.3.1 ***Gliome*** 21

1.2.3.2 ***Méningiome*** 21

1.2.3.3 ***Tumeur hypophysaire (pituitary tumor)*** 21

1.3 *Etat de l'art* 22

1.3.1 *Intelligence Artificielle pour la segmentation et la classification des tumeurs cérébrales* 22

1.3.1.1 *Segmentation des tumeurs cérébrales : approche classiques* 22

1.3.1.2 *Segmentation des tumeurs cérébrales par Deep Learning* 26

Segmentation des tumeurs cérébrales par deep learning 26

1) CNN (Réseaux de neurones convolutifs 26

2) U-Net (2D et 3D) 28

3) U-Net++ 30

4) Représentations multi-échelles et Feature Pyramid Network (FPN) 31

5) DeepLabv3+ pour la segmentation sémantique des tumeurs cérébrales 33

6) Pipeline du nnU-Net pour la segmentation automatique des images médicales 34

- Configuration nnU-Net 2D 36

- Configuration nnU-Net 3D full resolution : 36

- Configuration nnU-Net 3D cascade (coarse-to-fine) 37

7) Fonctions de perte et métriques d'évaluation 38

1.3.1.3 *Prétraitement des images* 40

1.3.1.4 *Classification des tumeurs cérébrales par IA* 45

1.3.2 *Réalité Augmentée en Neuro-imagerie* 54

1.3.2.1 *Historique de la réalité augmentée* 54

1.3.2.2 *RA pour la visualisation 3D des tumeurs* 55

1.3.2.3 *RA pour la planification préopératoire* 56

1.3.2.4 *Limites actuelles de la RA médicale* 57

1.3.3 *Approches hybrides Intelligence Artificielle-Réalité Augmentée* . . . 58

1.3.3.1 *Fusion IA-RA pour une visualisation intelligente* 58

1.3.4 *Systèmes hybrides existants* 60

1.3.4.1 *Difficultés et verrous scientifiques* 60

1.4 ***Conclusion*** 61

Chapitre II :Conception et Développement du système. **62**

2.1 ***Introduction*** 63

2.2 ***Architecture globale du système*** 63

2.3 ***Base de données*** 65

2.3.1 ***Description des données d'imagerie(BRISC 2025)*** 65

2.3.1.1 ***Augmentation de la base de données*** 66

2.3.1.2 ***Format des images et structures des données*** 67

2.3.1.3 ***Organisation des données (train/test)*** 67

2.3.2 ***Prétraitement des données*** 68

2.3.3 ***Création du Dataset et Data Loader*** 69

2.3.4 ***Segmentation des tumeurs cérébrales*** 70

2.3.4.1 ***Architecture retenue : FPN avec encodeur ResNet34*** 70

2.3.4.2 ***Apport du transfert d'apprentissage (Transfer Learning)*** . 70

2.3.4.3 ***Paramètres d'entraînement*** 71

2.3.4.4 ***Méthode de validation*** 72

2.3.5 ***Classification des tumeurs cérébrales*** 72

2.3.5.1 ***Architectures de classification étudiées*** 73

2.3.5.2 ***Paramètres d'entraînement*** 74

2.3.6 ***Visualisation anatomique et représentation enrichie des résultats*** . 74

2.3.6.1 ***Analyse visuelle des résultats de segmentation*** 75

2.3.6.2 ***Décomposition morphologique des régions tumorales*** 75

2.3.6.3 ***Fusion anatomique et visualisation enrichie*** 76

2.3.6.4 ***Gestion des orientations anatomiques*** 77

2.3.7 ***Description du dataset BraTS2020*** 79

2.3.8 ***Organisation des données et structuration pour nnU-Net*** 80

2.3.9 ***Prétraitement des données*** 81

2.3.10 ***Architecture et modules du framework nnU-Net v2*** 82

2.3.11 ***Segmentation des tumeurs cérébrales*** 82

2.3.11.1 ***Architecture de segmentation et configurations du framework*** 82

2.3.11.2 ***Configuration automatique du pipeline nnU-Net*** 84

2.3.11.3 ***Utilisation de nnU-Net v2*** 85

2.3.12 ***Visualisation tridimensionnelle et réalité augmentée*** 85

2.3.12.1 ***Reconstruction tridimensionnelle et visualisation en réalité augmenté***e 85

2.3.12.2 ***Organisation du projet Unity*** 86

2.3.12.3 ***Architecture logicielle de l'application*** 86

Gestion des patients : 86

Visualisation des modèles 3D : 87

Interaction utilisateur en réalité augmentée : 87

Fonctionnalités complémentaires : 87

2.3.12.4 ***Interface utilisateur*** 87

2.3.12.5 ***Organisation de la scène de réalité augmentée*** 88

2.3.12.6 ***Reconstruction tridimensionnelle et algorithme Marching Cubes*** 89

2.3.12.7 ***Exportation des modèles via script Python*** 90

2.3.12.8 ***Intégration dans Unity et environnement de réalité augmentée*** 90

2.3.12.9 ***Architecture logicielle globale*** 91

2.3.13 ***Intégration de l'intelligence artificielle et de la réalité augmentée*** . 91

2.3.13.1 ***Vue globale de la chaîne de traitement*** 91

Schéma global de l'architecture : 92

2.4 ***Conclusion*** 93

Chapitre III : Résultats expérimentaux **94**

3.1 ***Introduction*** 95

3.2 ***Transition de la réalité augmentée 2D vers la réalité augmentée 3D*** 95

3.2.1 ***Réalité augmentée 2D*** 96

3.2.1.1 ***Resultats de prétraitement des images IRM*** 96

3.2.1.2 ***Résultats de segmentation 2D*** 97

3.2.2 ***Visualisation anatomique et représentation enrichie des résultats*** . 104

3.2.2.1 ***Analyse globale des résultats de segmentation*** 104

3.2.3 ***Réalité augmentée 3D*** 106

3.2.3.1 ***Résultats du prétraitement automatique par nnU-Net*** 106

3.2.3.2 ***Segmentation des tumeurs cérébrales*** 107

3.2.4 ***Post-traitement, analyse quantitative et génération des modèles 3D pour la réalité augmentée*** 112

3.2.5 ***Visualisation qualitative par superposition*** 113

3.2.6 ***Évaluation quantitative des performances*** 113

3.2.7 ***Reconstruction tridimensionnelle et visualisation en réalité augmentée*** 115

3.2.7.1 ***Interface utilisateur*** 115

3.2.7.2 ***Résultat final de visualisation*** 116

3.2.7.3 ***Analyse comparative réalité augmentée 2D vs réalité augmentée 3D et discussion*** 119

3.2.8 ***De la segmentation à la classification : impact de l'enrichissement des données*** 120

3.2.8.1 ***Benchmarking des architectures de classification*** 120

3.2.8.2 ***Résultats sans réalité augmentée*** 121

3.2.8.3 ***Résultats avec réalité augmentée*** 123

3.2.8.4 ***Sélection du meilleur modèle*** 125

3.2.8.5 ***Analyse du modèle retenu (EfficientNet-B0)*** 125

3.3 ***Conclusion*** 132

Conclusion générale **134**

Annexes **135**

Table des figures

1.1 Comparaison des séquences IRM T1, T2 et FLAIR pour la visualisation des tumeurs cérébrales [7]. 20

1.2 Exemples d'images IRM cérébrales illustrant différentes catégories utilisées pour la classification des tumeurs cérébrales : (a) cerveau sain (no tumor), (b) gliome, (c) méningiome et (d) tumeur pituitaire[10]. 21

1.3 Image IRM cérébrale présentant un gliome infiltrant avec des contours tumoraux diffus [13]. 22

1.4 : Résultat de la segmentation par seuillage sur une image IRM cérébrale [14]. 23

1.5 : Résultat de la segmentation par croissance de régions sur une image IRM cérébrale : (a) Image originale, (b) Image filtrée, (c) Segmentation par croissance de régions [15]. 23

1.6 Résultat de la segmentation par watershed sur une image IRM cérébrale [16]. 24

1.7 Segmentation d'une image cérébrale CT utilisant la méthode des contours actifs [17]. 24

1.8 Résultat de segmentation par Fuzzy C-Means sur une image IRM cérébrale [12]. 25

1.9 Combinaison des méthodes Fuzzy C-Means et seuillage pour la segmentation d'images cérébrales [12]. 25

1.10 Architecture d'un réseau de neurones convolutifs (CNN) appliqué à la segmentation médicale [2]. 27

1.11 Principe de la segmentation patch-based appliquée aux images IRM [18]. 28

1.12 : Schéma de l'architecture U-Net pour la segmentation d'images biomédicales [20]. 29

1.13 Architecture 3D U-Net pour la segmentation médicale volumétrique [21]. 29

1.14 Architecture U-Net++ avec skip connections imbriquées et deep supervision [28] 30

1.15 Comparaison des approches multi-échelles (Image Pyramid, Single Scale, Feature Hierarchy, FPN) [22] 31

1.16 Architecture DeepLabv3+ avec backbone, ASPP et décodeur de segmentation [24] 33

1.17 Pipeline complet du framework nnU-Net pour la segmentation des images IRM cérébrales [25] 34

1.18 Effet de la normalisation des intensités sur une image IRM [33] 41

1.19 Filtrage médian et gaussien suivi de la segmentation level-set [36] 42

1.20 Alignement multimodal IRM 3D / image RGB 2D par recalage médical [60] 43

1.21 Pipeline de segmentation par atlas basé sur le recalage affine et déformable [37] 44

1.22 Illustration des opérations de dilatation et d'érosion en traitement d'image [37] 45

1.23 Schéma illustrant le processus de classification en imagerie médicale [39]. 46

1.24 Ivan Sutherland, considéré comme l'un des précurseurs de la réalité augmentée [43]. 55

1.25 Modélisation tridimensionnelle du cerveau illustrant la localisation de la tumeur et ses relations avec les structures anatomiques environnantes dans un environnement de RA [44]. 56

1.26 : Planification préopératoire d'une tumeur cérébrale à l'aide de RA [46] :
A : Modèle 3D du cerveau reconstruit à partir d'images cérébrales d'IRM avec métastase cérébrale en rouge
B : Après enregistrement, l'image créée dans le cadre A est projetée sur la tête du patient dans la planification de l'incision cutanée et de la craniotomie
C : Projection image en peropératoire après la 1ère incision cutanée et avant la craniotomie avec localisation de la tumeur sur le crâne pour planifier une craniotomie sur mesure
D : Surface du cerveau après ouverture de la dure-mère. La tumeur n'est pas visible à l'œil directement
E : Projection de la tumeur à la surface du cerveau [Neurosurgery Blog]. 57

2.1 Architecture globale du système proposé. 65

2.2 Pipeline de classification des tumeurs cérébrales basé sur les résultats de segmentation. 73

2.3 Illustration des classes de segmentation 79

2.4 Pipeline du modèle de segmentation nnU-Net appliquée aux images IRM 2D [25]. 83

2.5 Organisation des scripts C# dans le projet Unity 86

2.6 Organisation du Canvas (Hierarchy Unity). 88

2.7 Organisation de la hiérarchie de la scène de réalité augmentée sous Unity 89

2.8 *Chaîne d'intégration entre le modèle de segmentation nnU-Net et l'application de réalité augmentée développée sous Unity.* 92

3.1 Pipeline de prétraitement des images IRM 2D (BRISC 2025) 96

3.2 Résumé de l'étude de l'impact du seuil de binarisation sur la segmentation des tumeurs 100

3.3 Performance moyenne (Dice Score) 100

3.4 Exemple de segmentation obtenue avec le modèle FPN (Image IRM – Ground Truth – Prediction) 103

3.5 Courbe d'apprentissage du modèle FPN pour un batch size de 8 104

3.6 Visualisation complète des résultats de segmentation : IRM, GT, prédiction, superposition, décomposition tumorale et fusion avec atlas cérébral 105

3.7 Courbe du Dice Score par image sur le jeu de test 105

3.8 Pretraitement automatique avec nnU Net sur une image 106

3.9 Structure des résultats de validation croisée (nnU-Net – K-Fold) 108

3.10 Contenu du dossier fold_1 généré après l'entraînement du modèle nnU-Net 108

3.11 Comparaison des Dice Scores obtenus avec les configurations nnU-Net 2D et 3D 109

3.12 Visualisation qualitative des résultats de segmentation sur un cas BraTS2020 110

3.13 Évolution du Dice Score, de la loss et du learning rate durant l'entraînement (Fold 0 – configuration 3D full resolution) 111

3.14 Structure des résultats générés automatiquement pour un patient 113

3.15 Superposition entre la vérité terrain et la segmentation prédite (overlay) 114

3.16 Interface utilisateur en scène 116

3.17 Visualisation du cerveau et de la tumeur dans l'environnement Unity 116

3.18 Navigation entre les patients et interaction avec les modèles 3D 117

3.19 Visualisation de la tumeur seule 117

3.20 Contrôle de l'opacité des structures anatomiques à l'aide des sliders 118

3.21 Courbe EfficientNet sur base de données originale 126

3.22 Courbe EfficientNet sur base de données fusionnée (segmentation) 126

3.23 Matrices de confusion du modèle EfficientNet-B0. (a) Matrice sur la base de données originale. (b) Matrice sur la base de données fusionnée (avec segmentation). 127

3.24 Courbes ROC du modèle EfficientNet-B0 : (a) base de données originale, (b) base de données fusionnée 129

25 Interface Unity Hub 141

26 Version Unity 6000.4.2f1 installée dans Unity Hub 142

27 Sélection du template AR Mobile 142

28 Packages de réalité augmentée installés via Unity Package Manager 143

29 Interface de connexion de la plateforme NeuroVision 144
30 Interface de sélection des rôles d'utilisateurs de NeuroVision 145
31 Interface principale du module Radiologist AI Viewer 145
32 Image IRM originale du patient 146
33 Masque réel (Ground Truth) de la tumeur 146
34 Masque prédit par le modèle de segmentation 147
35 Visualisation des contours de la tumeur superposés sur l'image IRM 147
36 Superposition du masque tumoral sur l'image IRM 148
37 Résultats de l'analyse quantitative de la segmentation 148
38 Superposition de la tumeur et de ses sous-classes sur un atlas cérébral 149
39 Résultat de classification de la tumeur cérébrale 149
40 Chargement d'un volume médical dans le module Neurosurgeon 150
41 Visualisation 2D de l'overlay tumoral 151
42 Résultats d'analyse du module Neurosurgeon 151
43 Résultat de classification de la tumeur 151
44 Visualisation d'une coupe IRM (View Slice) 152
45 Visualisation 3D interactive de l'IRM et de la tumeur sous Unity 153
46 Interface de recherche et profil patient 154
47 Affichage du contexte clinique structuré 154
48 Dossiers médicaux et visualisation des résultats 155
49 Business Model Canvas de NeuroVision présentant la proposition de valeur, les segments de clients, les partenaires clés, les ressources, les coûts et les sources de revenus de la plateforme. 161
50 Exemple des fichiers générés après le prétraitement automatique pour un patient du dataset 164
51 configuration 2D - fold 3 164
52 configuration 2D - fold 4 165
53 configuration 3D full resolution - fold 0 165
54 configuration 3D full resolution - fold 3 165
55 La configuration détaillée des paramètres d'entraînement 166

Liste des tableaux

1.1 Comparaison des architectures de réseaux de neurones pour la classification d'images médicales 50

1.2 Limitations, impacts cliniques et solutions potentielles de la réalité augmentée en neurochirurgie 58

2.1 Répartition des données pour les tâches de classification et de segmentation [62] 66

2.2 Paramètres d'entraînement utilisés pour la classification des tumeurs cérébrales 74

2.3 Difficultés et défis liés à la décomposition morphologique et à l'intégration de l'atlas MNI 78

2.4 Comparaison des configurations du framework nnU-Net[25] 84

3.1 Caractéristiques des données après prétraitement 97

3.2 Performances des architectures de segmentation en fonction du nombre d'époques 97

3.3 Performance des architectures en fonction du seuil de binarisation 99

3.4 Optimisation du modèle FPN : influence du batch size et du nombre d'époques sur le Dice Score 101

3.5 Description des fichiers générés par patient 113

3.6 Métriques d'évaluation pour cinq patients du dataset 114

3.7 Mesures quantitatives moyennes de la reconstruction 3D 115

3.8 Performances des architectures de classification sans réalité augmentée 121

3.9 Performances des architectures de classification avec réalité augmentée 123

3.10 Comparaison des performances de segmentation 2D sur BRISC 2025 130

3.11 Comparaison des performances de segmentation 3D sur BraTS 2020 131

3.12 Comparaison des performances de classification sur BRISC 2025 132

3.13 Récapitulatif des environnements de développement utilisés 140

Liste des acronymes

Acronyme	Signification
AUC	Area Under Curve
BraTS	Brain Tumor Segmentation Challenge
BRISC	Brain Tumor MRI Dataset
CHU	Centre Hospitalo-Universitaire
CNN	Convolutional Neural Network
DA	Dinar Algérien
Dice	Coefficient de Dice
ET	Enhancing Tumor
FN	False Negative
FP	False Positive
FPN	Feature Pyramid Network
FPR	False Positive Rate
GPU	Graphics Processing Unit
GT	Ground Truth
HD95	Hausdorff Distance at 95%
IA	Intelligence Artificielle
IoU	Intersection over Union
IRM	Imagerie par Résonance Magnétique
JSON	JavaScript Object Notation
MNI	Montreal Neurological Institute
nnU-Net	No-New-Net
PACS	Picture Archiving and Communication System
RA	Réalité Augmentée
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic
SaaS	Software as a Service
TC	Tumor Core
TN	True Negative
TP	True Positive
TPR	True Positive Rate
WHO	World Health Organization
WT	Whole Tumor

Introduction générale

Les progrès récents en imagerie médicale ont considérablement transformé le diagnostic et le suivi des maladies , en offrant aux professionnels de santé des outils précis et non invasifs pour explorer l'anatomie et le fonctionnement des organes .Parmi ces avancées ,l'imagerie par résonance magnétique (IRM) joue un rôle centrale dans l'évaluation des pathologies cérébrales , notamment des tumeurs, permettant de détecter et de localiser les lésions avec une grande précision , et d'évaluer leur évolution dans le temps .ces techniques ont ouvert la voie à une meilleure planification thérapeutique et à un suivi plus rigoureux , réduisant les risques liés aux interventions chirurgicales et améliorant la prise en charge des patients .

Cependant, malgré la précision des Images IRM, l'interprétation des volumes cérébraux demeure complexe et chronophage. La segmentation manuelle des tumeurs cérébrales par les cliniciens est souvent sujette à des variations inter-observateurs et à des erreurs, du fait de l'hétérogénéité des tissus tumoraux et de leur similitude avec le parenchyme cérébrale sain. De plus, la classification des tumeurs selon leur type nécessite une expertise approfondie et un temps considérable, ce qui limite la rapidité de la prise de décision clinique. Ces difficultés soulignent la nécessité de développer des systèmes d'aides au diagnostic capables de compléter l'expertise médicale et d'optimiser la précision des analyses.

Dans ce contexte, les technologies émergentes telles que l'intelligence artificielle (IA) et réalité augmentée(RA) offrent des solutions prometteuses. L'IA, notamment à travers les réseaux de neurones convolutifs (CNN) et leurs variantes, permet une segmentation automatique et une classification précise des tumeurs cérébrales à partir d'images IRM, en exploitant de grandes bases de données et en réduisant la dépendance à l'expertise humain. La RA, quant à elle, permet de visualiser les structures tumorales en superposition avec l'image originale, offrant une interaction immersive et facilitant la planification préopératoire ou l'analyse détaillée des volumes cérébraux. L'intégration de ces deux technologies constitue une approche hybride capable d'améliorer la qualité et la rapidité du diagnostic médical.

Malgré les avancées des techniques IA et RA, plusieurs défis subsistent. La segmentation manuelle reste lourde et sujette à des erreurs, tandis que les méthodes automatiques doivent encore atteindre un niveau de fiabilité comparable à celui des experts. La visualisation augmentée, bien que prometteuse, rencontre des contraintes techniques et ergonomiques, et son intégration dans un flux de travail clinique n'est pas encore pleinement réalisée.

Le présent travail vise à concevoir un système intelligent d'aide au diagnostic des tumeurs cérébrales, combinant IA pour la segmentation et classification des tumeurs et RA pour la visualisation interactive des zones tumorales. L'objectif est de fournir aux cliniciens un outil capable de soutenir la prise de décision, d'améliorer la précision des diagnostics, et d'optimiser la compréhension spatiale des lésions dans le cerveau.

Comment l'intégration de l'intelligence artificielle et de la réalité augmentée peut-elle améliorer le processus diagnostique des tumeurs cérébrales, en considérant les critères de performance, les limites techniques et éthiques, et les contraintes cliniques pour un système d'aide au diagnostic ?

Le projet est organisé en plusieurs chapitres. Le chapitre 1 présente le contexte médical des tumeurs cérébrales ainsi que l'état de l'art des méthodes basées sur l'intelligence artificielle et la réalité augmentée appliquées à l'analyse des images IRM.

Le chapitre 2 décrit la conception et le développement du système proposé, incluant les étapes de prétraitement, de segmentation, de classification et de visualisation en réalité augmentée 2D et 3D.

Le chapitre 3 expose les expérimentations réalisées ainsi que les résultats obtenus et leur analyse.

Ainsi, ce travail vise à démontrer la faisabilité et l'efficacité d'une approche combinant IA et RA pour l'aide au diagnostic des tumeurs cérébrales, contribuant à l'amélioration de la qualité des soins et à l'optimisation de la prise de décision médicale.

CHAPITRE I

CONTEXE MEDICAL ET L'ETAT DE L'ART

1.1 *Introduction*

Le diagnostic des tumeurs cérébrales représente un défi majeur en raison de la complexité anatomique du cerveau et de l'hétérogénéité des lésions tumorales, notamment en imagerie par résonance magnétique (IRM), ont considérablement amélioré la visualisation des structures cérébrales et la détection des anomalies. Toutefois, l'interprétation des images IRM demeure une tâche délicate, nécessitant une expertise approfondie et pouvant être sujette à variabilité inter-observateur.

Parallèlement, les progrès récents en intelligence artificielle (IA)ont ouvert de nouvelles perspectives dans l'analyse automatisée des images médicales , en particulier pour la segmentation et la classification des tumeurs cérébrales. De plus, l'émergence des technologies de réalité augmentée (RA) offre des possibilités innovantes de visualisation interactive et d'assistance à la prise de décision clinique.

Dans ce contexte, ce chapitre présente d'abord les bases médicales nécessaires à la compréhension des tumeurs cérébrales et de leur exploration par IRM. Il expose ensuite un état de l'art détaillé des approches basées sur l'intelligence artificielle, des applications de la réalité augmentée en médecine, ainsi que des travaux intégrant ces deux technologies dans des systèmes hybrides d'aide au diagnostic.

1.2 *Contexte médical*

1.2.1 *L'imagerie médicale*

L'imagerie médicale constitue aujourd'hui un pilier fondamental du diagnostic clinique moderne [5]. Elle regroupe l'ensemble des techniques permettant de visualiser les structures internes du corps humain de manière non invasive, afin de détecter, localiser et caractériser les anomalies pathologiques. Grâce aux avancées technologiques, ces modalités offrent des images de haute résolution permettant une analyse anatomique et parfois fonctionnelle des organes.

Dans le domaine neurologique ,l'imagerie médicale occupe une place centrale en raison de la complexité structurelle et fonctionnelle du cerveau .Les principales techniques utilisées incluent la tomodensitométrie (TDM) , la topographie par émission de positrons (TEP) et l'imagerie par résonance magnétique (IRM) [5] [6] .Parmi ces modalités, l'IRM s'est imposée comme la technique de référence pour l'exploration du cerveau, en raison de sa capacité à fournir un excellent contraste des tissus mous sans exposition aux rayonnements ionisants [6] [7]. Ainsi, l'imagerie médicale constitue la base sur laquelle reposent les systèmes modernes d'aide au diagnostic, notamment ceux intégrant des approches d'intelligence artificielle.

1.2.1.1 *IRM*

L'imagerie par résonance magnétique est une technique d'imagerie non invasive reposant sur les propriétés magnétiques des noyaux d'hydrogène présents en abondance dans les tissus biologiques [6] [4]. Lorsqu'un patient est placé dans un champ magnétique intense, les protons s'alignent selon ce champ. L'application d'une impulsion radiofréquence perturbe cet équilibre, et le signal émis lors du retour à l'état initial est capté puis transformé en image [5]. L'IRM présente un avantage majeur dans l'exploration cérébrale grâce à son excellente résolution en contraste, permettant de distinguer la

matière grise, la matière blanche et le liquide céphalorachidien. Plusieurs séquences sont couramment utilisées dans l'étude des tumeurs cérébrales :

- La séquence T1, utile pour l'analyse anatomique détaillée.
- La séquence T2, sensible aux variations hydriques et aux œdèmes.
- La séquence FLAIR, particulièrement efficace pour détecter les lésions proches des ventricules.

Ces différentes séquences fournissent des informations complémentaires, essentielles pour la détection et la caractérisation des lésions tumorales [6].

Afin d'illustrer les differences de contraste et d'information fournie par les principales séquences IRM utilisées en neuro-imagerie, la figure suivante présente une comparaison entre les séquences T1, T2 et FLAIR dans le cas des tumeurs cérébrales (Figure 1.1).

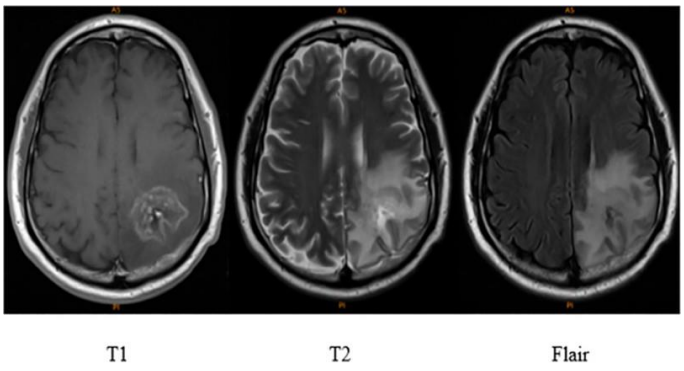


Figure 1.1 : Comparaison des séquences IRM T1, T2 et FLAIR pour la visualisation des tumeurs cérébrales [7].

1.2.2 *Anatomie cérébrale*

Le cerveau est l'organe centrale du système nerveux et se compose principalement de matière grise et de matière blanche. La matière grise, située principalement au niveau du cortex cérébral, est constituée des corps cellulaires neuronaux responsables du traitement de l'information. La matière blanche, quant à elle, est formée d'axones myélinisés assurant la transmission des signaux entre différentes régions cérébrales [7].

On distingue également les ventricules cérébraux, cavités contenant le liquide céphalorachidien, ainsi que des structures spécifiques telles que la glande hypophysaire (pituitaire) [7], localisée à la base du cerveau et jouant un rôle majeur dans la régulation hormonale.

La connaissance précise de cette anatomie est indispensable pour identifier les zones pathologiques et évaluer l'extension tumorale sur les images IRM.

1.2.3 *Les tumeurs cérébrales*

Une tumeur cérébrale correspond à une prolifération anormale et incontrôlée de cellules au sein du tissu cérébral. Elle peut être bénigne ou maligne, primaire ou secondaire (métastatique) [10]. Les

tumeurs cérébrales posent un défi diagnostique majeur en raison de leur hétérogénéité morphologique et de leur caractère parfois infiltrant[8],[9].

1.2.3.1 *Gliome*

Les gliomes sont des tumeurs issues des cellules gliales, qui assurent le soutien et la protection des neurones. Ils représentent les tumeurs cérébrales primaires les plus fréquentes. Leur caractère infiltrant rend leur délimitation difficile sur les images IRM, ce qui complique à la fois le diagnostic et la planification thérapeutiques [9].

1.2.3.2 *Méningiome*

Le méningiome est une tumeur se développant à partir des méninges, membranes enveloppant le cerveau et la moelle épinière. Il est généralement bien délimité et souvent bénin, ce qui facilite son identification sur les images IRM comparativement aux gliomes infiltrants [10].

1.2.3.3 *Tumeur hypophysaire (pituitary tumor)*

Les tumeurs hypophysaires se développent au niveau de la glande pituitaire. Elles peuvent entrainer des troubles hormonaux en raison de l'altération de la fonction endocrine. Leur localisation spécifique nécessite une analyse ciblée des coupes IRM centrées sur la région sellaire [11].

La figure suivante illustre quelques exemples des différentes classes de tumeurs cérébrales (Figure 1.2).

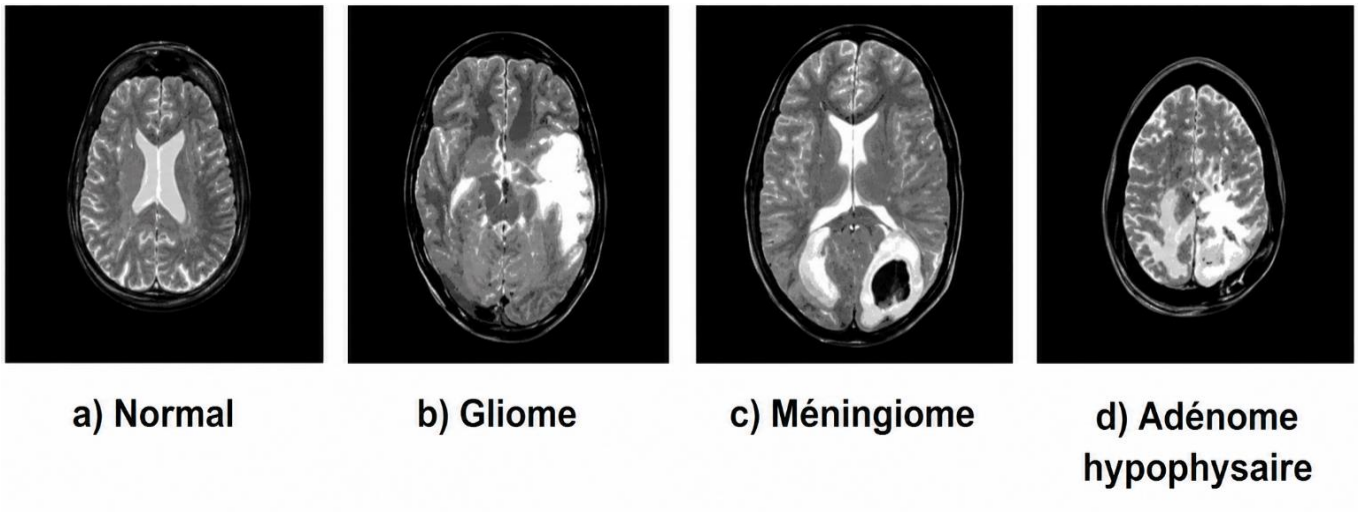


Figure 1.2 : Exemples d'images IRM cérébrales illustrant différentes catégories utilisées pour la classification des tumeurs cérébrales : (a) cerveau sain (no tumor), (b) gliome, (c) méningiome et (d) tumeur pituitaire[10].

1.3 *Etat de l'art*

1.3.1 *Intelligence Artificielle pour la segmentation et la classification des tumeurs cérébrales*

1.3.1.1 *Segmentation des tumeurs cérébrales : approche classiques*

La segmentation des tumeurs cérébrales est une étape essentielle dans l'analyse des images IRM, permettant de délimiter précisément les régions tumorales.

Avant l'émergence de deep learning, cette tâche reposait principalement sur des méthodes classiques de traitement d'image. Ces approches exploitent les variation d'intensité et de contraste pour distinguer les tissus sains des tissus pathologiques [12]. Toutefois, la complexité et l'hétérogénéité des tumeurs cérébrales rendent cette segmentation particulièrement difficile.

1) Méthodes basées sur le traitement d'image

Avant l'essor des méthodes d'apprentissage profond, la segmentation des tumeurs cérébrales en imagerie par résonance magnétique(IRM) reposait principalement sur des techniques classiques de traitement d'image. Ces approches visaient à exploiter les différences d'intensité, de texture ou de contraste entre les tissus sains et pathologiques.

Cependant, dans de nombreux cas, les tumeurs cérébrales présentent des contours irréguliers et des zones infiltrantes qui rendent leur délimitation difficile. La figure suivante illustre un exemple d'image IRM montrant une tumeur cérébrale infiltrante et la complexité associée à sa segmentation (Figure 1.3)[13].

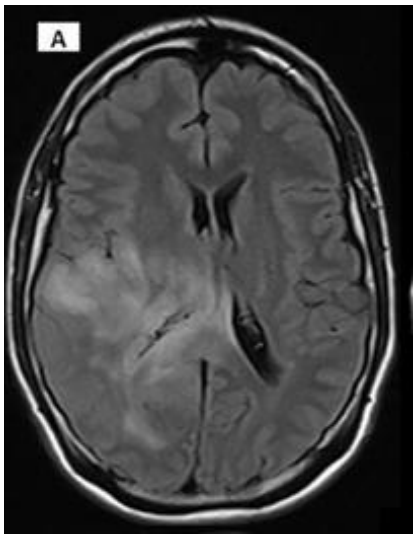


Figure 1.3 : Image IRM cérébrale présentant un gliome infiltrant avec des contours tumoraux diffus [13].

Parmi les méthodes les plus utilisées, on distingue :

a) Le seuillage adaptatif

Le seuillage consiste à segmenter l'image en séparant les pixels selon un seuil d'intensité. Dans le

cas des tumeurs cérébrales, un seuil est choisi afin de distinguer les régions hyperintenses (souvent associées à la tumeur ou à l'œdème) du tissu cérébral normal. Le seuillage adaptatif améliore cette approche en tenant compte des variations locales d'intensités [14]. Cependant, en IRM cérébrale, les variations d'illumination, le bruit et les hétérogénéités internes de la tumeur rendent difficile la détermination d'un seuil optimal universel.

La figure suivante montre un exemple de segmentation par seuillage appliquée a une image IRM cérébrale (Figure 1.4).

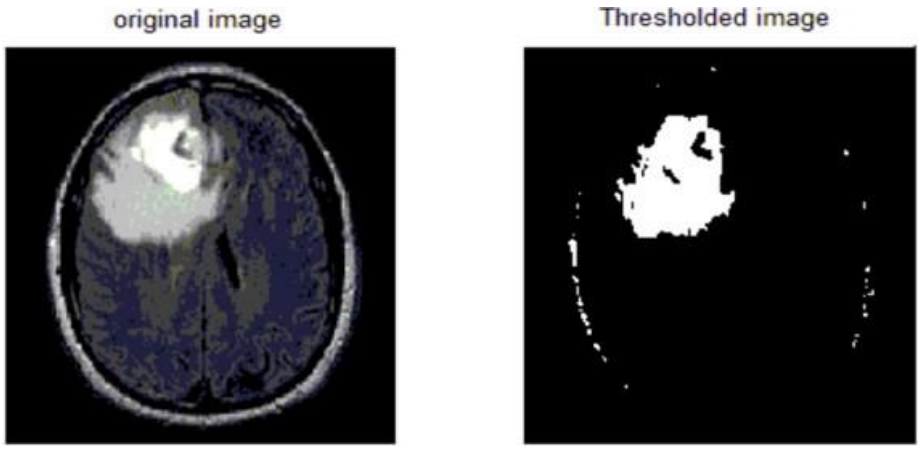


Figure 1.4 : : Résultat de la segmentation par seuillage sur une image IRM cérébrale [14].

b) La croissance de région (Region Growing)

La méthode de croissance de région [15] débute à partir d'un ou plusieurs pixels semences (seed points) sélectionnées dans la zone tumorale. Les pixels voisins sont ensuite ajoutés à la région si leurs caractéristiques (intensité, similarité statistique) satisfont un critère donne.

Bien que cette méthode permette une segmentation relativement cohérente, elle dépend fortement du choix initial des semences et du critère d'homogénéité, ce qui limite sa robustesse inter-patients.

La figure suivante illustre le processus de segmentation par croissance de régions, depuis l'image originale jusqu'au résultat final après filtrage (Figure 1.5).

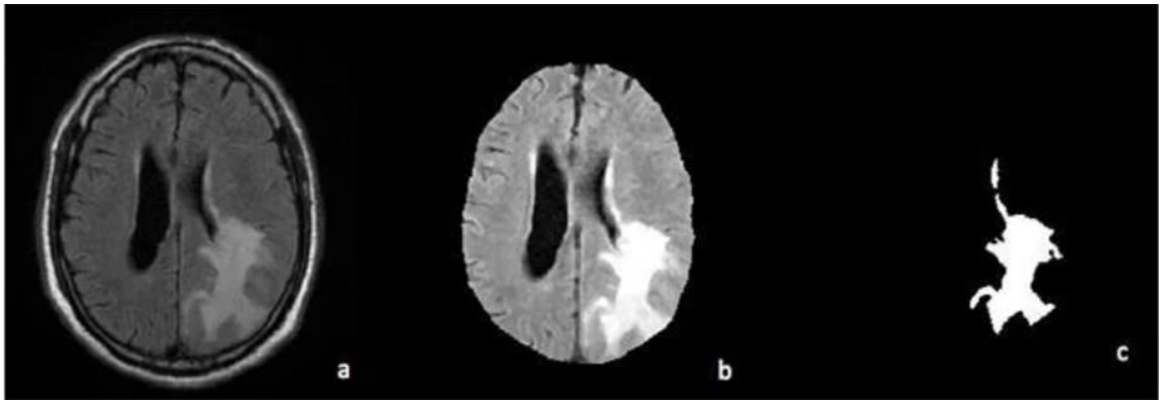


Figure 1.5 : : Résultat de la segmentation par croissance de régions sur une image IRM cérébrale : (a) Image originale, (b) Image filtrée, (c) Segmentation par croissance de régions [15].

c) La méthode Watershed

L'algorithme watershed [16] repose sur une interprétation topographique de l'image, considérée comme une surface ou les intensités représentent des altitudes. Les lignes de partage des eaux délimitent les régions segmentées.

Appliquée aux IRM cérébrales, cette méthode souffre d'un problème majeur de sur-segmentation, en particulier en présence de bruit ou de faibles variations d'intensité. Exemple de segmentation par Watershed illustrant la décomposition de l'image en régions (Figure 1.6).

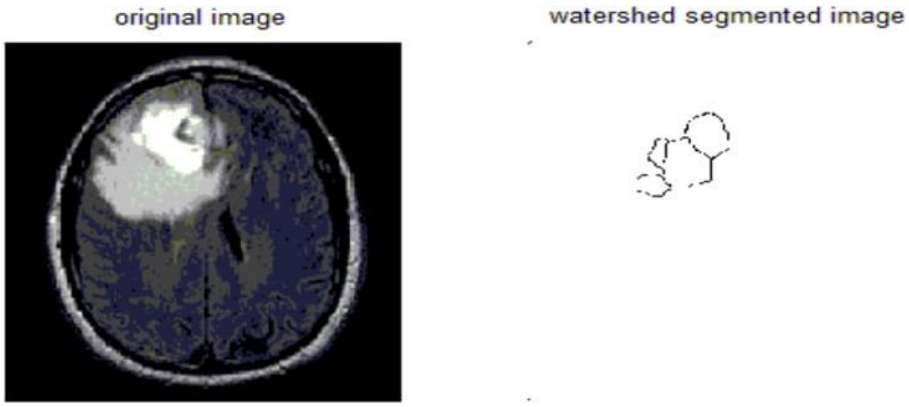


Figure 1.6 : Résultat de la segmentation par watershed sur une image IRM cérébrale [16].

d) Les contours actifs(Snakes)

Les méthodes de contours actifs, également appelés Snakes, utilisent une courbe déformable qui évolue sous l'effet de forces internes(régularité)et externes (gradient d'image) afin d'épouser les frontières tumorales [17].

Bien que plus sophistiquée que les méthodes précédentes, cette approche nécessite une initialisation proche de la frontière réelle et peut échouer lorsque le contraste tumoral est faible ou lorsque les contours sont irréguliers, comme dans le cas des gliomes infiltrants.

La figure suivante illustre la segmentation par contours actifs (snakes) appliquée à une image médicale (Figure 1.7).

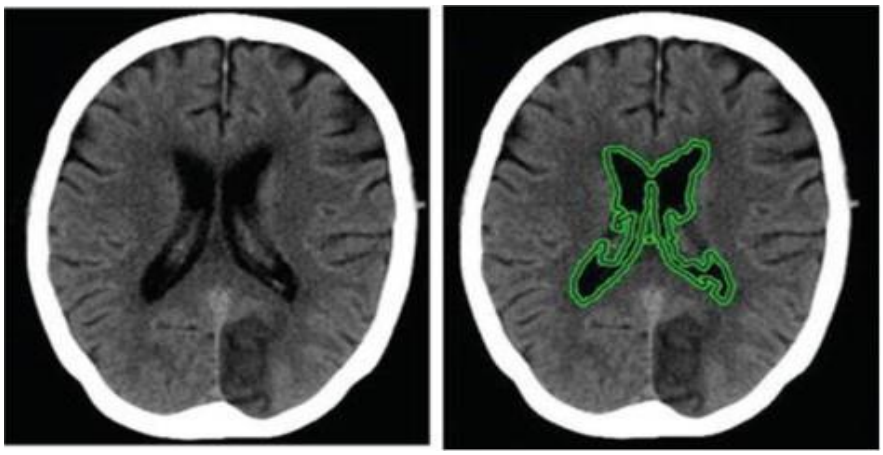


Figure 1.7 : Segmentation d'une image cérébrale CT utilisant la méthode des contours actifs [17].

e) Le Fuzzy C-Means(FCM)

Le Fuzzy C-Means est une méthode de classification non supervisée basée sur la logique floue. Chaque pixel appartient à plusieurs classes avec un degré d'appartenance variable, ce qui est particulièrement adapté aux frontières floues des tumeurs cérébrales.

Cependant, cette méthode reste sensible au bruit et ne prend pas toujours en compte les informations spatiales, ce qui peut conduire à des segmentations incohérentes dans les zones hétérogènes [12].

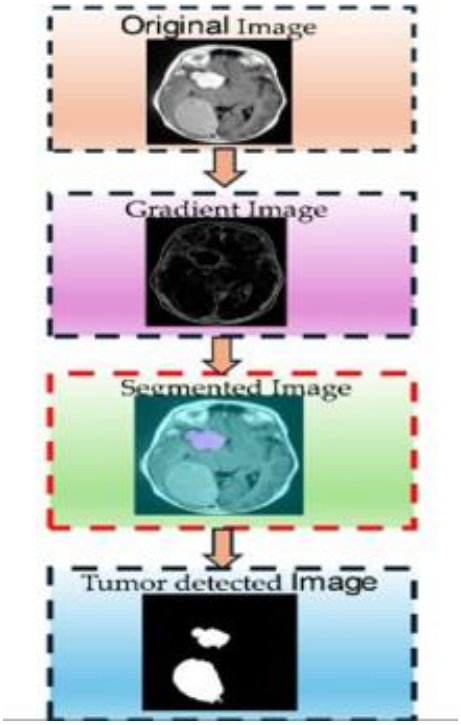


Figure 1.8 : Résultat de segmentation par Fuzzy C-Means sur une image IRM cérébrale [12].

La figure suivante présente le déroulement complet de la méthode proposée, depuis l'image originale jusqu'à l'obtention des contours finaux des régions cérébrales et tumorales (Figure 1.9).

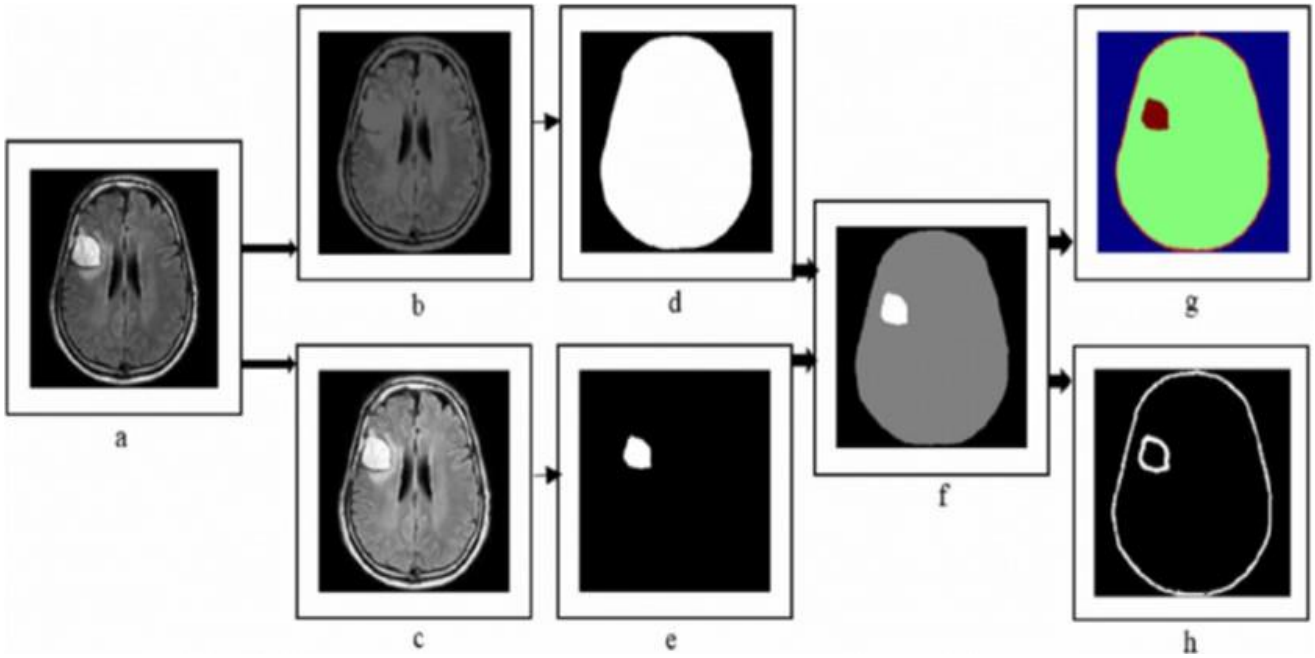


Figure 1.9 : Combinaison des méthodes Fuzzy C-Means et seuillage pour la segmentation d'images cérébrales [12].

Exemple de fonctionnement de la méthode proposée :

(a) image originale,

(b) et (c) images après prétraitement,

(d) résultat après segmentation par FCM,

(e) résultat après seuillage,

(f) combinaison des deux segmentations,

(g) image combinée avec carte de couleurs,

(h) carte finale des contours des régions cérébrales et tumorales.

Performances et limites des approches classiques :

Les performances de ces méthodes traditionnelles sont généralement évaluées à l'aide du coefficient de Dice. Dans la majorité des études, les valeurs obtenues restent inferieures a 75%, notamment pour les tumeurs de haut grade caractérisées par une forte hétérogénéité et des zones infiltrantes diffuses [15]. Plusieurs limitations majeures peuvent être identifiées :

- Forte dépendance aux paramètres manuels.
- Sensibilité élevée au bruit et aux artefacts IRM.
- Faible capacité de généralisation entre différents patients.
- Difficulté à segmenter les gliomes de haut grade.
- Incapacité à intégrer simultanément des caractéristiques complexes (texture, contexte global, multi modalité IRM)

Ces contraintes ont mis en évidence la nécessité de méthodes plus robustes, capables d'apprendre automatiquement des représentations discriminantes à partir de grandes bases de données annotées.

1.3.1.2 *Segmentation des tumeurs cérébrales par Deep Learning*

La segmentation des tumeurs cérébrales par deep learning constitue aujourd'hui l'un des domaines les plus avancés en imagerie médicale. Les approches modernes reposent sur des architectures profondes capables d'exploiter les informations spatiales et contextuelles des images IRM multimodales. Les performances sont principalement évaluées sur la base BraTS (Brain Tumor Segmentation Challenge), qui constitue la référence internationale en segmentation tumorale cérébrale. Les modèles récents atteignent des performances élevées en Dice score, variant selon la complexité architecturale et les modalités utilisées [BraTS-MICCAI].

1) CNN (Réseaux de neurones convolutifs - Les réseaux de neurones convolutifs (CNN) constituent la base fondamentale de la vision par ordinateur en imagerie médicale. Ils fonctionnent en appliquant des filtres de convolution sur l'image afin d'extraire automatiquement des caractéristiques locales telles que les contours, textures et variations d'intensité. Chaque couche convolutionnelle

apprend des représentations de plus en plus abstraites, permettant de passer de simples pixels à des structures anatomiques complexes. Les opérations de pooling réduisent progressivement la dimension spatiale afin d'augmenter le champ réceptif du réseau. Les fonctions d'activation non linéaires comme ReLU permettent d'introduire de la complexité dans l'apprentissage[2].

Dans le contexte des IRM cérébrales, les CNN permettent de détecter les régions tumorales en apprenant des patterns discriminants à partir de grandes bases de données annotées. Toutefois, ils restent limités dans la capture du contexte global sans architectures avancées.

Les CNN fonctionnent par application successive de filtres de convolution permettant d'extraire des caractéristiques locales de plus en plus abstraites. Cette opération mathématique est définie par [19] :

$$(I * K)(x, y) = \sum_i \sum_j I(x - i, y - j)\, K(i, j) \tag{1.1}$$

où I représente l'image d'entrée et K le noyau de convolution.

Sur la base BraTS, les CNN classiques atteignent un Dice score compris entre 0.70 et 0.78 [27]. Ils constituent néanmoins la base de toutes les architectures modernes de segmentation.

Le principe de cette architecture est présenté dans la figure suivante (Figure 1.10).

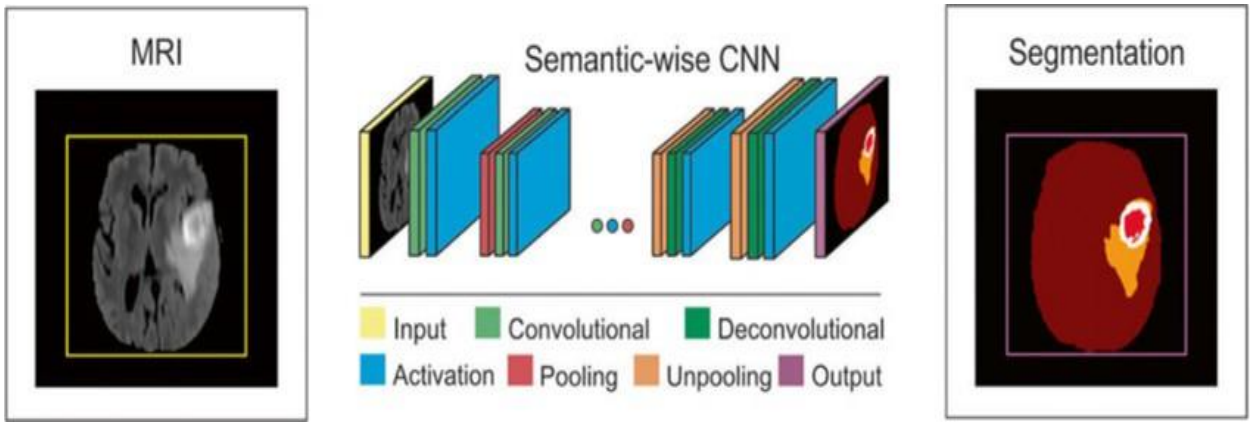


Figure 1.10 : Architecture d'un réseau de neurones convolutifs (CNN) appliqué à la segmentation médicale [2].

Avantages :

- Extraction automatique des caractéristiques
- Bonne efficacité sur images bruitées
- Architecture simple et rapide

Limites :

- Faible prise en compte du contexte global
- Dépendance aux données annotées

- Sensibilité aux variations d'intensité

Parmi les méthodes de segmentation fondées sur les CNN :

- Segmentation "patch-based " : Dans cette approche, l'image est divisée en petits patchs (sous-images). Chaque patch est ensuite passé individuellement dans le réseau pour prédire la classe du pixel central ou du patch entier[18].

- Avantages :

o Facilité d'entraînement sur des volumes de données limités.

o Moindre consommation de mémoire.

- Limites :

o Contexte local restreint, le modèle ne capture pas l'information globale de l'image.

o Sensibilité aux artefacts et variations d'intensités.

o Temps de traitement élevé pour des volumes 3D complets.

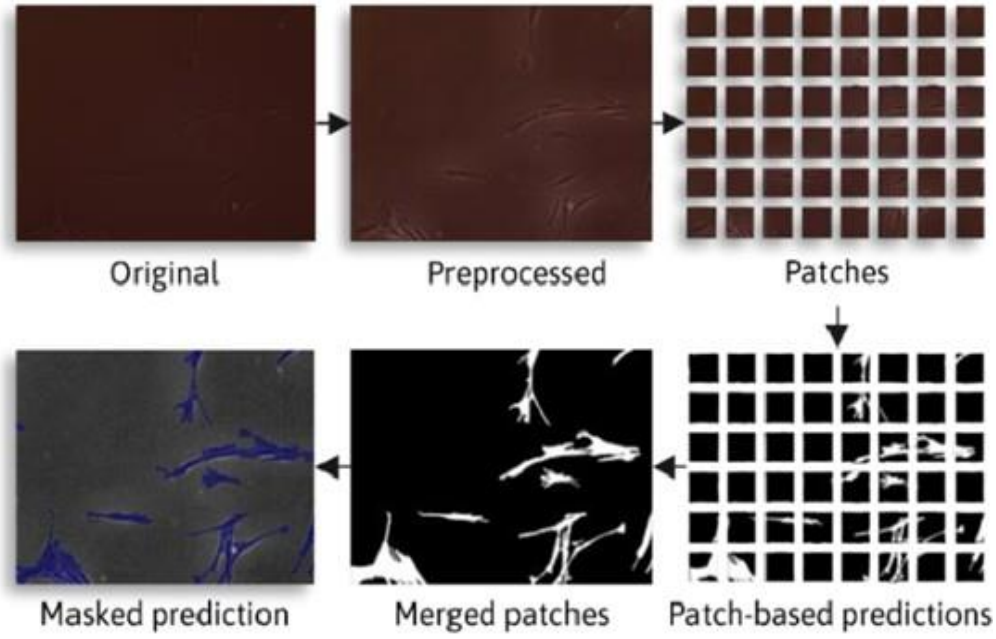


Figure 1.11 : Principe de la segmentation patch-based appliquée aux images IRM [18].

2) U-Net (2D et 3D) - U-Net est une architecture encoder–decoder spécialement conçue pour la segmentation biomédicale. L'encodeur [74] applique une série de convolutions et de pooling afin de réduire progressivement la dimension spatiale et extraire des caractéristiques profondes. Le décodeur effectue l'opération inverse en reconstruisant une carte de segmentation de même taille que l'image d'entrée. Les skip connections jouent un rôle essentiel en transférant les informations de bas niveau de l'encodeur vers le décodeur afin de préserver les détails spatiaux. le principe de cette architecture est présenté dans la figure suivante (Figure 1.12)[20].

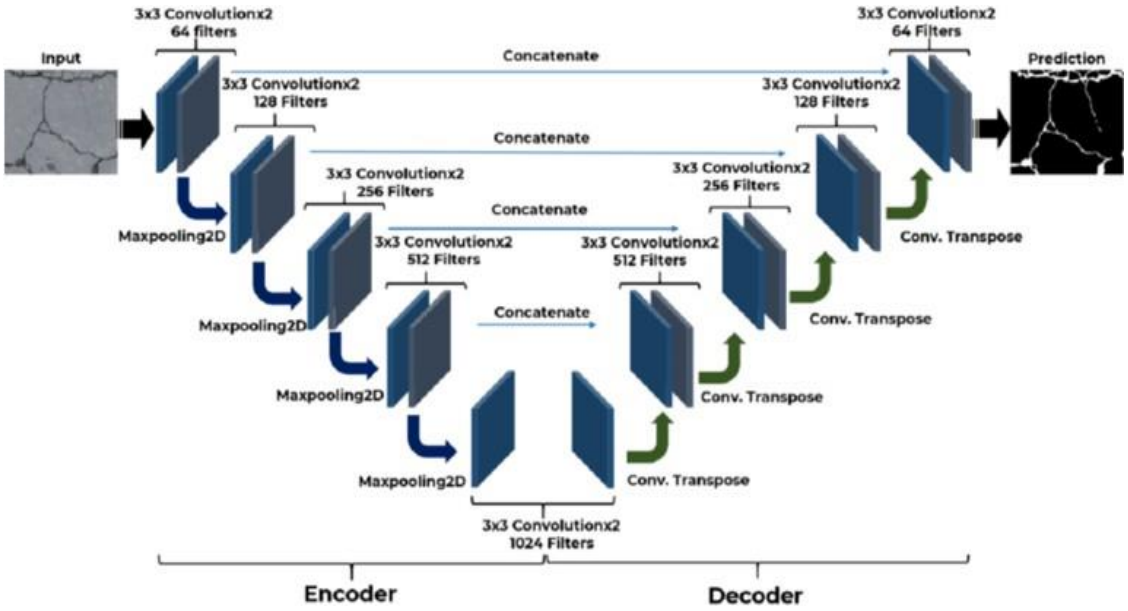


Figure 1.12 : : Schéma de l'architecture U-Net pour la segmentation d'images biomédicales [20].

La version 2D traite les coupes IRM indépendamment, tandis que la version 3D exploite directement les volumes complets, permettant une meilleure cohérence spatiale. Les convolutions 3D améliorent la compréhension du contexte volumique mais augmentent fortement la complexité computationnelle.

Cette approche permet d'exploiter directement les volumes d'images en trois dimensions, comme illustre dans la figure suivante (Figure 1.12)[21].

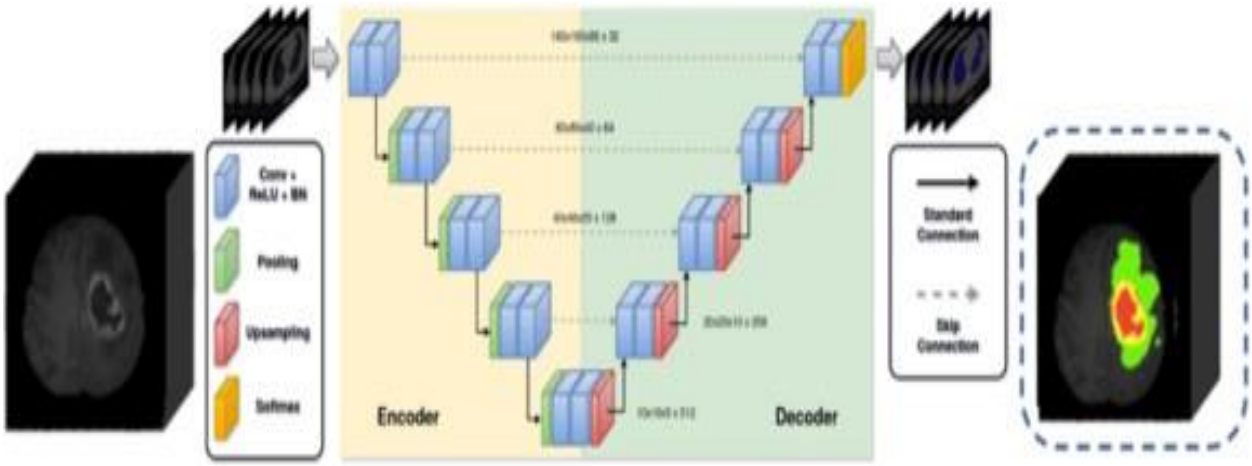


Figure 1.13 : Architecture 3D U-Net pour la segmentation médicale volumétrique [21].

U-Net est optimisé pour les données médicales où les structures sont irrégulières et peu contrastées. Sur BraTS, U-Net atteint environ 0.82 Dice en 2D [26] et jusqu'à 0.87–0.89 en 3D [26]. Cette architecture reste une référence en segmentation médicale.

Avantages :

- Très bonne précision en segmentation médicale
- Préservation des détails anatomiques
- Architecture robuste et éprouvée

Limites :

- U-Net 2D ignore la cohérence volumique
- U-Net 3D très coûteux en GPU
- Sensible aux annotations bruitées

3) U-Net++ - L'architecture U-Net++ est une amélioration avancée du U-Net classique introduite par Zhou et al. [33], visant à réduire le décalage sémantique entre les caractéristiques de l'encodeur [74] et du décodeur. Comme illustré dans la figure correspondante, cette architecture repose sur une structure de skip connections imbriquées (nested skip pathways), remplaçant les connexions directes simples du U-Net par une série de blocs convolutionnels intermédiaires.

Contrairement au U-Net classique où les features de l'encodeur sont directement concaténées avec celles du décodeur, U-Net++ introduit des chemins intermédiaires constitués de plusieurs couches convolutionnelles successives. Ces chemins permettent d'affiner progressivement les représentations avant la fusion, ce qui réduit significativement l'écart sémantique entre les deux branches du réseau. Chaque nœud intermédiaire représente une version transformée des caractéristiques issues de niveaux inférieurs, permettant une meilleure homogénéisation des informations multi-niveaux[22].

Un autre élément clé de cette architecture est l'utilisation de la *deep supervision*. Cette stratégie consiste à superviser plusieurs sorties intermédiaires du réseau (notamment les nœuds de la première ligne du décodeur), et non uniquement la sortie finale. Ainsi, les prédictions intermédiaires participent au calcul de la fonction de perte globale, ce qui améliore la convergence du modèle et sa capacité de généralisation. Cette approche permet également de sélectionner dynamiquement le niveau de profondeur optimal lors de l'inférence, réduisant ainsi le temps de calcul sans dégrader significativement les performances.

De plus, U-Net++ permet une forme de "pruning" architectural lors de l'inférence. Selon le niveau de supervision utilisé, différentes profondeurs du réseau peuvent être activées (L1, L2, L3, L4), offrant un compromis entre précision et coût computationnel. Les expériences menées par les auteurs montrent qu'une version intermédiaire (L3) permet de réduire le temps d'inférence d'environ 32% tout en conservant des performances proches du modèle complet, avec une perte de performance inférieure à 1% en IoU [26].

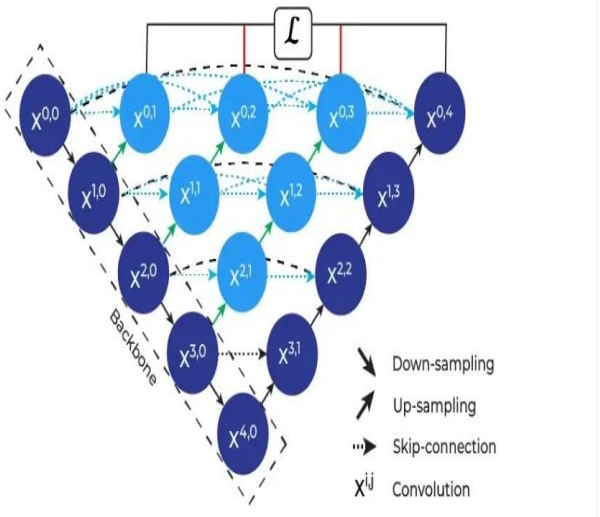


Figure 1.14 : Architecture U-Net++ avec skip connections imbriquées et deep supervision [28]

Sur BraTS 2019, U-Net++ atteint environ 0.85 à 0.89 Dice . Il est souvent utilisé comme amélioration directe de U-Net [26].

Avantages :

- Meilleure fusion multi-échelle
- Amélioration des petites lésions
- Réduction du gap encodeur-décodeur

Limites :

- Complexité élevée
- Temps d'entraînement important
- Consommation mémoire élevée

4) Représentations multi-échelles et Feature Pyramid Network (FPN) - La segmentation des structures anatomiques complexes, notamment les tumeurs cérébrales, nécessite une prise en compte des variations d'échelle présentes dans les images médicales. En effet, les lésions peuvent apparaître sous différentes tailles et formes, ce qui rend indispensable l'utilisation de représentations multi-résolutions. Les approches basées sur les pyramides de caractéristiques visent précisément à améliorer la robustesse des modèles face à cette variabilité spatiale [22].

La figure associée illustre différentes stratégies de construction de représentations multi-échelles, allant des méthodes classiques basées sur la pyramide d'images jusqu'aux architectures profondes comme le Feature Pyramid Network (FPN) [76].

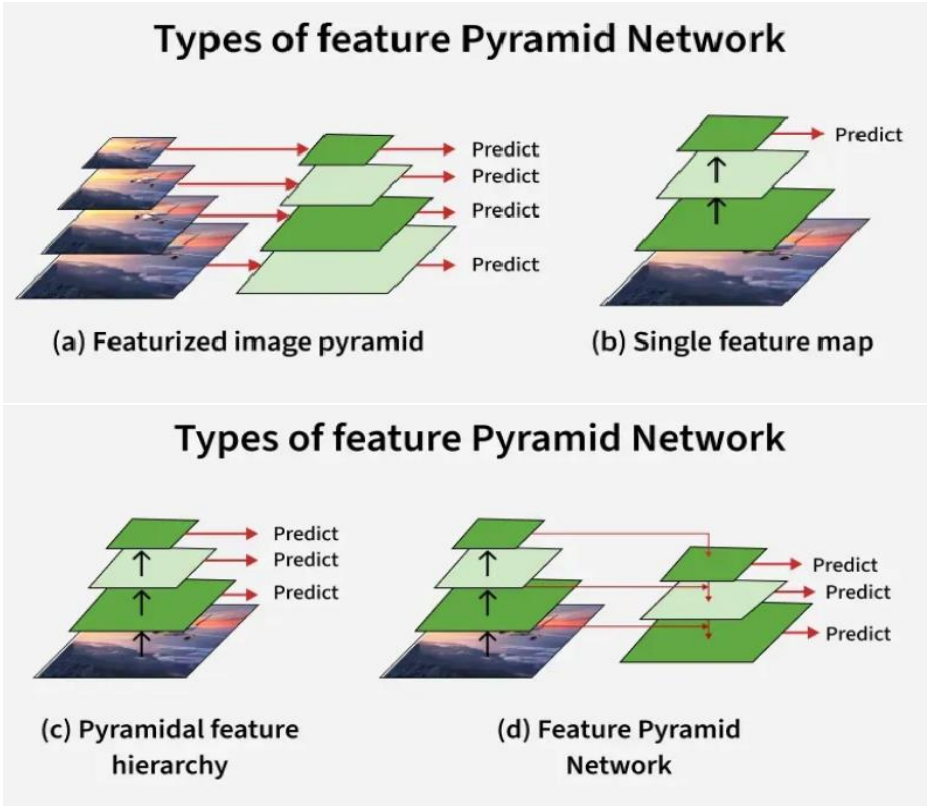


Figure 1.15 : Comparaison des approches multi-échelles (Image Pyramid, Single Scale, Feature Hierarchy, FPN) [22]

- Image Pyramid

L'approche par pyramide d'images consiste à redimensionner l'image d'entrée à plusieurs résolutions distinctes. Chaque version de l'image est ensuite traitée indépendamment par un extracteur de caractéristiques. Cette méthode permet de capturer des informations multi-échelles explicites, mais elle augmente fortement le coût computationnel, car chaque échelle nécessite un passage complet dans le réseau. De plus, il n'existe pas de mécanisme de fusion entre les différentes échelles, ce qui limite l'exploitation globale des informations.

- Single Scale Feature Extraction

Dans cette approche, l'image est traitée à une seule échelle fixe. Le réseau extrait les caractéristiques une seule fois sans tenir compte explicitement des variations de taille des objets. Cette méthode est plus simple et plus rapide, mais elle est limitée lorsqu'il s'agit de détecter des structures de tailles très variables, comme les tumeurs cérébrales de formes irrégulières.

- Pyramidal Feature Hierarchy

Les réseaux de neurones convolutifs profonds génèrent naturellement une hiérarchie de caractéristiques. Les premières couches capturent des informations locales (bords, textures), tandis que les couches profondes extraient des informations plus globales et sémantiques. Cette hiérarchie peut être exploitée pour représenter implicitement plusieurs niveaux d'échelle, sans nécessiter de redimensionnement explicite de l'image.

- Feature Pyramid Network (FPN)

Le Feature Pyramid Network (FPN) constitue une amélioration significative des approches multi-échelles. Il combine les avantages de la hiérarchie intrinsèque des CNN avec un mécanisme explicite de fusion des caractéristiques. L'architecture repose sur une structure encodeur–décodeur inversée, où les features de haut niveau (sémantiques) sont progressivement remontées vers les niveaux inférieurs via un chemin top-down. Des connexions latérales permettent de fusionner ces informations avec les features issues du backbone CNN.

Ce mécanisme permet d'obtenir des représentations riches et cohérentes à toutes les échelles, améliorant ainsi la détection des petites et grandes lésions dans les images médicales. FPN est particulièrement efficace dans les tâches de segmentation où les objets présentent des variations importantes de taille et de contraste [22].

Avantages du FPN :

- Excellente gestion des objets multi-échelles
- Fusion efficace des informations locales et globales
- Amélioration de la détection des petites structures
- Compatible avec différents backbones CNN

Limites du FPN :

- Architecture plus complexe que les CNN standards

- Consommation mémoire élevée
- Sensibilité au choix du backbone

Les approches multi-échelles, et en particulier le Feature Pyramid Network, jouent un rôle essentiel dans l'amélioration des performances des modèles de segmentation médicale. Leur capacité à intégrer des informations à différentes résolutions permet une meilleure robustesse face à la variabilité des structures anatomiques, ce qui est crucial dans l'analyse des tumeurs cérébrales [35].

Sur BraTS, FPN atteint environ 0.83 à 0.88 Dice [27].

5) DeepLabv3+ pour la segmentation sémantique des tumeurs cérébrales - DeepLabv3+ est une architecture avancée de segmentation sémantique proposée par Chen et al. [24], largement utilisée dans les tâches de vision par ordinateur et particulièrement efficace pour l'analyse des images médicales. Cette architecture améliore les performances des modèles classiques en intégrant des mécanismes capables de capturer simultanément les informations locales et globales à différentes échelles.

La figure associée présente l'architecture globale de DeepLabv3+, incluant le backbone CNN, le module ASPP (Atrous Spatial Pyramid Pooling) ainsi que le décodeur de raffinement utilisé pour améliorer la précision des contours de segmentation.

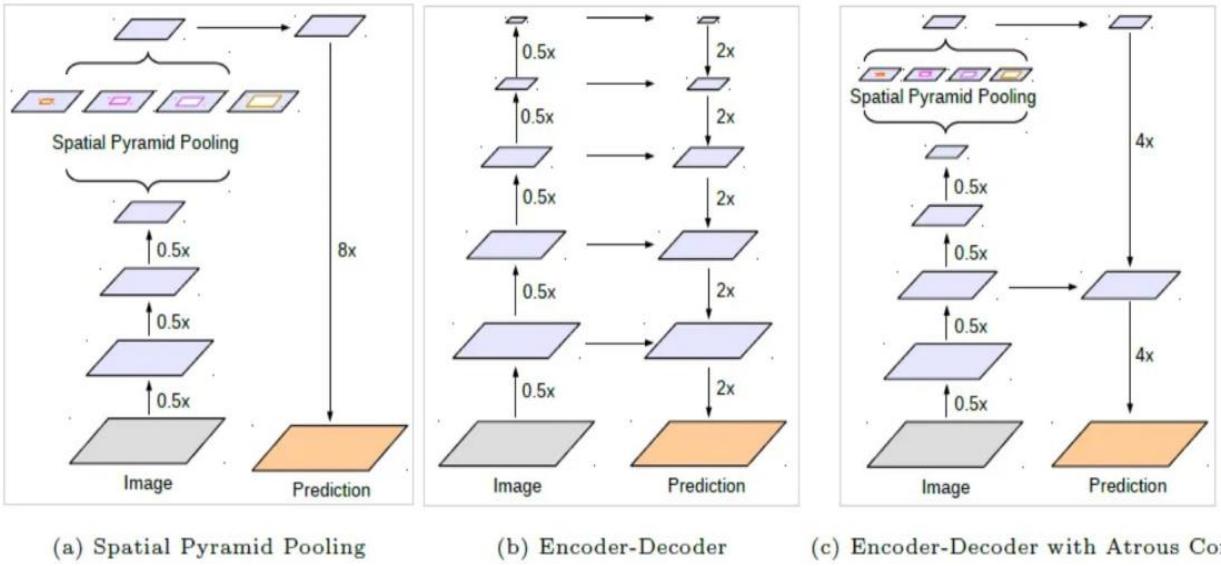


Figure 1.16 : Architecture DeepLabv3+ avec backbone, ASPP et décodeur de segmentation [24]

Le fonctionnement de DeepLabv3+ repose principalement sur les convolutions dilatées (atrous convolutions), qui permettent d'augmenter le champ réceptif du réseau sans augmenter le nombre de paramètres ni réduire la résolution spatiale des cartes de caractéristiques. Cette propriété est particulièrement importante en imagerie médicale, où la préservation des détails fins est essentielle pour la détection des tumeurs cérébrales.

Le module central de l'architecture est le Atrous Spatial Pyramid Pooling (ASPP). Ce module applique plusieurs convolutions dilatées en parallèle avec différents taux de dilatation, ce qui permet de capturer des informations multi-échelles dans une même représentation. Les sorties de ces convolutions sont ensuite concaténées afin de former une représentation riche intégrant à la fois des informations locales et globales.

Le décodeur de DeepLabv3+ joue un rôle essentiel dans l'amélioration des contours de segmentation. Il prend en entrée les caractéristiques extraites par l'ASPP ainsi que les informations de bas

niveau provenant du backbone afin de restaurer progressivement la résolution spatiale. Cette étape permet d'obtenir des segmentations plus précises, notamment au niveau des frontières des lésions tumorales[24].

Avantages de DeepLabv3+ :

- Excellente capture du contexte global et local
- Très bonne performance sur les structures complexes
- Utilisation efficace des convolutions dilatées
- Haute précision de segmentation en imagerie médicale

Limites de DeepLabv3+ :

- Coût computationnel élevé
- Temps d'entraînement important
- Complexité de réglage des hyperparamètres (taux de dilatation)

DeepLabv3+ constitue une architecture performante pour la segmentation des tumeurs cérébrales, grâce à sa capacité à intégrer des informations multi-échelles via le module ASPP et à améliorer la précision des contours grâce à son décodeur. Cette combinaison en fait une méthode robuste et largement utilisée dans les applications médicales modernes [24].

DeepLabv3+ est robuste mais computationnellement coûteux. Sur BraTS, il atteint environ 0.83 à 0.88 Dice [27].

6) Pipeline du nnU-Net pour la segmentation automatique des images médicales Le nnU-Net [25]est un framework auto-configurable conçu pour la segmentation d'images médicales, reposant sur une stratégie entièrement automatisée d'adaptation aux caractéristiques des données. Contrairement aux architectures classiques nécessitant un réglage manuel des hyperparamètres, nnU-Net utilise une approche basée sur l'analyse des propriétés du dataset afin de générer automatiquement un pipeline optimal de segmentation [25].

La figure associée illustre le fonctionnement global du pipeline nnU-Net, depuis l'analyse des données jusqu'à la génération des prédictions finales.

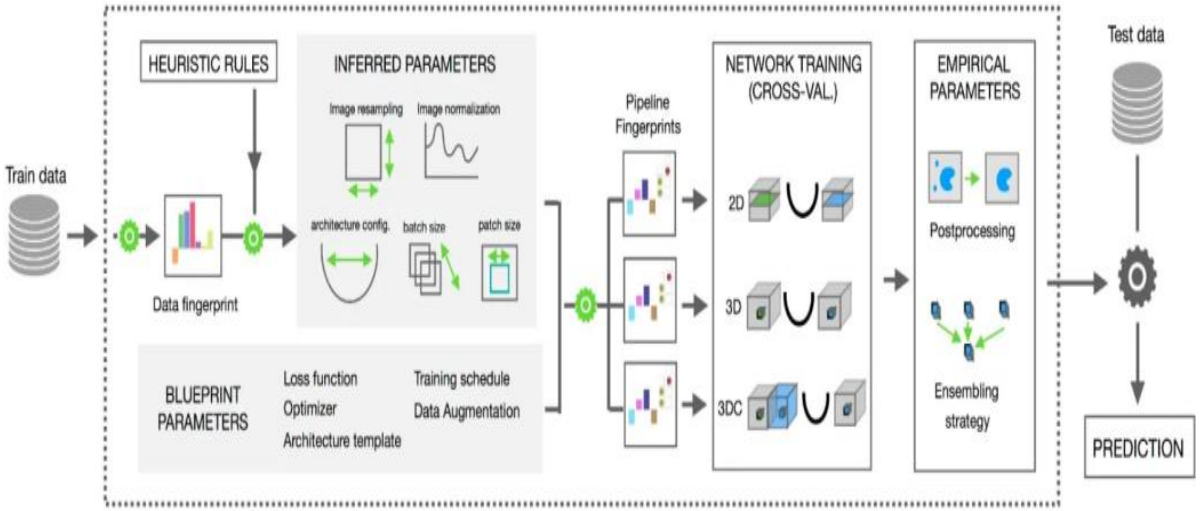


Figure 1.17 : Pipeline complet du framework nnU-Net pour la segmentation des images IRM cérébrales [25]

Le fonctionnement du nnU-Net repose sur une première étape appelée *data fingerprinting*. Cette étape consiste à analyser les caractéristiques intrinsèques des données d'entrée, telles que la résolution spatiale, la taille des volumes, le nombre de modalités (IRM T1, T2, FLAIR) ainsi que la distribution des classes. Ces informations permettent de décrire précisément la structure du dataset et constituent la base de la configuration automatique du modèle.

À partir de ce *data fingerprint*, le framework génère ensuite un ensemble de paramètres appelés *pipeline fingerprint*. Ces paramètres incluent à la fois des éléments prédéfinis (blueprint parameters), tels que la fonction de perte, l'optimiseur et l'architecture de base du réseau U-Net, ainsi que des paramètres inférés automatiquement, comme la stratégie de normalisation, la taille des patchs, la méthode de rééchantillonnage et la taille des batchs. L'ensemble de ces paramètres permet de construire un pipeline entièrement adapté aux données disponibles[25].

Le nnU-Net entraîne ensuite plusieurs configurations de réseaux, notamment des modèles 2D, 3D full resolution et 3D cascade. Chaque configuration est optimisée selon les paramètres générés automatiquement et évaluée indépendamment sur les données d'entraînement. Une étape d'ensemble (ensemble learning) est ensuite appliquée afin de combiner les prédictions des différents modèles, ce qui permet d'améliorer la robustesse et la précision globale de la segmentation[73],[74].

Enfin, une phase de post-traitement est utilisée afin de sélectionner la meilleure configuration en fonction du coefficient de Dice moyen obtenu. Le modèle retenu est ensuite utilisé pour effectuer les prédictions finales sur les données de test.

Avantages du nnU-Net :

- Automatisation complète du pipeline de segmentation
- Adaptation automatique aux caractéristiques du dataset
- Excellentes performances sans réglage manuel
- Robustesse grâce à l'ensemble de modèles

Limites du nnU-Net :

- Temps d'entraînement élevé
- Forte consommation de ressources GPU
- Complexité du pipeline global

Le nnU-Net représente une approche de référence en segmentation médicale, grâce à sa capacité unique à adapter automatiquement l'ensemble du pipeline d'apprentissage. Cette automatisation permet d'obtenir des performances élevées et stables sur différents types de données médicales, notamment dans le cadre de la segmentation des tumeurs cérébrales [73].

Cette approche permet une optimisation automatique sans intervention humaine. Elle est considérée comme une référence state-of-the-art en segmentation médicale. Sur BraTS 2020–2021, nnU-Net atteint environ 0.90 à 0.93 Dice [26].

Analyse et discussion des configurations du nnU-Net pour la segmentation des tumeurs cérébrales

Le framework nnU-Net constitue une avancée majeure dans la segmentation automatique des images médicales, car il ne propose pas une architecture unique mais un ensemble de configurations adaptatives. Ces configurations permettent d'exploiter efficacement les données IRM en fonction de leurs caractéristiques spatiales, de leur résolution et de leur dimensionnalité. Dans cette section, nous analysons en détail les trois principales configurations du nnU-Net, à savoir la configuration 2D, la configuration 3D full resolution et la configuration 3D cascade, afin de mettre en évidence leurs principes de fonctionnement, leurs avantages et leurs limites respectives [25].

- Configuration nnU-Net 2D La configuration nnU-Net 2D repose sur le traitement des images IRM coupe par coupe, où chaque slice est considérée indépendamment comme une image bidimensionnelle. Le réseau applique des convolutions 2D classiques afin d'extraire les caractéristiques locales présentes dans chaque coupe. Cette approche simplifie fortement la complexité du modèle et permet un entraînement efficace même avec des ressources matérielles limitées. Le pipeline inclut automatiquement des étapes de normalisation, de rééchantillonnage et de sélection des patchs les plus informatifs. Toutefois, l'absence de prise en compte de la dimension volumique entraîne une perte d'information sur la continuité entre les différentes coupes IRM. Cela peut affecter la cohérence spatiale des segmentations obtenues, notamment pour les structures tumorales complexes.

Sur le plan expérimental, la configuration 2D du nnU-Net obtient généralement un Dice score compris entre 0.84 et 0.87 sur la base de données BraTS, ce qui reste compétitif malgré sa simplicité [27].

Avantages :

- Faible consommation mémoire GPU
- Temps d'entraînement réduit
- Adapté aux petits jeux de données
- Facile à implémenter et stabiliser

Limites :

- Absence de contexte volumique 3D
- Incohérence entre les coupes adjacentes
- Performance limitée pour les structures complexes

- Configuration nnU-Net 3D full resolution : La configuration nnU-Net 3D full resolution exploite directement les volumes IRM en trois dimensions, en appliquant des convolutions 3D afin de capturer les relations spatiales entre les différentes coupes. Cette approche permet de préserver l'intégrité anatomique des structures cérébrales et d'améliorer significativement la cohérence des segmentations. Le réseau traite l'ensemble du volume ou de larges sous-volumes, ce qui permet une meilleure compréhension globale de la morphologie tumorale. Le pipeline conserve les mêmes principes

d'auto-configuration que le nnU-Net, notamment l'adaptation automatique de la normalisation et de la taille des patchs.

Cependant, cette configuration nécessite une grande capacité mémoire GPU, ce qui limite son utilisation sur des machines standards. Le temps d'entraînement est également plus élevé en raison de la complexité des convolutions 3D.

Sur BraTS 2020, cette configuration atteint des performances comprises entre 0.90 et 0.92 en Dice score, ce qui en fait une méthode très compétitive pour la segmentation volumique [74].

Avantages :

- Exploitation complète du contexte 3D
- Meilleure cohérence anatomique
- Haute précision de segmentation
- Très bonnes performances sur BraTS

Limites :

- Forte consommation mémoire GPU
- Temps d'entraînement élevé
- Complexité computationnelle importante

- Configuration nnU-Net 3D cascade (coarse-to-fine) La configuration nnU-Net cascade 3D repose sur une stratégie en deux étapes visant à améliorer la précision tout en réduisant la complexité du traitement des volumes volumineux. Dans une première étape, un réseau à basse résolution est entraîné afin de localiser globalement la région tumorale dans l'ensemble du volume IRM. Cette étape permet de réduire la zone d'intérêt et d'identifier les régions pertinentes. Dans une seconde étape, un second réseau est appliqué à haute résolution uniquement sur les régions détectées afin d'affiner la segmentation et de récupérer les détails fins des contours tumoraux.

Cette approche hiérarchique permet de combiner efficacité computationnelle et précision élevée, tout en réduisant la charge mémoire par rapport à une approche 3D full resolution unique. Elle est particulièrement adaptée aux volumes IRM de grande taille et aux tumeurs complexes présentant des variations importantes de forme et de localisation[25].

Sur la base BraTS 2021, la configuration cascade du nnU-Net atteint généralement les meilleures performances avec un Dice score pouvant atteindre 0.91 à 0.93, ce qui représente l'état de l'art actuel en segmentation cérébrale [27].

Avantages :

- Très haute précision de segmentation
- Réduction de la complexité globale du problème
- Meilleure détection des petites lésions

- Adapté aux grands volumes IRM

Limites :

- Temps de calcul élevé
- Pipeline complexe en deux étapes
- Besoin important en ressources GPU

Globalement, les trois configurations du nnU-Net illustrent un compromis entre précision, complexité et ressources computationnelles. La configuration 2D privilégie la rapidité et la simplicité, tandis que la configuration 3D full resolution privilégie la cohérence spatiale et la précision. La configuration cascade représente une approche intermédiaire optimisée, combinant localisation globale et raffinement local. Ce cadre adaptatif explique pourquoi le nnU-Net est considéré comme une référence dans les challenges internationaux de segmentation médicale [25].

Le nnU-Net représente une approche de référence en segmentation médicale, grâce à sa capacité unique à adapter automatiquement l'ensemble du pipeline d'apprentissage. Cette automatisation permet d'obtenir des performances élevées et stables sur différents types de données médicales, notamment dans le cadre de la segmentation des tumeurs cérébrales [25].

Cette approche permet une optimisation automatique sans intervention humaine. Elle est considérée comme une référence state-of-the-art en segmentation médicale.

Sur BraTS 2020–2021, nnU-Net atteint environ 0.90 à 0.93 Dice [26].

7) Fonctions de perte et métriques d'évaluation La performance d'un modèle de segmentation médicale dépend fortement de deux éléments fondamentaux : les fonctions de perte utilisées lors de l'apprentissage et les métriques d'évaluations employées pour quantifier la qualité des prédictions.

Les fonctions de perte permettent d'optimiser le réseau pendant l'entraînement, tandis que les métriques permettent d'évaluer sa performance de manière objective sur des données non vues[28].

a) Fonctions de perte

Dice Loss

La Dice Loss est l'une des fonctions les plus utilisées en segmentation médicale, en particulier dans les cas de déséquilibre important entre les classes (ex : petite zone tumorale vs grand volume cérébral sain).

Elle est basée sur le coefficient de Dice et mesure le chevauchement entre la prédiction P et la vérité terrain T :

$$Dice\ Loss = -\frac{2|P \cap T|}{|P| + |T|} \tag{1.2}$$

Elle favorise directement le recouvrement des régions segmentées [28].

Cross-Entropy Loss La Cross-Entropy est une fonction de perte standard utilisée dans les problèmes classification voxel par voxel.

Elle mesure la distance entre la distribution prédite et la distribution réelle.

Elle est efficace lorsque les classes sont bien équilibrées, mais moins performante seule en cas de forte déséquilibre [28].

Focal Loss

La Focal Loss est une extension de la Cross-Entropy, introduite pour donner plus de poids aux exemples difficiles à classer.

Elle est particulièrement utile dans la segmentation des tumeurs, où certaines régions sont très petites ou peu visibles [30].

IoU Loss (Jaccard Loss)

La IoU Loss mesure le rapport entre l'intersection et l'union des régions prédites et réelles :

$$IoU = \frac{|P \cap T|}{|P \cup T|} \tag{1.3}$$

Elle est complémentaire à la Dice Loss et renforce la cohérence globale de la segmentation[29].

b) Métriques d'évaluation

Les métriques permettent d'évaluer la qualité du modèle après entraînement. Dans le domaine médical, plusieurs indicateurs sont utilisés afin d'obtenir une évaluation complète[28].

Dice Coefficient

- Mesure principale en segmentation médicale.
- Il quantifie le recouvrement entre la prédiction et la vérité terrain.
- Plus il est proche de 1, meilleure est la segmentation.

IoU (Intersection over Union)

- Mesure la proportion d'intersection entre prédiction et vérité terrain par rapport à leur union.
- Plus strict que Dice.

HD95 (Hausdorff Distance at 95%)

- Mesure la distance maximale entre les contours de la prédiction et de la vérité terrain, en ignorant les 5% des distances les plus extrêmes.
- Elle évalue la précision des bords et la cohérence géométrique de la segmentation.
- Plus la valeur est faible, meilleure est la segmentation, car les contours prédits sont proches des contours réels.

Ces différentes fonctions de perte et métriques jouent un rôle complémentaire dans l'optimisation. Le choix approprié de ces éléments est essentiel pour garantir des performances robustes, en particulier dans des contextes médicaux où le déséquilibre des classes est important.

1.3.1.3 *Prétraitement des images*

Le prétraitement constitue une étape essentielle visant à améliorer la qualité des données avant leur exploitation par les modèles d'intelligence artificielle. Il permet de réduire les artefacts, homogénéiser les intensités et garantir la cohérence des données[31].

1) Normalisation des intensités :

Les images IRM présentent souvent des variations d'intensité dues aux différences d'appareils ou de protocoles d'acquisition. La normalisation permet de standardiser ces intensités.

Une méthode classique est la normalisation min-max [32], définie par :

$$I_{\text{norm}} = \frac{I - I_{\text{min}}}{I_{\text{max}} - I_{\text{min}}} \tag{1.4}$$

Où :

- I : intensité du pixel[32],
- I_{max}, I_{min} : intensité minimale et maximale.

Une autre approche consiste à utiliser la standardisation (z-score) [33] :

$$I_{\text{norm}} = \frac{I - \mu}{\sigma} \tag{1.5}$$

Où [33] :

- μ : moyenne des intensités,
- σ : écart-type.

Cette étape permet de rendre les données comparables et d'améliorer la convergence des modèles de deep learning.

La figure ci-dessous illustre l'effet de la normalisation des intensités sur une image IRM, en comparant l'image originale avec l'image après standardisation. On observe une homogénéisation des valeurs d'intensité, ce qui permet de réduire les variations liées aux conditions d'acquisition et d'améliorer la qualité des données pour les étapes de traitement ultérieures.

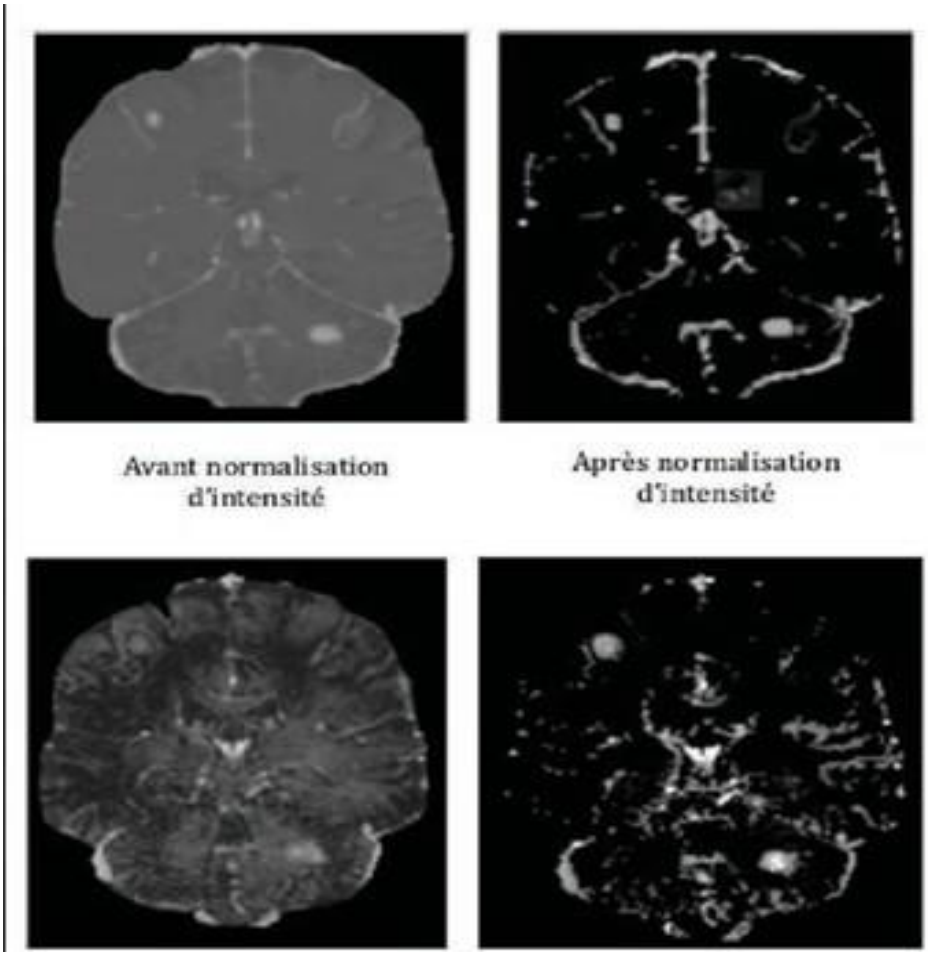


Figure 1.18 : Effet de la normalisation des intensités sur une image IRM [33]

2) Réduction du bruit :

Les images IRM sont souvent affectées par différents types de bruit qui peuvent dégrader la qualité des images et compliquer la détection des structures anatomiques. Afin d'améliorer la qualité des données, plusieurs techniques de filtrage peuvent être utilisées[34].

Filtrage gaussien

Le filtre gaussien est une méthode classique de lissage qui repose sur la convolution de l'image avec un noyau gaussien [35] :

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1.6}$$

Ce filtre permet de réduire efficacement le bruit aléatoire en moyennant les intensités des pixels voisins.

Limite : Cependant, cette méthode présente l'inconvénient de flouter les contours, ce qui peut entraîner une perte d'informations importantes, notamment au niveau des limites tumorales.

Filtrage médian

Le filtre médian est une technique non linéaire qui remplace chaque pixel par la médiane des valeurs de son voisinage. Il est particulièrement efficace pour éliminer le bruit impulsionnel.

Limite : Toutefois, dans le cas des images IRM, ce type de bruit est moins fréquent, ce qui rend cette méthode moins adaptée. De plus, elle peut altérer certaines structures fines[36].

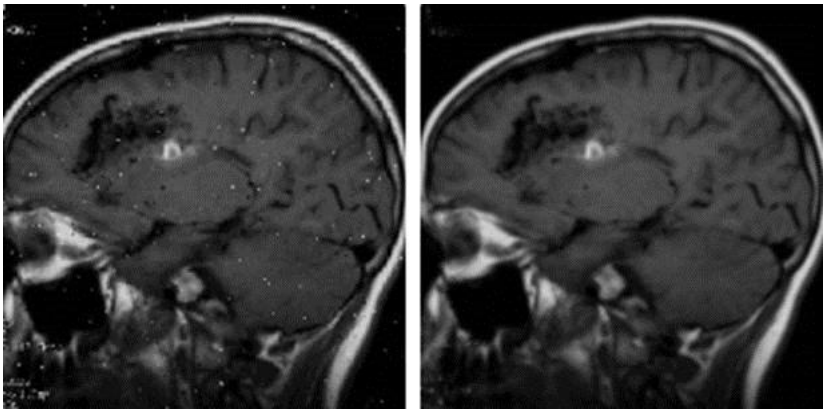


Figure 1.19 : Filtrage médian et gaussien suivi de la segmentation level-set [36]

3) Recalage et alignement des images (Registration) :

Le recalage d'images [37], également appelé registration, constitue une étape fondamentale dans le traitement des images médicales. Il consiste à aligner spatialement plusieurs images acquises à partir de différentes modalités ou à différents instants, de manière à ce qu'elles correspondent à une même référence anatomique. Cette étape est particulièrement indispensable dans le cas des images IRM multimodales (T1, T2, FLAIR), où chaque séquence met en évidence des caractéristiques spécifiques des tissus cérébraux.

En pratique, le recalage permet de superposer correctement les différentes informations issues de ces séquences, facilitant ainsi l'analyse conjointe et améliorant la précision des étapes ultérieures telles que la segmentation et la classification des tumeurs [14].

Types de recalage

Selon le degré de transformation nécessaire, plusieurs types de recalage peuvent être distingués [37] :

- **Recalage rigide :** Ce type de recalage repose uniquement sur des transformations géométriques simples, telles que la translation et la rotation. Il est généralement utilisé lorsque les structures anatomiques restent globalement inchangées, par exemple entre deux acquisitions proches dans le temps [59],[60].
- **Recalage affine :** En plus des transformations rigides, le recalage affine intègre des opérations de mise à l'échelle (scaling) et de cisaillement (shearing). Il permet ainsi de corriger certaines déformations globales liées à l'acquisition ou aux différences inter- [61].

- **Recalage non rigide (ou déformable) :** Ce type de recalage est le plus avancé. Il permet de modéliser des déformations locales complexes, notamment celles dues à des variations anatomiques importantes ou à la présence de tumeurs. Il est particulièrement adapté aux applications en neuro-imagerie [61],[37].

Le problème du recalage peut être formulé comme la recherche d'une transformation T permettant d'aligner une image mobile I_2 sur une image de référence I_1 [37] :

$$I_2(x) = I_1(T(x)) \tag{1.7}$$

Où :

- I_1 représente l'image de référence,
- I_2 l'image à aligner,
- $T(x)$ la transformation appliquée aux coordonnées spatiales.

La figure ci-dessous illustre le principe du recalage d'images médicales, consistant à aligner une IRM 3D préopératoire avec une image optique 2D acquise en peropératoire. Cette étape est essentielle pour localiser précisément les structures corticales et les zones tumorales afin de guider l'intervention chirurgicale. Le recalage multimodal est particulièrement complexe en raison des différences de modalités, de résolution et de contraste entre les images.

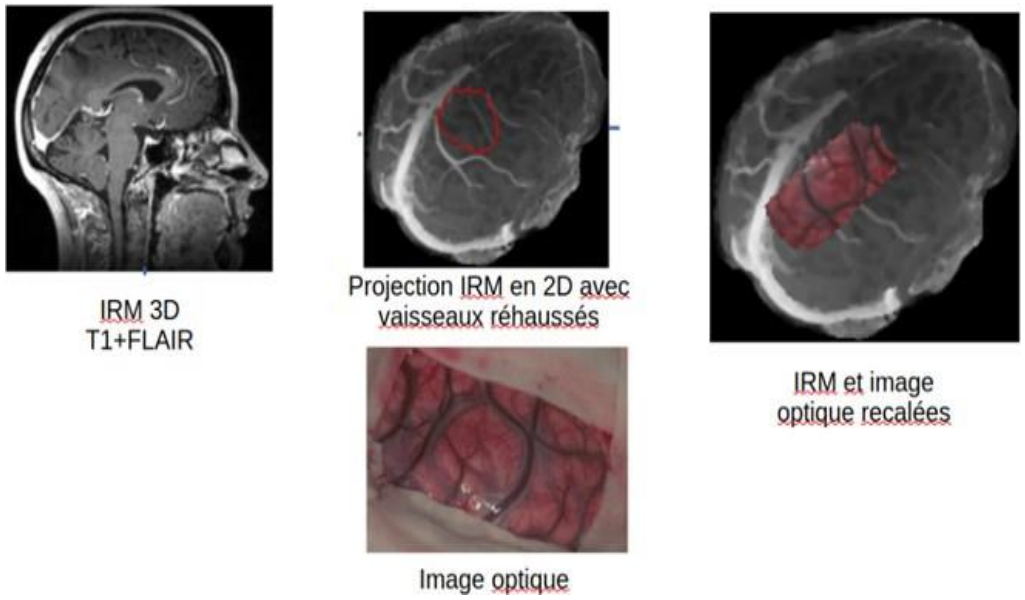


Figure 1.20 : Alignement multimodal IRM 3D / image RGB 2D par recalage médical [60]

Mesures de similarité

Afin d'évaluer la qualité du recalage, différentes mesures de similarité sont utilisées [37] :

- Corrélation croisée : adaptée aux images de même modalité.
- Information mutuelle : très utilisée en IRM multimodale, car elle permet de mesurer la dépendance statistique entre les intensités des images.

Limites du recalage

Malgré son importance, le recalage présente certaines limitations :

- Sensibilité au bruit et aux artefacts d'acquisition
- Difficulté dans les cas de déformations importantes (tumeurs volumineuses)
- Temps de calcul élevé, notamment pour les méthodes non rigides
- Dépendance aux paramètres et à l'initialisation

La figure ci-dessous présente la méthode de segmentation basée sur un atlas, combinant un recalage rigide (affine) et un recalage déformable en 3D. L'image cible est d'abord alignée avec un atlas de référence à l'aide d'une transformation rigide, puis une déformation non linéaire permet d'adapter précisément les structures anatomiques. La transformation finale est ensuite appliquée à l'atlas afin d'obtenir une segmentation cohérente avec l'image d'entrée.

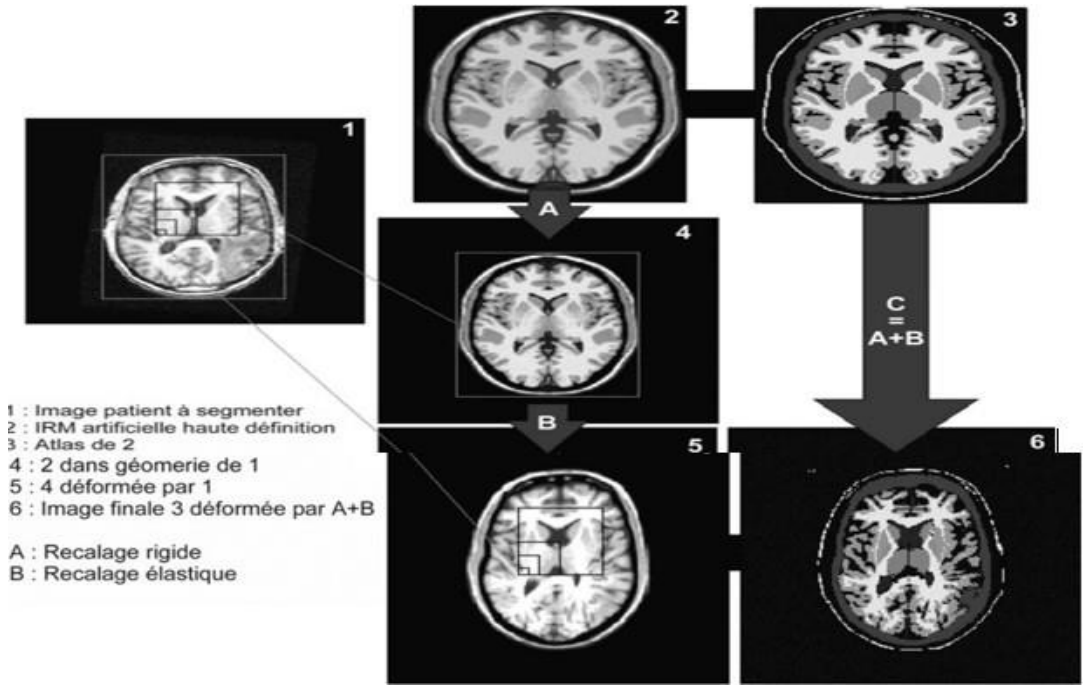


Figure 1.21 : Pipeline de segmentation par atlas basé sur le recalage affine et déformable [37]

4) Opérations morphologique :

Les opérations morphologiques [57] constituent des outils fondamentaux en traitement d'images, permettant de manipuler la structure des objets en fonction de leur forme [36]. Elles sont couramment utilisées en complément des méthodes de segmentation afin d'améliorer la qualité des résultats, notamment en supprimant les artefacts, en régularisant les contours et en renforçant la cohérence des régions d'intérêt.

Parmi les opérations de base, l'érosion et la dilatation jouent un rôle essentiel. L'érosion consiste à réduire les objets en supprimant les pixels situés à leurs frontières. Elle permet d'éliminer les petites structures indésirables et de diminuer l'effet du bruit. D'un point de vue mathématique, l'érosion d'une image A par un élément structurant B est définie par [36] :

$$A \ominus B = \{x \in Z^2 \mid B_x \subseteq A\} \tag{1.8}$$

Où B_x représente la translation de l'élément structurant B au point x.

À l'inverse, la dilatation [58] permet d'agrandir les objets en ajoutant des pixels à leur contour. Elle est particulièrement utile pour combler les petites discontinuités et connecter des régions proches. Elle est définie par [36] :

$$A \oplus B = \{x \in Z^2 \mid (B^*)_x \cap A \neq \emptyset\} \tag{1.9}$$

Où B^* désigne la symétrie de l'élément structurant.

En pratique, ces deux opérations sont souvent combinées afin de réaliser des transformations plus élaborées. L'ouverture, définie par $(A \ominus B) \oplus B$ [57], permet de supprimer les petits objets tout en préservant la forme globale des structures. À l'inverse, la fermeture, définie par $(A \oplus B) \ominus B$ [57], permet de combler les trous et de lisser les contours.

Dans le contexte de la segmentation des tumeurs cérébrales, ces opérations morphologiques sont généralement appliquées en post-traitement afin d'améliorer la qualité des masques de segmentation. Elles contribuent notamment à éliminer les régions isolées, à affiner les limites tumorales et à renforcer la cohérence spatiale des structures détectées, ce qui est essentiel pour une interprétation clinique fiable [59].

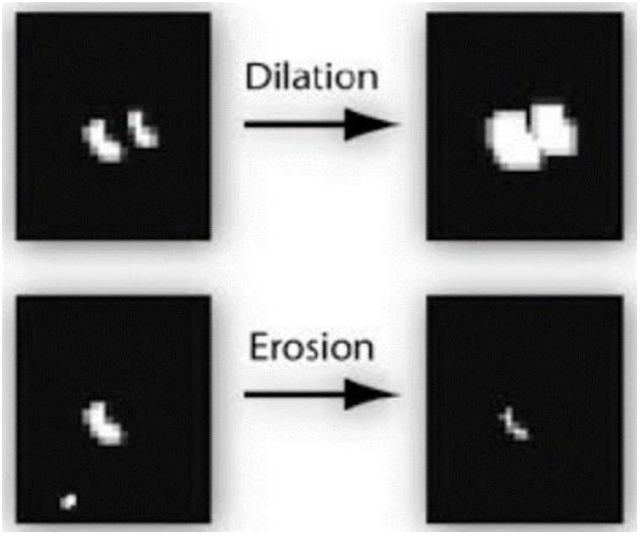


Figure 1.22 : Illustration des opérations de dilatation et d'érosion en traitement d'image [37]

1.3.1.4 *Classification des tumeurs cérébrales par IA*

La classification des tumeurs cérébrales constitue une étape complémentaire à la segmentation. Alors que la segmentation permet de délimiter précisément la région tumorale, la classification vise à identifier le type histopathologie de la tumeur à partir des caractéristiques extraites de l'image IRM [38].

Avec l'émergence du deep learning, les approches basées sur les réseaux neuronaux convolutifs (CNN) ont progressivement ont remplacé les méthodes classiques fondées sur l'extraction manuelle de caractéristiques.

1) Méthodes classiques de classification

Avant l'utilisation des réseaux profonds, la classification reposait sur[38] :

- Extraction manuelle de caractéristiques (texture, forme, intensité, histogramme)
- Méthodes statistiques et apprentissage automatique traditionnel :

i. SVM ii. K-Nearest Neighbors iii. Random Forest

Ces approches [39] ont permis d'obtenir des précisions comprises entre 75% et 85% selon les bases utilisées.

Le processus global de classification des images médicales est illustre dans la figure suivante (Figure 1.15).

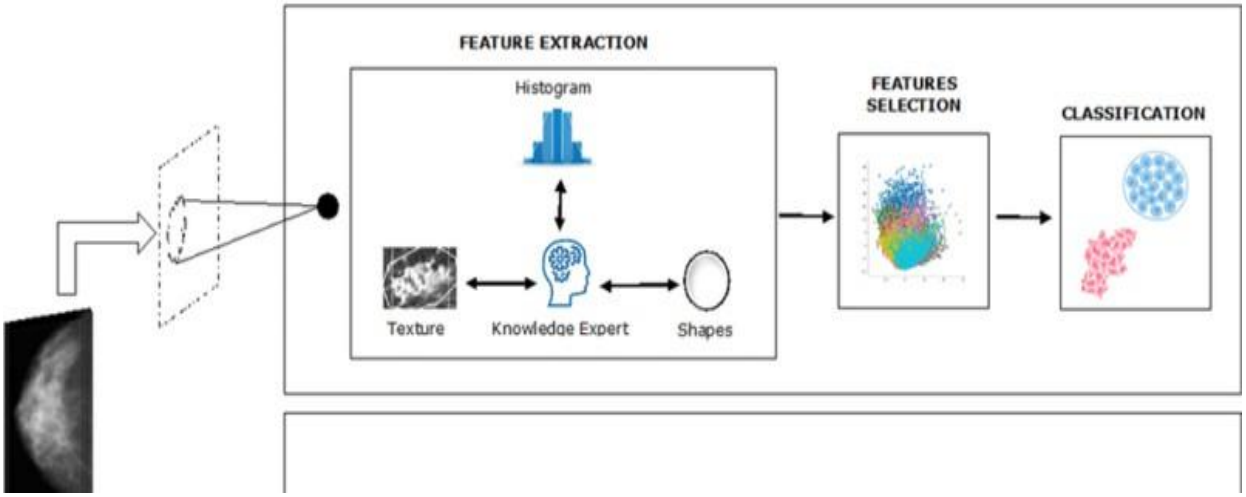


Figure 1.23 : Schéma illustrant le processus de classification en imagerie médicale [39].

2) Réseaux de neurones pour la classification

1. Convolutional Neural Networks (CNN)

Les réseaux de neurones convolutifs (CNN) constituent la base des architectures modernes de classification d'images médicales. Leur principe repose sur l'extraction hiérarchique de caractéristiques à partir des images, permettant de transformer progressivement les données brutes en représentations discriminantes [32].

L'opération fondamentale d'un CNN est la convolution [40] :

$$y(i,j) = \sum_m \sum_n x(i+m, j+n) \cdot k(m,n) \tag{1.10}$$

Où x représente l'image d'entrée, k le noyau de convolution, et y la carte caractéristique obtenue.

Cette opération permet d'extraire des motifs locaux tels que les contours, textures et variations d'intensité, qui sont essentiels pour différencier les tissus sains et pathologiques.

La non-linéarité est introduite par une fonction d'activation, généralement ReLU [41] :

$$ReLU(x) = max(0, x) \tag{1.11}$$

La fonction de perte utilisée pour la classification multi-classes est la cross-entropy [42] :

$$L_{CE} = -\sum_{k=1} y_k \log(P_k) \tag{1.12}$$

Les CNN offrent une bonne capacité d'apprentissage automatique des caractéristiques, mais restent limites par leur difficulté à capturer le contexte global et leur sensibilité au déséquilibre des classes.

2. ResNet (Residual Networks)

Les réseaux résiduels (ResNet) [76] ont été introduits afin de résoudre le problème de dégradation des performances dans les réseaux profonds principalement cause par la disparition du gradient [42] :

$$\frac{\partial I}{\partial t} \to 0 \tag{1.13}$$

ResNet propose une reformulation de la fonction à apprendre sous forme résiduelle [42] :

$$H(x) = F(x) + x \tag{1.14}$$

Où F(x) est la transformation apprise et x la connexion directe (skip connection).

Ainsi, la sortie d'un bloc résiduel est donnée par [42] :

$$Y = F(x, W) + x \tag{1.15}$$

Cette structure améliore la propagation du gradient [42] :

$$\frac{\partial y}{\partial x} = \frac{\partial F}{\partial x} + 1 \tag{1.16}$$

Les réseaux ResNet permettent ainsi d'entrainer des architectures tres profondes tout en maintenant une bonne stabilité.

- Avantages
  - Amélioration de la convergence
  - Réduction du problème de vanishing gradient
  - Robustesse sur les tâches complexes
- Limites
  - Coût computationnel élevé
  - Architecture plus complexe

3. DenseNet (Dense Convolutional Network)

DenseNet [77] repose sur une stratégie de connexions denses, où chaque couche reçoit en entrée toutes les sorties des couches précédentes [38] :

$$x_l = H_l([x_0, x_1, ..., x_{l-1}]) \tag{1.17}$$

Où [.] représente une concaténation.

Cette architecture favorise la réutilisation des caractéristiques et améliore la propagation de l'information.

- Avantages
  - Meilleur circulation du gradient
  - Réutilisation maximale des features
  - Réduction du nombre de paramètres
- Limites
  - Consommation mémoire importante
  - Complexité computationnelle

4. EfficientNet

EfficientNet [78] introduit une méthode de scaling uniforme du réseau, permettant d'optimiser simultanément la profondeur, la largeur et la résolution de l'image [39].

Le principe repose sur une mise à l'échelle composée[39] :

$$depth = \alpha^{\varphi}, \quad width = \beta^{\varphi}, \quad resolution = \gamma^{\varphi} \tag{1.18}$$

Avec la contrainte [39] :

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \tag{1.19}$$

Où φ est un coefficient de scaling global.

Cette approche permet d'obtenir un meilleur compromis entre précision et complexité.

- Avantages
  - Excellente performance avec moins de paramètres
  - Optimisation du rapport précision/coût
  - Adapté aux images médicales haute résolution
- Limites

- Architecture plus difficile à comprendre
- Nécessite un réglage précis

5. MobileNet

MobileNet [79] est une architecture légère conçue pour réduire le coût computationnel, notamment grave à l'utilisation de convolutions séparables en profondeur [38].

Une convolution standard est décomposée en :

Conv standard = Depthwise Conv + Pointwise Conv

Cela réduit le coût de calcul :

Coût réduit ≈ (formule)

Où N est le nombre de filtres.

- Avantages
  - Faible consommation mémoire
  - Rapidité d'exécution
  - Adapté aux systèmes embarqués
- Limites
  - Précision légèrement inférieure aux modèles lourds
  - Moins performant pour structures complexes

6. Pipeline de prédiction

Afin de mieux analyser les différentes architectures de classification utilisées dans ce travail, une synthèse comparative est présentée dans le Tableau 1.1 . Ce tableau met en évidence leurs principes de fonctionnement, leurs objectifs ainsi que leurs avantages et limites dans le contexte de la classification des tumeurs cérébrales.

Architecture	Principe fondamental	Objectif principal	Avantages	Limites
CNN	Extraction locale par convolution	Apprentissage des caractéristiques de base	Simplicité, efficacité, bonne performance en imagerie	Manque de contexte globale, sensible au déséquilibre
ResNet	Connexions résiduelles $y = F(x) + x$	Amélioration de la profondeur et stabilité	Réduction du vanishing gradient, convergence rapide	Complexité plus élevée
DenseNet	Connexions denses entre couches	Réutilisation maximale des caractéristiques	Meilleure propagation du gradient, efficacité des features	Consommation mémoire importante
EfficientNet	Scaling composé (profondeur, largeur, résolution)	Optimisation précision/complexité	Haute précision avec peu de paramètres	Implémentation plus complexe
MobileNet	Convolutions séparables (depthwise+pointwise)	Réduction du coût computationnel	Rapide, léger, adapté aux systèmes embarqués	Précision légèrement inférieure

TAB. 1.1 : Comparaison des architectures de réseaux de neurones pour la classification d'images médicales

Comme illustre dans le Tableau 1.1, chaque architecture présente des caractéristiques spécifiques qui répondent à des besoins particuliers.

Les CNN assurent l'extraction des caractéristiques de base, tandis que ResNet et DenseNet améliorent respectivement la stabilité de l'apprentissage et la réutilisation des informations. Par ailleurs, EfficientNet optimise le compromis entre précision et complexité, alors que MobileNet vise à réduire le cout computationnel. Ainsi, le choix des ces architectures permet d'explorer différents compromis adaptés aux contraintes des données médicales et aux ressources disponibles.

Le pipeline de prédiction constitue la phase opérationnelle du système de classification des tumeurs cérébrales. Il regroupe l'ensemble des traitements appliques de manière séquentielle afin de transformer une image IRM brute en une décision clinique exploitable.

Contrairement à la partie théorique décrite précédemment, cette étape met en évidence l'organisation réelle du système implémenté dans ce travail.

Dans ce contexte, le pipeline repose sur une approche en deux étapes : segmentation automatique suivie d'une classification supervisée.

Organisation du pipeline

Le processus global peut être représenté par la chaine suivante [38] :

$$x \rightarrow M \rightarrow x_{ROI} \rightarrow f_0 \rightarrow \hat{y}$$

Où :

- x : image IRM d'entrée
- M : masque de segmentation généré automatiquement
- x_{ROI} : region tumorale extraite
- f_0 : modèle de classification
- $\hat{y}$: prédiction finale

Rôle de segmentation

La segmentation constitue une étape préalable essentielle. Elle permet de délimiter automatiquement la zone tumorale au sein de l'image IRM.

Cette étape présente un rôle fondamental [18] :

- Suppression des régions non informatives
- Isolation de la tumeur
- Réduction de la variabilité anatomique entre patients

Ainsi, l'image traitée devient :

$$x_{ROI} = x \odot M \tag{1.20}$$

Ce qui permet de concentrer l'analyse uniquement sur la zone pathologique.

Étape de classification

Une fois la région d'intérêt extraite, celle-ci est introduite dans un modèle de classification supervisée afin de déterminer la classe de la tumeur.

Le modèle produit une distribution probabiliste [39] :

$$P(y = k|x) = softmax(z_k) \tag{1.21}$$

Et la classe finale est obtenue par [38] :

$$\hat{y} = \arg\max_k P(y = k|x) \tag{1.22}$$

Intérêt du pipeline combiné

L'approche segmentation et classification présente plusieurs avantages importants dans le contexte médical :

- Amélioration de la qualité des caractéristiques extraites
- Réduction du bruit provenant des tissus non tumoraux
- Meilleure robustesse du modèle face aux variations des images IRM
- Augmentation de la précision globale du système

Interprétation du pipeline de prédiction

Le pipeline de prédiction peut être interprété comme une transformation progressive de l'information, où l'image IRM brute est d'abord traitée afin de localiser la tumeur, puis analysée pour en extraire des caractéristiques pertinentes, avant de conduire à une décision clinique basée sur un modèle de classification.

Métriques d'évaluation

L'évaluation des performances du modèle de classification est une étape essentielle permettant de mesurer la qualité des prédictions générées[41].

Dans le domaine médical, plusieurs métriques complémentaires sont utilisées afin d'obtenir une analyse fiable et robuste.

Matrice de confusion

La matrice de confusion constitue une représentation fondamentale des performances du modèle. Elle permet de visualiser les différentes catégories de prédictions [41] :

- TP : vrais positifs
- TN : vrais négatifs
- FP : faux positifs
- FN : faux négatifs

Elle sert de base au calcul de toutes les métriques d'évaluation.

Accuracy (exactitude globale)

Elle mesure la proportion de bonnes prédictions sur l'ensemble des données [41] :

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1.23}$$

Elle donne une vision globale mais peut être biaisée en cas de déséquilibre des classes [62].

Précision

Elle mesure la fiabilité des prédictions positives [63] :

$$Precision = \frac{TP}{TP + FP} \tag{1.24}$$

Elle est importante pour limiter les faux diagnostics positifs [41].

Recall (sensibilité)

Elle mesure la capacité du modèle à détecter les cas positifs :

$$Recall = \frac{TP}{TP + FN} \tag{1.25}$$

En imagerie médicale, elle est critique pour éviter les cas non détectés [41].

AUC (Area Under Curve)

Elle mesure la capacité de séparation entre les classes [41] :

$$AUC = \int_0^1 TPR(FPR)\, dFPR \tag{1.26}$$

Avec :

$$TPR = \frac{TP}{TP + FN}, \quad FPR = \frac{FP}{FP + TN} \tag{1.27}$$

Cohen's kappa

Le coefficient de kappa permet de mesurer le degré d'accord entre les prédictions du modèle et la vérité terrain, tout en tenant compte de l'accord obtenu par hasard [41].

$$k = \frac{P_0 - P_e}{1 - P_e} \tag{1.28}$$

Où :

- P_0 Correspond à l'accord observé (accuracy)
- P_e Correspond à l'accord attendu par hasard

Interprétation [41] :

- k = 1 : accord parfait
- k = 0 : accord aléatoire
- k < 0 : accord faible

Cette métrique est particulièrement pertinente en imagerie médicale, où les classes sont souvent déséquilibrées [65].

L'utilisation combinée de plusieurs métriques est indispensable afin d'obtenir une évaluation complète du modèle. Chaque métrique apporte une information complémentaire permettant d'analyser la performance sous différents angles (global, sensibilité, précision, séparation des classes).

Courbe ROC (classification)

La courbe ROC (Receiver Operating Characteristic) est utilisée en classification pour évaluer la capacité du modèle à distinguer les différentes classes. Elle représente le taux de vrais positifs (TPR) en fonction du taux de faux positifs (FPR) pour différents seuils de décision. L'aire sous la courbe (AUC) quantifie la performance globale du modèle : plus l'AUC est proche de 1, meilleur est le pouvoir de discrimination du modèle[40].

1.3.2 *Réalité Augmentée en Neuro-imagerie*

La réalité augmentée (RA) en neuro-imagerie représente une avancée technologique permettant de superposer des informations virtuelles sur des images réelles, améliorant ainsi l'interprétation et l'analyse des données IRM cérébrales. En neurochirurgie, la RA facilite la visualisation des tumeurs [37], la planification des interventions et la communication avec le patient.

1.3.2.1 *Historique de la réalité augmentée*

La réalité augmentée (RA) trouve ses origines dans les travaux pionniers de Ivan Sutherland dans les années 1960, avec le premier système de visualisation appelé "Sword of Damocles" (Figure 1.24) .

Le terme réalité augmentée a ensuite été formalisé par Ronald Azuma en 1997, qui définit la RA comme un système combinant le réel et le virtuel, interactif en temps réel et aligné en trois dimensions avec l'environnement réel. Depuis, les avancées technologiques ont permis son application en médecine, notamment en neurochirurgie pour la visualisation et l'assistance opératoire [43].

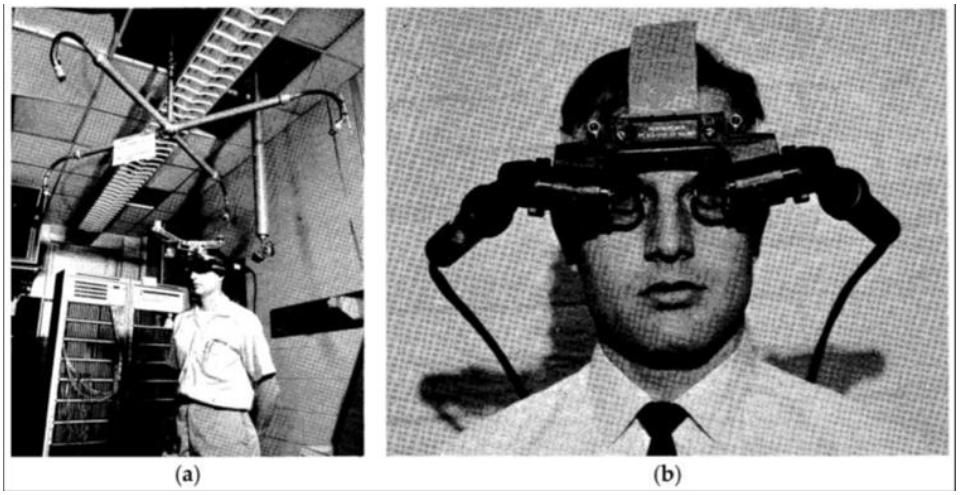


Figure 1.24 : Ivan Sutherland, considéré comme l'un des précurseurs de la réalité augmentée [43].

1.3.2.2 *RA pour la visualisation 3D des tumeurs*

La RA permet de construire un modèle volumique de la tumeur [44] à partir des images IRM et de superposer à l'anatomie réelle du patient.

Cette visualisation interactive offre un repérage spatial précis des structures cérébrales et de la tumeur.

Fonctionnalités :

- Reconstruction volumique de la tumeur et des tissus environnants.
- Superposition du modèle 3D sur les images IRM ou le patient réel.
- Interaction dynamique : rotation, zoom, découpage virtuel.

Applications :

- Formation des étudiants et résidents en neurochirurgie.
- Communication médecin-patient pour expliquer l'emplacement et l'étendue de la tumeur.

Limites :

- Difficultés de registration entre le modèle 3D et l'anatomie réelle.
- Risque d'erreurs de superposition, pouvant induire une mauvaise interprétation.

La figure suivante illustre un exemple de visualisation tridimensionnelle du cerveau permettant de représenter la tumeur et ses relations avec les structures anatomiques environnantes (Figure 1.25) [44][45].

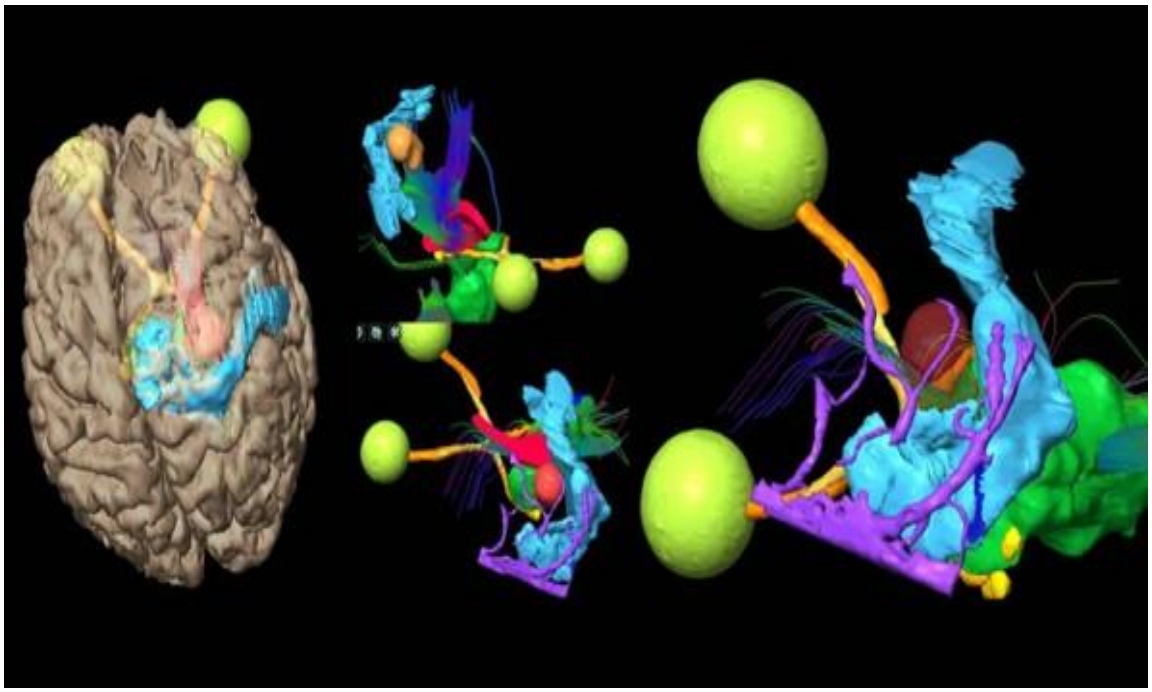


Figure 1.25 : Modélisation tridimensionnelle du cerveau illustrant la localisation de la tumeur et ses relations avec les structures anatomiques environnantes dans un environnement de RA [44].

1.3.2.3 *RA pour la planification préopératoire*

En neurochirurgie, la RA sert également à planifier l'intervention [46]. Les modèles 3D peuvent être utilisés pour déterminer la trajectoire opératoire, délimiter les marges tumorales et simuler l'intervention.

Fonctionnalités :

- Navigation neurochirurgicale guidée par RA.
- Délimitation précise des marges tumorales.
- Simulation d'intervention pour tester différents angles ou stratégies.

Limites :

- Latence dans l'affichage en temps réel
- Cout élevé du matériel spécialisé (casques RA, capteurs)
- Précision spatiale limitée si la calibration ou la registration est imparfaite

La figure suivante illustre les différentes étapes de la planification préparatoire en neurochirurgie assistée par réalité augmentée, depuis la reconstruction 3D jusqu'à sa projection en salle opératoire (Figure 1.18).

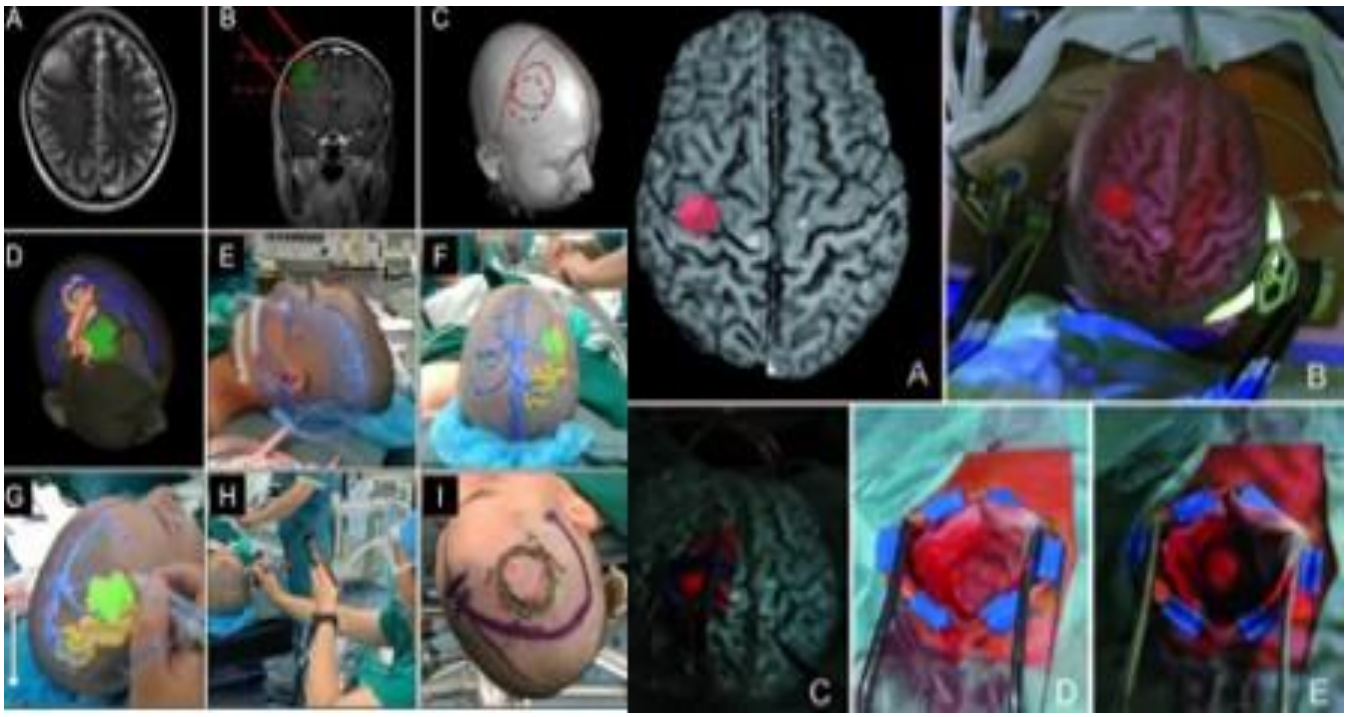


Figure 1.26 : : Planification préopératoire d'une tumeur cérébrale à l'aide de RA [46] :
A : Modèle 3D du cerveau reconstruit à partir d'images cérébrales d'IRM avec métastase cérébrale en rouge
B : Après enregistrement, l'image créée dans le cadre A est projetée sur la tête du patient dans la planification de l'incision cutanée et de la craniotomie
C : Projection image en peropératoire après la 1ère incision cutanée et avant la craniotomie avec localisation de la tumeur sur le crâne pour planifier une craniotomie sur mesure
D : Surface du cerveau après ouverture de la dure-mère. La tumeur n'est pas visible à l'œil directement
E : Projection de la tumeur à la surface du cerveau [Neurosurgery Blog].

1.3.2.4 *Limites actuelles de la RA médicale*

Malgré ses avantages, la RA en neuro-imagerie présente encore des défis importants [47] :

- Intégration limitée de l'intelligence artificielle pour l'identification automatique des structures cliniques
- Peu de systèmes temps réel fiables pour une utilisation en salle d'opération
- Manques de validation cliniques à grande échelle, limitant son adoption universelle

Le Tableau 1.2 met en évidence que, malgré le potentiel considérable de la réalité augmentée en neurochirurgie, plusieurs défis techniques et cliniques doivent encore être surmontes afin d'assurer une intégration fiable et sécurisée en pratique médicale [47].

Limitation	*Impact clinique*	*Solutions potentielles*
Erreurs de registration (mauvais alignement virtuel-réel)	Mauvaise localisation de la tumeur ou des structures critiques (vaisseaux, aires fonctionnelles).	Algorithmes de recalage plus robustes, recalibration en temps réel.
Latence du système.	Retard dans l'affichage pouvant perturber la précision gestuelle du chirurgien.	Optimisation des algorithmes, utilisation de matériel à haute performance.
Précision spatiale limitée.	Risque de déviation lors de la navigation chirurgicale.	Intégration de capteurs de suivi plus précis (tracking optique ou électromagnétique).
Coût élevé des dispositifs RA.	Accessibilité limitée dans les hôpitaux à ressources réduites.	Développement de solutions logicielles compatibles avec du matériel standard.
Manque de validation clinique à grande échelle.	Adoption limitée dans la pratique clinique courante.	Réalisation d'essais cliniques multicentriques et études comparatives.

TAB. 1.2 : Limitations, impacts cliniques et solutions potentielles de la réalité augmentée en neurochirurgie

1.3.3 *Approches hybrides Intelligence Artificielle-Réalité Augmentée*

L'évolution récente des technologies médicales a conduit à l'émergence de systèmes hybrides combinant l'intelligence artificielle(IA) et la réalité augmentée (RA) [41]. Dans le contexte de la neuro-imagerie, cette convergence vise à dépasser les limites des approches isolées en associant la puissance analytique des algorithmes d'apprentissage profond a la capacité immersive et interactive de la RA[48].

Alors que l'IA permet l'analyse automatique des images IRM cérébrales, notamment la segmentation et la classification des tumeurs. La réalité augmentée transforme ces résultats numériques en représentations spatiales exploitables en environnement clinique. L'approche hybride constitue ainsi une nouvelle génération de systèmes d'aides à la décision, intégrant détection intelligente et visualisation avancée dans un cadre unifié.

1.3.3.1 *Fusion IA-RA pour une visualisation intelligente*

La fusion IA-RA repose généralement sur une architecture en plusieurs niveaux intégrés [49] :

1. Analyse automatique par IA

Les images IRM sont traites par des modèles de deep learning (U-Net, ResNet, Transformers, ect.) afin d'identifier, segmenter et éventuellement classifier la tumeur cérébrale.

Cette étape permet d'extraire avec précision la région pathologique et ses caractéristiques morphologiques[49].

2. Reconstruction tridimensionnelle

Les résultats de segmentation sont convertis en modèles volumiques 3D représentant la tumeur et les structures anatomiques environnantes. Cette reconstruction améliore la perception des rapports spatiaux complexes[49].

3. Projection en réalité augmentée

Le modèle 3D généré est ensuite superposé à l'environnement réel (patient, mannequin, anatomique ou imagerie projetée) à l'aide d'un dispositif de RA. L'utilisateur peut interagir avec la scène virtuelle : rotation, zoom, filtrage des structures[49].

Cette intégration permet une visualisation intelligente contextuelle, ou les résultats algorithmiques ne restent pas abstraits mais deviennent directement interprétables dans l'espace anatomique réel.

Plusieurs systèmes hybrides combinant intelligence artificielle et réalité augmentée ont été proposés dans la littérature. Par exemple, des plateformes basées sur Microsoft HoloLens ont été utilisées pour la visualisation de tumeurs cérébrales segmentées automatiquement.

D'autres travaux récents ont explore l'intégration de modèles d'apprentissage profond (deep learning) avec des systèmes de navigation chirurgicale en temps réel, permettant d'améliorer la précision du guidage opératoire [49].

Apports cliniques des approches hybrides Les travaux récents consacrés à l'intégration de l'IA et RA en neurochirurgie mettent en évidence plusieurs bénéfices potentiels [48]. Ces apports peuvent être résumés comme suit :

- **Amélioration de la compréhension spatiale des rapports tumeur-structure critiques**

La visualisation tridimensionnelle issue de la segmentation automatique permet au clinicien d'analyser avec davantage de précision les relations anatomiques entre la tumeur, les vaisseaux sanguins et les zones fonctionnelles du cerveau. Cette perception volumique réduit les ambiguïtés liées à l'interprétation des coupes bidimensionnelles classiques.

- **Facilitation de la planification préopératoire**

La projection du modèle tumoral dans un environnement augmente autorise la simulation de trajectoires chirurgicales et l'anticipation des difficultés opératoires. Cela contribue à une préparation plus rigoureuse et à une meilleure gestion des risques.

- **Réduction de l'incertitude dans l'interprétation des images complexes**

L'IA automatise l'identification des zones pathologiques, tandis que la RA contextualise ces résultats dans l'espace réel. Cette complémentarité améliore la lisibilité des informations radiologiques.

- **Renforcement de la communication médecin-patient**

La représentation immersive de la tumeur facilite l'explication du diagnostic et du plan thérapeutique, rendant l'information plus accessibles au patient.

Cependant, il convient de souligner que ces bénéfices restent principalement démontrés dans des environnements expérimentaux ou semi-cliniques, et nécessitent encore des validations à grande échelle.

1.3.4 *Systèmes hybrides existants*

À l'heure actuelle, les systèmes combinant intelligence artificielle et réalité augmentée en neuro-chirurgie demeurent en phase de développement avancé mais non totalement mature. On distingue principalement [49] :

- **Des prototypes académiques,** développés dans un cadre de recherche pour démontrer la faisabilité technique de l'approche hybride.
- **Des plateformes expérimentales,** intégrant des modules de segmentation automatique couples a des environnements de visualisation immersive.
- **Des démonstrateurs technologiques,** présentés lors de projets pilotes ou d'études préliminaires.

Malgré leur potentiel, leur intégration complète dans la pratique hospitalière reste limitée en raison de plusieurs facteurs :

- **Validation clinique restreinte,** souvent réalisée sur de petits échantillons de patients.
- **Absence de standardisation,** rendant difficile l'harmonisation des protocoles d'utilisation.
- **Complexité d'intégration dans le flux opératoire réel,** ou la stabilité et la rapidité du système sont essentielles.
- **Contraintes matérielles et logicielles élevées,** notamment en termes de performances de calcul et de compatibilité des équipements.

Ainsi, bien que prometteurs, ces systèmes ne sont pas encore largement adoptés comme outils standards d'aide au diagnostic ou à la navigation chirurgicale.

1.3.4.1 *Difficultés et verrous scientifiques*

La mise en œuvre d'un système hybride IA-RA robuste soulève plusieurs défis multidimensionnels.

Verrous techniques :

- **Registration précise IA-RA**

L'alignement exact entre la segmentation générée par l'algorithme et l'anatomie réelle du patient constitue un enjeu central. Toute erreur de recalage peut engendrer un décalage spatial susceptible d'affecter la précision clinique.

- **Traitement en temps réel**

La combinaison de modèles de deep learning et de moteurs de rendu 3D impose des exigences élevées en termes de puissance de calcul afin de limiter la latence et garantir une interaction fluide.

- **Optimisation des ressources matérielles(GPU/CPU)**

L'exécution simultanée des modèles d'intelligence artificielle et de visualisation nécessite une architecture informatique performante et correctement optimisée [50].

Verrous cliniques :

- **Acceptation par les professionnels de santé**

L'adoption clinique de ces systèmes repose en grande partie sur le niveau de confiance que les professionnels de santé accordent aux algorithmes d'intelligence artificielle, notamment en termes de fiabilité, d'exploitabilité et de robustesse.

- **Responsabilité médico-légale**

L'introduction d'outils intelligents soulève des interrogations quant à la responsabilité en cas d'erreur diagnostique ou de complication chirurgicale [51].

- **Fiabilité décisionnelle**

Les modèles doivent démontrer leur capacité à généraliser face à la variabilité inter-patients et inter-centres.

Enjeux éthiques :

- **Transparence et exploitabilité des modèles d'IA,** afin de permettre une compréhension des décisions algorithmiques.

- **Biais potentiels lies aux bases de données d'apprentissage,** pouvant affecter l'équité des résultats.

- **Protection et confidentialité des données médicales,** conformément aux règlementations en vigueur [50]

Ces différents aspects conditionnent l'intégration responsable et durable des systèmes hybrides dans le domaine médical.

1.4 *Conclusion*

Ce chapitre a présenté le contexte et l'état de l'art des approches d'intelligence artificielle et de réalité augmentée en neuro-imagerie. Les méthodes classiques de segmentation, limitées par le bruit et la variabilité inter-patients, ont laissé place aux réseaux profonds et architectures avancées (U-Net, CNN...), améliorant la précision et la prise en compte du contexte spatial.

La réalité augmentée apporte une visualisation 3D interactive et facilite la planification préopératoire, mais reste confrontée à des limites techniques et cliniques. Les systèmes hybrides IA-RA montrent un fort potentiel pour l'aide à la décision et la communication, tout en nécessitant une validation, une optimisation et une attention aux aspects éthiques et légaux.

Cette synthèse prépare le terrain pour l'étude des méthodes de classification des tumeurs cérébrales,abordée dans le chapitre suivant.

CHAPITRE II

Conception et développement du système

2.1 *Introduction*

Le présent chapitre décrit la conception et le développement du système intelligent proposé dans ce travail, destiné à assister le diagnostic des tumeurs cérébrales à travers l'intégration de l'intelligence artificielle et de la réalité augmentée. L'objectif principal est de proposer une chaîne de traitement complète allant de la segmentation automatique des structures tumorales jusqu'à leur classification et leur visualisation interactive en environnement immersif.

Contrairement au chapitre précédent, qui a présenté le contexte médical ainsi que l'état de l'art des approches existantes en imagerie médicale, en intelligence artificielle et en réalité augmentée, ce chapitre se concentre sur la mise en œuvre pratique du système développé. Il met en évidence les différents choix techniques adoptés ainsi que l'architecture globale du pipeline proposé.

Ainsi, ce chapitre détaille successivement les différentes étapes de conception, notamment les bases de données utilisées, les méthodes de prétraitement appliquées, les modèles d'intelligence artificielle employés pour la segmentation et la classification, ainsi que l'intégration des résultats dans un environnement de réalité augmentée.

L'objectif est de montrer comment les concepts théoriques présentés précédemment sont transformés en une solution fonctionnelle, structurée et adaptée à une utilisation dans un contexte d'analyse et d'aide à la décision médicale.

2.2 *Architecture globale du système*

L'objectif principal du système développé dans ce travail est de proposer une plateforme intégrée permettant l'analyse des tumeurs cérébrales à partir d'images d'imagerie par résonance magnétique (IRM), en combinant les capacités de l'intelligence artificielle et de la réalité augmentée. Cette approche vise non seulement à améliorer la détection et la visualisation des régions tumorales, mais également à fournir des outils d'aide au diagnostic et à l'interprétation des résultats.

L'architecture proposée repose sur une chaîne de traitement complète organisée en plusieurs modules complémentaires. Chaque module assure une fonction spécifique dans le traitement des données médicales et contribue à la transformation progressive des images brutes en informations cliniquement exploitables.

Le processus débute par l'acquisition des données médicales provenant des bases de données utilisées dans cette étude. Ces données peuvent être constituées d'images IRM bidimensionnelles ou de volumes tridimensionnels selon l'objectif visé. Les images sont généralement fournies sous des formats standards de l'imagerie médicale tels que DICOM (Digital Imaging and Communications in Medicine) ou NIfTI (Neuroimaging Informatics Technology Initiative), largement utilisés pour le stockage et l'échange des données neuro-radiologiques.

Après leur acquisition, les données sont soumises à une étape de prétraitement destinée à améliorer leur qualité et à les rendre compatibles avec les modèles d'apprentissage profond. Cette phase comprend notamment la normalisation des intensités, le redimensionnement des images, l'harmonisation des dimensions ainsi que différentes opérations de préparation des données. Ces traitements permettent de réduire les variations liées aux conditions d'acquisition et d'améliorer la robustesse des modèles utilisés par la suite.

Les images prétraitées sont ensuite transmises au module de segmentation automatique. Dans ce travail, cette étape repose sur des architectures de deep learning capables d'identifier automatiquement les régions tumorales présentes dans les images cérébrales. La segmentation permet de localiser avec précision les zones pathologiques et de générer des masques représentant les différentes structures d'intérêt.

Afin d'améliorer la qualité des résultats obtenus, une étape de post-traitement est appliquée aux masques prédits. Cette phase vise à éliminer les régions parasites, réduire les faux positifs et améliorer la cohérence anatomique des contours segmentés. Les opérations réalisées permettent d'obtenir des segmentations plus propres et mieux adaptées aux étapes ultérieures du pipeline.

Les résultats de segmentation sont ensuite exploités dans deux approches complémentaires de visualisation en réalité augmentée.

La première approche correspond à la réalité augmentée 2D. Elle consiste à superposer les masques segmentés directement sur les images IRM originales afin de mettre en évidence les régions tumorales détectées. Cette représentation facilite l'interprétation visuelle des résultats tout en conservant le contexte anatomique fourni par les images médicales initiales.

La seconde approche repose sur la réalité augmentée 3D. Les masques segmentés sont convertis en modèles tridimensionnels représentant les différentes structures anatomiques cérébrales. Ces modèles sont ensuite intégrés dans un environnement immersif développé sous Unity, permettant leur visualisation, leur manipulation et leur exploration interactive à l'aide d'un smartphone compatible avec les technologies de réalité augmentée.

En parallèle de la visualisation, les informations issues de la segmentation sont également utilisées dans une étape de classification automatique des tumeurs cérébrales. Cette phase a pour objectif d'identifier la catégorie tumorale à partir des caractéristiques présentes dans les images médicales. Plusieurs architectures de deep learning sont étudiées afin d'évaluer leur capacité à distinguer les différents types de tumeurs.

Au-delà des aspects techniques, cette architecture répond à plusieurs objectifs fonctionnels. Elle contribue d'une part à l'amélioration du diagnostic assisté par ordinateur grâce à l'automatisation de l'analyse des images. Elle facilite également la planification et l'interprétation des résultats grâce aux outils de visualisation avancée. Enfin, l'intégration de la réalité augmentée offre une assistance visuelle supplémentaire permettant une meilleure compréhension spatiale des structures anatomiques et pathologiques.

La Figure 2.1 présente le schéma bloc général de l'architecture proposée, illustrant l'enchaînement des différentes étapes du pipeline depuis l'acquisition des données jusqu'à la classification finale des tumeurs cérébrales.

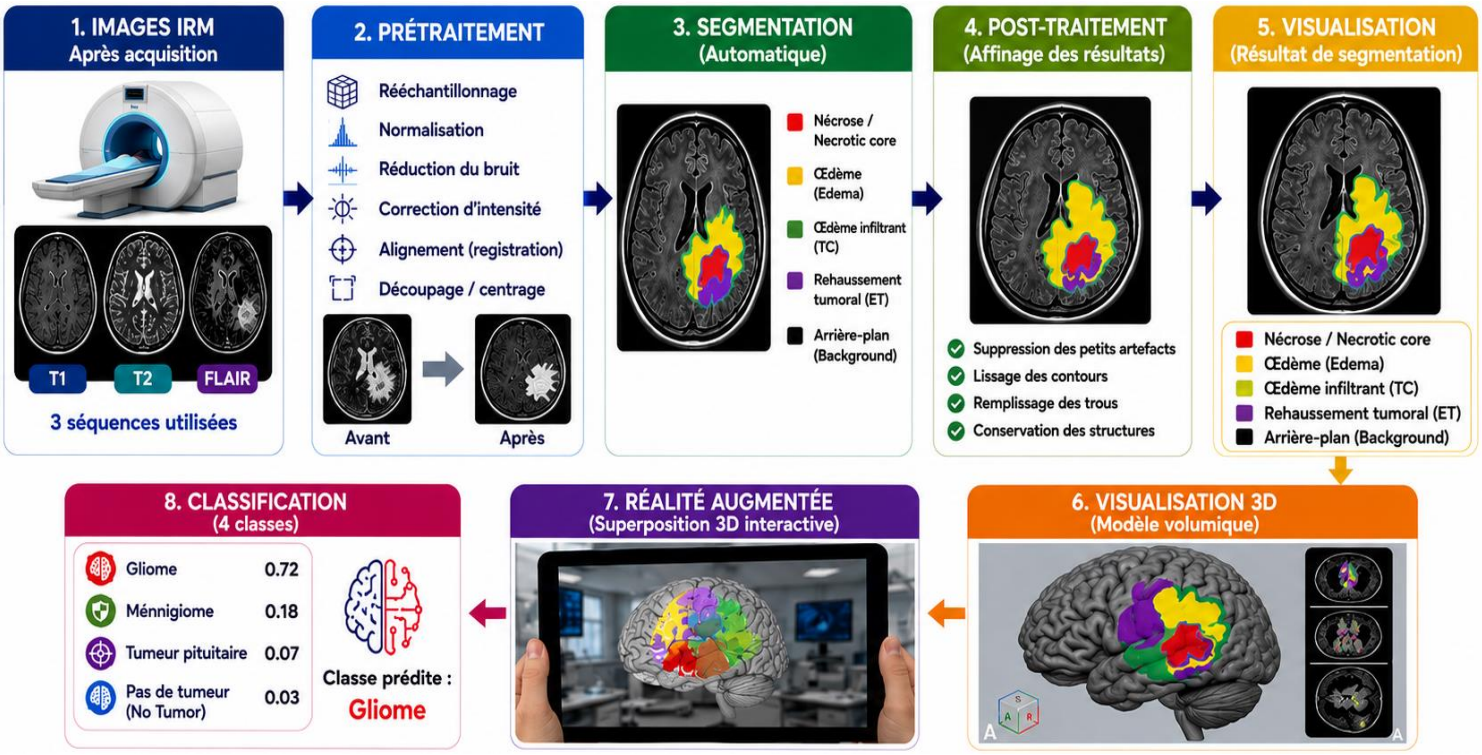


Figure 2.1 : Architecture globale du système proposé.

2.3 *Base de données*

1) BRISC 2025

Dans le cadre de la première approche de visualisation en réalité augmentée 2D, ce travail repose principalement sur la base de données BRISC 2025 ***(Brain Tumor Image Segmentation and Classification)***, un ensemble de données médicales dédié aux tâches de segmentation et de classification des tumeurs cérébrales à partir d'images IRM. Cette base a été sélectionnée en raison de son adéquation avec les approches de visualisation bidimensionnelle et de son adaptation aux méthodes d'apprentissage profond.

BRISC 2025 constitue une ressource particulièrement adaptée au développement du pipeline de réalité augmentée 2D, grâce à la disponibilité d'annotations précises et à la diversité des cas cliniques qu'elle contient. Elle permet ainsi l'entraînement et l'évaluation de modèles capables de segmenter les régions tumorales et de les superposer aux images IRM originales dans un environnement de visualisation enrichi.

2.3.1 *Description des données d'imagerie(BRISC 2025)*

La base de données BRISC 2025 [62] est constituée d'images IRM cérébrales pondérées en T1, organisées autour de deux tâches principales : la segmentation et la classification des tumeurs cérébrales. Les données sont réparties en ensembles d'entraînement (train) et de test (test), afin de permettre l'apprentissage et l'évaluation des modèles d'intelligence artificielle dans des conditions cohérentes.

Chaque image correspond à une coupe bidimensionnelle du cerveau obtenue selon différents plans anatomiques, notamment axial, coronal et sagittal. Dans le cadre de la tâche de classification, les données sont réparties selon quatre catégories principales :

- Gliome ;
- Méningiome ;
- Tumeur hypophysaire ;
- Absence de tumeur (No Tumor).

Cette organisation permet le développement de modèles capables d'identifier automatiquement le type tumoral à partir des caractéristiques présentes dans les images IRM.

Dans le cadre de la tâche de segmentation, chaque image est associée à un masque binaire représentant les régions tumorales au niveau pixel. Cette annotation permet la localisation précise des zones pathologiques et facilite l'apprentissage des modèles de segmentation automatique. Par ailleurs, chaque échantillon suit une organisation structurée permettant d'identifier différentes informations relatives aux données, notamment :

- Le type de tumeur ;
- Le plan anatomique d'acquisition ;
- Le type de séquence IRM utilisée. Cette structuration des données facilite leur intégration dans les pipelines de traitement d'image et les modèles d'apprentissage profond développés dans ce travail. Pour présenter l'organisation des données, le tableau 2.1 résume la répartition des images entre les ensembles d'apprentissage et de test.

Tache	***Train Images***	***Test Images***	***Total***
Classification	Glioma :1147 Meningioma :1328 No-Tumor :1066 Pituitary :1456	Glioma :254 Meningioma :305 No-Tumor :139 Pituitary :299	5995
Segmentation	4998	997	5995

TAB. 2.1 : Répartition des données pour les tâches de classification et de segmentation [62]

Afin de pouvoir exploiter efficacement cette base de données dans un cadre d'apprentissage profond, les images brutes doivent subir une étape de préparation et de standardisation. Cette étape, appelée prétraitement des données, permet d'adapter les données d'entrée aux exigences des modèles utilisés et constitue une phase essentielle du pipeline global de traitement.

2.3.1.1 *Augmentation de la base de données*

Afin d'améliorer la capacité de généralisation du modèle et de réduire le risque de surapprentissage, une stratégie d'augmentation de données a été appliquée directement sur les images d'entraînement. Cette étape est essentielle en imagerie médicale, où la quantité de données annotées est généralement limitée. Dans ce travail, les transformations suivantes ont été utilisées :

- Rotation aléatoire des images (jusqu'à $\pm$ 15) ;
- Retournement horizontal et vertical ;
- Variation de la luminosité et du contraste ;
- Ajout de bruit gaussien.
- Redimensionnement uniforme des images.

Ces opérations permettent de simuler différentes conditions d'acquisition et d'augmenter la variabilité des données sans modifier leur contenu anatomique. L'objectif principal de cette augmentation est d'améliorer la robustesse du modèle face aux variations inter-patients et aux différences d'acquisition IRM. Ainsi, le modèle est capable de mieux généraliser ses performances sur des données non vues lors de l'entraînement.

2.3.1.2 *Format des images et structures des données*

Les images de la base de données BRISC 2025 sont fournies principalement au format JPEG/PNG, ce qui les distingue des datasets volumétriques classiques en format NIfTI. Chaque image représente une coupe bidimensionnelle indépendante du cerveau, ce qui permet leur traitement direct par des réseaux de neurones convolutifs 2D. La structure des données est organisée comme suit :

- Images/ : contient les images IRM originales
- Masks/ : contient les masques de segmentation correspondants. Chaque masque est aligné spatialement avec son image associée, garantissant une correspondance pixel à pixel.

Les images sont ensuite redimensionnées à une taille uniforme afin de les adapter aux contraintes des réseaux de neurones profonds.

2.3.1.3 *Organisation des données (train/test)*

Dans ce travail, la base de données BRISC 2025 est fournie avec une séparation préétablie en deux ensembles distincts : un ensemble d'entraînement (train) et un ensemble de test (test). Cette organisation est directement exploitée dans le pipeline de traitement sans modification, ce qui permet de garantir une évaluation objective des performances du modèle sur des données non vues lors de l'apprentissage.

L'ensemble d'entraînement est utilisé pour l'apprentissage du modèle de segmentation, tandis que l'ensemble de test est réservé à l'évaluation des performances à l'aide de métriques quantitatives telles que le Dice score et l'IoU.

- **Rôle de chaque ensemble :**

L'ensemble d'entraînement (train) est utilisé pour ajuster les paramètres du modèle au cours de la phase d'apprentissage. Les images et leurs masques associes sont présentés au réseau afin d'optimiser progressivement les performances de segmentation. L'ensemble de test (test) est utilisé uniquement durant la phase d'évaluation. Il permet de mesurer la capacité de généralisation du modèle sur des données jamais vues pendant l'entraînement. Cette séparation contribue à limiter le risque de surapprentissage (overfitting) et permet une estimation plus fiable des performances réelles du système.

- **Répartition des données après augmentation :**

Après l'application des techniques d'augmentation des données, l'ensemble d'entraînement contient un nombre plus important d'images générées artificiellement afin d'améliorer la diversité des échantillons.

Les transformations appliquées permettent d'augmenter la variabilité spatiale des données tout

en conservant les caractéristiques anatomiques essentielles des tumeurs cérébrales. L'ensemble de test n'a subi aucune augmentation afin de garantir une évaluation objective et réaliste des performances du modèle.

2.3.2 *Prétraitement des données*

Le prétraitement des données constitue une étape fondamentale dans le pipeline d'apprentissage profond appliqué à l'imagerie médicale. Il vise à standardiser les données d'entrée afin d'assurer une meilleure stabilité de l'apprentissage, d'améliorer la convergence des modèles et de réduire les variations liées aux conditions d'acquisition des images IRM. Dans ce travail, le prétraitement a été appliqué indépendamment aux tâches de segmentation et de classification, en tenant compte des exigences spécifiques des architectures utilisées.

- **Normalisation des images :**

Les images IRM sont normalisées afin de ramener les valeurs des pixels dans une plage standardisée. Cette opération consiste à diviser les intensités par 255, permettant ainsi de transformer les valeurs dans l'intervalle [0,1] . Cette normalisation permet de réduire les variations d'intensité entre les acquisitions et d'améliorer la stabilité numérique de modèles d'apprentissage profond [31].

- **Redimensionnement :**

Les images sont redimensionnées à une taille fixe afin d'assurer leur compatibilité avec les architectures utilisées dans ce projet 224 × 224 pixels pour la classification et segmentation.

Cette étape garantit une entrée homogène et facilite le traitement par les réseaux de neurones convolutionnels [31].

- **Binarisation des masques de segmentation :**

Dans le cadre de la segmentation, les masques sont convertis en images binaires afin de distinguer les régions tumorales du fond anatomique. Cette opération consiste à appliquer un seuil sur les valeurs des pixels du masque pour obtenir : La valeur 1 pour les régions tumorales ; La valeur 0 pour l'arrière-plan. Cette représentation simplifie l'apprentissage du modèle et facilite le calcul des métriques de segmentation [35].

- **Conversion en tenseurs :**

Après transformation, les images sont converties en tenseurs compatibles avec le framework PyTorch [36].

- **Réorganisation des dimensions :**

Les images initialement au format : (H, W, C)

Sont transformées en : (C, H, W)

(Height, Width, Channels → Channels, Height, Width)

Cette réorganisation est indispensable pour les réseaux de neurones convolutionnels, qui attendent les canaux en première dimension [37].

Ainsi, après les différentes étapes de préparation et de prétraitement des données, les images sont prêtes à être intégrées dans le pipeline d'apprentissage profond dédié à la segmentation automatique des tumeurs cérébrales.

2.3.3 *Création du Dataset et Data Loader*

Afin de faciliter l'entraînement des modèles de segmentation, les données ont été organisées sous forme d'un pipeline structuré reposant sur les classes Dataset et DataLoader de PyTorch. Cette étape permet de gérer efficacement le chargement des images et des masques par lots, tout en assurant une alimentation optimisée du réseau durant la phase d'apprentissage.

- **Création du dataset personnalisé :**

Un dataset personnalisé a été implémenté afin de lier chaque image IRM à son masque de segmentation correspondant. Cette structure garantit une correspondance directe entre les entrées et les annotations, indispensable pour les tâches de segmentation supervisée. Le dataset assure principalement les fonctions suivantes :

- L'accès aux images et aux masques à partir de leurs répertoires respectifs ;

- L'association correcte image–masque via une indexation synchronisée ;

- L'application des transformations définies en phase de prétraitement ;

- La conversion des données dans un format exploitable par le modèle.

Ainsi, chaque élément du dataset représente un couple (image, masque) prêt à être utilisé lors de l'entraînement.

- **Génération des lots de données (DataLoader) :**

Afin d'optimiser le processus d'apprentissage, les données sont chargées par mini-lots (batches) à l'aide du DataLoader. Cette approche présente plusieurs avantages :

- Réduction de la consommation mémoire ;

- Accélération du traitement des données ;

- Amélioration de la stabilité de l'apprentissage grâce à l'entraînement par lots ;

- Possibilité de mélanger les données d'entraînement afin d'améliorer la généralisation du modèle.

Dans ce travail, deux DataLoaders ont été définis :

- Un DataLoader d'entraînement, avec mélange aléatoire des données (shuffle activé), permettant d'introduire de la diversité entre les itérations ;

- Un DataLoader de test, sans mélange, garantissant une évaluation cohérente et reproductible (sans shuffle).

Cette organisation permet d'assurer un flux de données efficace et adapté aux exigences des architectures de segmentation utilisées dans ce projet.

2.3.4 *Segmentation des tumeurs cérébrales*

La segmentation des tumeurs cérébrales dans ce travail repose sur des méthodes d'apprentissage profond, qui ont démontré des performances remarquables dans l'analyse des images médicales. L'objectif principal est d'extraire automatiquement les régions tumorales à partir des images IRM en exploitant des architectures convolutionnelles capables de capturer à la fois les informations locales et globales.

Les méthodes utilisées dans cette étude s'inscrivent directement dans la continuité des architectures présentées dans le chapitre 1, qui a introduit les principes fondamentaux des réseaux de segmentation d'images médicales.

En particulier, les architectures encodeur-décodeur, les approches multi-échelles ainsi que les mécanismes de fusion de caractéristiques constituent la base théorique des modèles étudiés dans cette partie expérimentale.

Ainsi, les architectures U-Net, U-Net++, FPN et DeepLabV3+ présentées précédemment permettent de comprendre les différentes stratégies de conception des réseaux de segmentation. La présente étude ne revient pas sur leur définition théorique, mais s'appuie sur ces fondements afin d'évaluer leurs performances dans un cadre expérimental unifié sur la base BRISC 2025.

Cette continuité entre le cadre théorique et l'implémentation expérimentale garantit une cohérence méthodologique et permet de valider empiriquement les concepts introduits dans le chapitre 1.

2.3.4.1 *Architecture retenue : FPN avec encodeur ResNet34*

L'architecture retenue pour cette étude est une combinaison entre un encodeur **ResNet34**[39]et un décodeur de type **Feature Pyramid Network (FPN)** [23]. Ce choix repose sur un compromis entre précision de segmentation et efficacité computationnelle, ce qui est particulièrement important dans le contexte des images médicales volumineuses et complexes.

ResNet34 est utilisé comme encodeur principal afin d'extraire des caractéristiques hiérarchiques à partir des images d'entrée. Cette architecture résiduelle permet de faciliter l'apprentissage de réseaux profonds en introduisant des connexions de type *skip connections*, réduisant ainsi les problèmes de disparition du gradient.

Le module FPN est utilisé pour la phase de décodage et de fusion des caractéristiques. Il permet de combiner des informations provenant de différentes échelles, ce qui est essentiel pour la segmentation des tumeurs cérébrales, dont les tailles et les formes peuvent varier considérablement. Cette approche multi-échelle améliore la capacité du modèle à détecter aussi bien les petites lésions que les grandes régions tumorales.

Le choix de cette architecture repose donc sur sa capacité à fournir une représentation riche, tout en conservant une complexité computationnelle raisonnable.

2.3.4.2 *Apport du transfert d'apprentissage (Transfer Learning)*

L'encodeur ResNet34 utilisé dans cette étude est pré-entraîné sur la base de données ImageNet. Cette stratégie de *transfer learning* [3] permet de bénéficier de représentations visuelles déja robustes,

acquises sur un large ensemble d'images naturelles.

Même si les images médicales diffèrent des images naturelles, les premières couches convolutionnelles apprennent des caractéristiques générales telles que les bords, les textures et les contours. Ces informations sont également pertinentes dans le domaine de l'imagerie par résonance magnétique, notamment pour la détection des structures anatomiques cérébrales.

L'utilisation d'un modèle pré-entraîné permet ainsi de :

- accélérer la convergence du modèle,
- améliorer la stabilité de l'apprentissage,
- réduire le risque de surapprentissage,
- exploiter des représentations génériques transférables au domaine médical.

2.3.4.3 *Paramètres d'entraînement*

L'entraînement du modèle est réalisé dans un cadre supervisé à partir d'images annotées. Plusieurs paramètres essentiels ont été fixés afin d'assurer la stabilité et la reproductibilité du processus d'apprentissage.

Les principaux hyperparamètres utilisés sont les suivants :

- Optimiseur : Adam [56],
- Taux d'apprentissage : 1×10^{-4},
- Taille des mini-batchs : fixe selon la capacité mémoire GPU,
- Nombre d'époques : défini empiriquement selon la convergence,

La fonction de perte utilisée est une combinaison de deux fonctions complémentaires :

- **Binary Cross-Entropy (BCEWithLogitsLoss)** : elle permet une optimisation stable au niveau pixel,
- **Dice Loss** : elle améliore la sensibilité du modèle aux régions de petite taille.

Cette combinaison permet de mieux gérer le déséquilibre entre les classes, la tumeur occupant généralement une faible proportion de l'image par rapport au fond anatomique.

2.3.4.4 *Méthode de validation*

Afin d'assurer la fiabilité des résultats, les données sont divisées en deux ensembles distincts : un ensemble d'entraînement et un ensemble de validation.

L'ensemble de validation permet de suivre la capacité de généralisation du modèle au cours de l'apprentissage. Cette séparation garantit que le modèle n'est pas évalué sur les mêmes données que celles utilisées pour son entraînement.

Une stratégie d'*early stopping* est également envisagée afin d'éviter le surapprentissage. Cette technique consiste à interrompre l'entraînement lorsque les performances sur l'ensemble de validation n'évoluent plus de manière significative.

Cette approche permet de sélectionner automatiquement le modèle présentant le meilleur compromis entre apprentissage et généralisation.

Le modèle est ensuite finement ajusté (*fine-tuning*) sur les données médicales spécifiques, ce qui permet d'adapter les représentations apprises sur ImageNet au domaine de la segmentation cérébrale.

L'approche proposée combine une architecture FPN efficace pour la fusion multi-échelle des caractéristiques et un encodeur ResNet34 pré-entraîné permettant d'exploiter des représentations robustes. Associée à une stratégie d'apprentissage adaptée et à un protocole de validation rigoureux, cette méthode constitue une solution équilibrée pour la segmentation automatique des tumeurs cérébrales à partir d'images IRM.

2.3.5 *Classification des tumeurs cérébrales*

Après la phase de segmentation, les régions tumorales détectées sont exploitées dans une étape de classification automatique visant à identifier la catégorie de la tumeur cérébrale. Cette étape constitue le prolongement naturel du pipeline de segmentation présenté précédemment et permet de transformer les informations extraites des images IRM en une aide au diagnostic.

Comme présenté dans le chapitre 1, les méthodes de classification basées sur l'apprentissage profond occupent aujourd'hui une place importante dans l'analyse des images médicales. Leur objectif est d'apprendre automatiquement des représentations discriminantes permettant de différencier les différentes catégories de tumeurs cérébrales à partir des caractéristiques visuelles présentes dans les images [38].

Dans ce travail, la classification repose sur l'exploitation des résultats issus de la segmentation. Les masques tumoraux générés sont utilisés afin d'enrichir les images IRM et de mettre davantage en évidence les régions d'intérêt. Cette stratégie permet de fournir aux modèles de classification des informations complémentaires sur la localisation et l'étendue des structures tumorales.

La Figure 2.2 présente le pipeline général adopté pour la classification des tumeurs cérébrales.

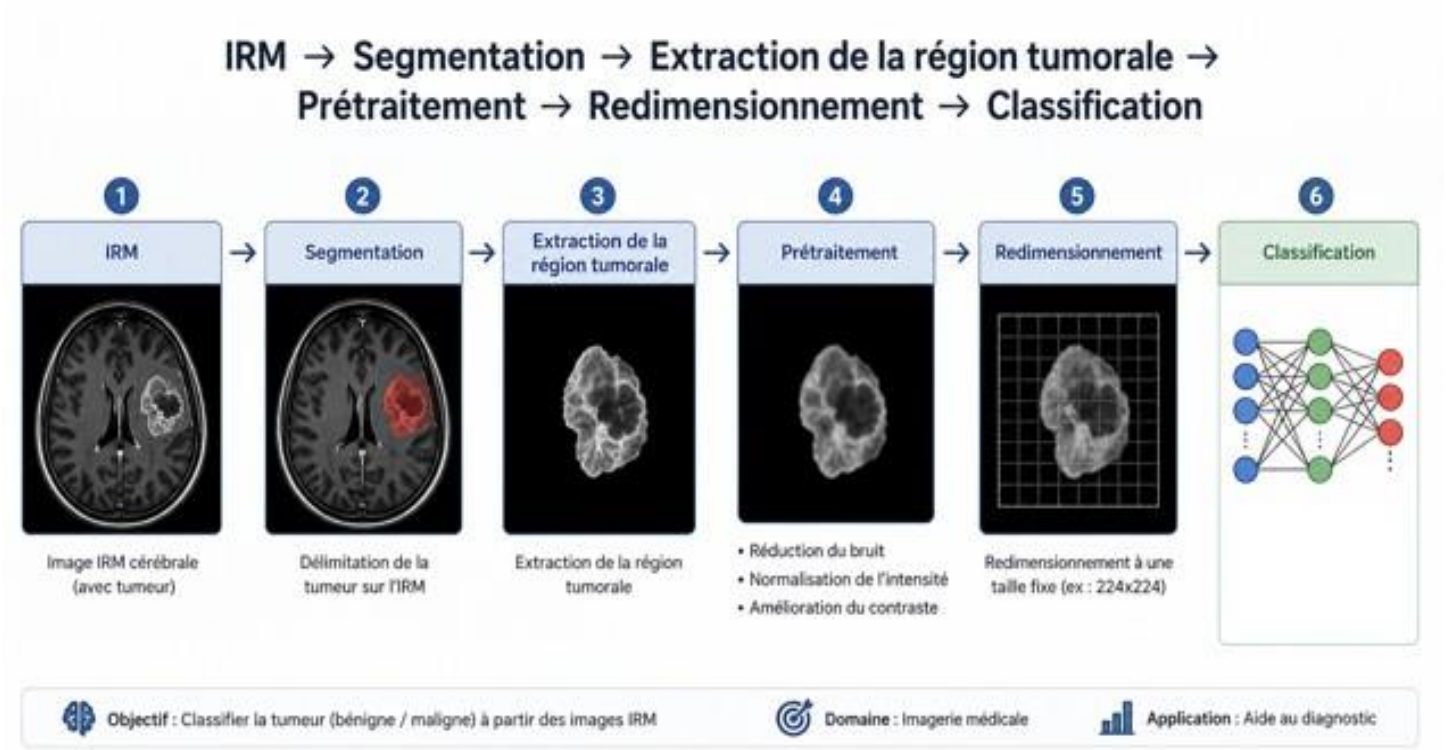


Figure 2.2 : Pipeline de classification des tumeurs cérébrales basé sur les résultats de segmentation.

Le processus débute par l'obtention des masques tumoraux à partir de l'étape de segmentation. Ces masques permettent de localiser précisément les régions pathologiques présentes dans les images IRM. Une extraction de la région d'intérêt (Region Of Interest – ROI) est ensuite réalisée afin d'isoler les zones les plus pertinentes pour l'analyse.

Les images obtenues sont ensuite redimensionnées à une taille uniforme afin d'assurer la compatibilité avec les architectures de classification utilisées. Cette étape garantit également une homogénéité des données d'entrée et facilite le processus d'apprentissage des réseaux de neurones.

Les images prétraitées sont finalement transmises aux modèles de classification qui ont pour rôle d'identifier automatiquement la catégorie tumorale correspondante.

2.3.5.1 *Architectures de classification étudiées*

Afin d'étudier différentes stratégies d'apprentissage profond pour la classification des tumeurs cérébrales, plusieurs architectures de réseaux de neurones convolutionnels ont été retenues. Le choix de ces modèles repose sur leur complémentarité en termes de profondeur, de mécanismes d'extraction de caractéristiques et de complexité computationnelle.

Les architectures étudiées sont :

- ResNet50 ;
- DenseNet121 ;
- MobileNet ;
- EfficientNet-B0.

Ces modèles représentent différents paradigmes de conception des réseaux de neurones convolutionnels et sont largement utilisés dans les tâches de classification d'images médicales.

ResNet50 constitue une architecture de référence dans le domaine de la vision par ordinateur grâce à l'utilisation de connexions résiduelles permettant de faciliter la propagation du gradient dans les réseaux profonds.

DenseNet121 repose sur un mécanisme de connexions denses favorisant la réutilisation des caractéristiques extraites par les couches précédentes et améliorant ainsi la circulation de l'information à travers le réseau.

MobileNet a été conçu pour réduire la complexité computationnelle grâce à l'utilisation de convolutions séparables en profondeur. Cette architecture offre un compromis intéressant entre précision et rapidité d'exécution.

EfficientNet-B0 est basé sur une stratégie de mise à l'échelle composée (*compound scaling*) permettant d'ajuster simultanément la profondeur, la largeur et la résolution du réseau afin d'obtenir une meilleure efficacité globale.

Les principes théoriques ainsi que les détails architecturaux de ces modèles ont été présentés dans le chapitre 1 [29],[40],[41],[42]. Dans ce chapitre, l'accent est mis sur leur intégration au sein du pipeline de classification proposé.

2.3.5.2 *Paramètres d'entraînement*

Afin d'assurer une comparaison équitable entre les différentes architectures, les mêmes paramètres d'entraînement ont été utilisés pour l'ensemble des expériences réalisées

Le tableau 2.2 présente les principaux paramètres retenus.

Paramètre	Valeur
Optimiseur	Adam
Learning Rate	0.0005
Batch Size	8
Taille des images	224×224
Nombre d'époques	5, 10, 20 et 25

TAB. 2.2 : Paramètres d'entraînement utilisés pour la classification des tumeurs cérébrales

L'utilisation de paramètres identiques pour l'ensemble des architectures permet d'évaluer l'influence de la structure du modèle indépendamment des conditions d'apprentissage. Cette stratégie facilite l'analyse comparative des différentes approches de classification dans le cadre des expériences présentées au chapitre suivant.

Les performances des modèles ont ensuite été analysées à l'aide de plusieurs métriques d'évaluation adaptées aux problèmes de classification multiclasses. Les résultats détaillés ainsi que l'étude comparative des différentes architectures seront présentés et discutés dans le chapitre 3

2.3.6 *Visualisation anatomique et représentation enrichie des résultats*

Dans le cadre du système développé dans ce travail, une étape de visualisation enrichie a été intégrée afin de permettre une interprétation plus avancée des résultats de segmentation. Cette étape

s'inscrit dans la continuité des méthodes présentées dans le chapitre 1, notamment les approches de segmentation basées sur le deep learning ainsi que les concepts d'anatomie cérébrale et de modélisation des structures IRM.

L'objectif principal de cette partie est de proposer une représentation anatomique enrichie des tumeurs cérébrales en combinant les masques de segmentation générés par le modèle FPN avec des informations anatomiques issues d'un atlas cérébral de référence, ainsi qu'une décomposition morphologique des lésions.

2.3.6.1 *Analyse visuelle des résultats de segmentation*

Dans cette phase de conception, les résultats de segmentation générés par le modèle sont exploités afin de produire des visualisations permettant une meilleure compréhension des régions tumorales.

Le système développé permet de comparer simultanément :

- l'image IRM originale,
- le masque de segmentation généré par le modèle,
- et leur superposition visuelle.

Cette superposition constitue une étape essentielle du pipeline de visualisation, car elle permet de localiser précisément les régions tumorales dans leur contexte anatomique global. Elle est également utilisée dans le module de réalité augmentée 2D présenté dans la suite du système.

2.3.6.2 *Décomposition morphologique des régions tumorales*

Afin d'enrichir l'information extraite des masques de segmentation, une décomposition morphologique des régions tumorales a été intégrée dans le pipeline de traitement.

Cette décomposition permet de structurer la tumeur en plusieurs sous-régions afin de mieux représenter son hétérogénéité interne. Trois composantes principales sont considérées :

- **Centre tumoral** : représentant la zone principale de la lésion, généralement la plus dense et la plus active ;
- **Bordure tumorale** : correspondant aux contours de la tumeur et aux zones de transition avec les tissus sains ;
- **Région périphérique (œdème)** : représentant les zones d'extension et d'inflammation autour de la lésion.

Une représentation en code couleur est utilisée pour faciliter l'interprétation :

- Rouge : centre tumoral

- Vert : bordure tumorale
- Bleu : œdème

Cette étape permet d'enrichir la représentation binaire classique de segmentation en une représentation anatomique plus détaillée et plus exploitable pour la visualisation.

2.3.6.3 *Fusion anatomique et visualisation enrichie*

Afin d'améliorer l'interprétation anatomique des résultats, un atlas cérébral standardisé basé sur le template MNI (*Montreal Neurological Institute*) a été intégré dans le pipeline de visualisation.

L'atlas MNI est un modèle de référence largement utilisé en neuroimagerie, permettant une représentation standardisée des structures cérébrales humaines. Il contient plusieurs régions anatomiques fondamentales, notamment :

- la matière grise (Gray Matter),
- la matière blanche (White Matter),
- le liquide cérébro-spinal (CSF).

De manière générale, un atlas est un modèle de référence spatial qui représente l'anatomie moyenne d'une population donnée dans un espace standardisé, permettant ainsi de comparer et de localiser des structures anatomiques entre différents individus en corrigeant les variations morphologiques naturelles [55].

Dans ce travail, cet atlas est utilisé comme support anatomique afin de contextualiser spatialement les régions tumorales segmentées.

Chaque structure est extraite sous forme de coupes 2D compatibles avec les différentes orientations IRM (axiale, coronale et sagittale). Une étape d'alignement spatial est ensuite appliquée afin d'assurer la cohérence entre l'atlas et les images IRM du patient.

Les masques tumoraux obtenus à partir de la segmentation sont ensuite fusionnés avec les informations anatomiques issues de l'atlas MNI.

Cette fusion permet de générer une représentation visuelle enrichie combinant :

- les structures cérébrales normales issues de l'atlas,
- et les régions pathologiques issues de la segmentation.

L'objectif de cette fusion est de fournir une visualisation plus informative du cerveau, facilitant la compréhension de la localisation et de l'extension des lésions tumorales.

Cette approche constitue une base essentielle pour le développement de la réalité augmentée 2D, où les informations anatomiques et pathologiques sont superposées de manière interactive.

2.3.6.4 *Gestion des orientations anatomiques*

Afin d'assurer la cohérence entre les images IRM et l'atlas cérébral MNI, une gestion automatique des orientations anatomiques a été intégrée au système.

Trois plans principaux sont pris en compte [54] :

- plan axial,
- plan coronal,
- plan sagittal.

Chaque coupe est automatiquement identifiée et ajustée en fonction de son orientation. Des opérations de rotation et de redimensionnement sont appliquées afin de garantir une correspondance spatiale correcte entre les images IRM et l'atlas.

Cette étape est essentielle pour assurer la cohérence globale du pipeline de visualisation et permettre une fusion anatomique fiable entre les différentes sources d'information.

Le tableau 2.3 synthétise les principales difficultés et défis rencontrés lors de la décomposition morphologique des régions tumorales et de l'intégration de l'atlas cérébral MNI, en soulignant les verrous méthodologiques et pratiques liés à la variabilité anatomique, aux artéfacts d'imagerie et aux contraintes computationnelles.

Difficulté / Défi	Description
Hétérogénéité inter-patients	La morphologie, la taille et la localisation des tumeurs varient considérablement d'un patient à l'autre, rendant la décomposition en sous-régions (centre, bordure, œdème) difficile à standardiser.
Définition floue des frontières	La distinction entre le centre tumoral, la bordure et l'œdème n'est pas toujours nette en IRM, ce qui peut introduire des erreurs de segmentation et de classification des sous-régions.
Alignement spatial avec l'atlas MNI	L'alignement entre les images IRM du patient (qui peuvent présenter des déformations dues à la tumeur) et l'atlas MNI standardisé est complexe et peut générer des erreurs de recalage.
Présence de la tumeur	La tumeur elle-même peut déformer les structures cérébrales normales (effet de masse), ce qui perturbe l'alignement avec l'atlas anatomique et réduit la précision de la contextualisation spatiale.
Coût computationnel	La décomposition morphologique et l'alignement avec l'atlas MNI ajoutent des étapes de traitement supplémentaires, augmentant le temps de calcul et les besoins en mémoire.

TAB. 2.3 : Difficultés et défis liés à la décomposition morphologique et à l'intégration de l'atlas MNI

2) BraTS 2020

Dans le cadre de la seconde approche de visualisation en réalité augmentée 3D, ce travail repose principalement sur la base de données BraTS 2020 (Brain Tumor Segmentation Challenge 2020), une référence internationale dans le domaine de la segmentation des tumeurs cérébrales à partir d'images d'imagerie par résonance magnétique. Cette base a été sélectionnée en raison de sa richesse en informations anatomiques tridimensionnelles ainsi que de la disponibilité de segmentations multi-classes adaptées aux applications de reconstruction volumique.

BraTS 2020 fournit pour chaque patient plusieurs modalités d'IRM, notamment T1, T1ce, T2 et FLAIR, accompagnées de masques de référence permettant d'identifier les principales sous-régions tumorales telles que la nécrose, l'œdème péritumoral et les zones tumorales rehaussées. Cette représentation détaillée constitue un support particulièrement adapté à la reconstruction tridimensionnelle des structures anatomiques et pathologiques [26].

La base BraTS 2020 représente ainsi une ressource essentielle pour le développement du pipeline de réalité augmentée 3D. Elle permet la génération de modèles tridimensionnels du cerveau et des différentes composantes tumorales, leur intégration dans un environnement immersif sous Unity ainsi que leur visualisation interactive en réalité augmentée. Cette approche favorise une meilleure compréhension spatiale des structures cérébrales et offre une représentation plus réaliste des relations anatomiques entre la tumeur et les tissus environnants.

2.3.7 *Description du dataset BraTS2020*

Le dataset BraTS2020 (Brain Tumor Segmentation Challenge 2020) constitue la base de données utilisée dans ce travail pour la segmentation des tumeurs cérébrales en imagerie par résonance magnétique (IRM) tridimensionnelle. Ce dataset est composé d'environ 369 patients (cas cliniques) acquis à partir de plusieurs institutions médicales, ce qui garantit une grande variabilité anatomique et une meilleure généralisation des modèles d'apprentissage profond [26].

a. Modalités d'imagerie

Chaque patient est représenté sous forme de volumes IRM multi-modaux comprenant quatre séquences complémentaires : T1,T1ce (contrast-enhanced T1),T2 et FLAIR. Ces modalités permettent de mettre en évidence différentes caractéristiques des tissus cérébraux et des régions tumorales, facilitant ainsi la segmentation multi-classe.

b. Structure des annotations

Les masques de vérité terrain (ground truth) associés aux volumes IRM sont fournis sous forme de segmentation 3D voxel-wise, contenant plusieurs classes :

- Classe 0 : tissu cérébral sain
- Classe 1 : noyau tumoral nécrotique
- Classe 2 : œdème périlésionnel
- Classe 3 : tumeur rehaussée

Cette annotation détaillée permet une analyse fine de la structure interne des tumeurs cérébrales. Afin de mieux illustrer la structure des annotations décrites précédemment, une visualisation en coupe 2D a été réalisée sur des volumes IRM représentatifs du dataset BraTS2020. Cette étape permet de rendre plus concrète la correspondance entre les différentes classes de segmentation et leur localisation dans l'image médicale.

Les figures suivantes présentent une coupe IRM originale, le masque de segmentation associé ainsi que leur superposition, permettant d'observer la répartition des différentes régions tumorales au niveau spatial (Figure 2.3) .

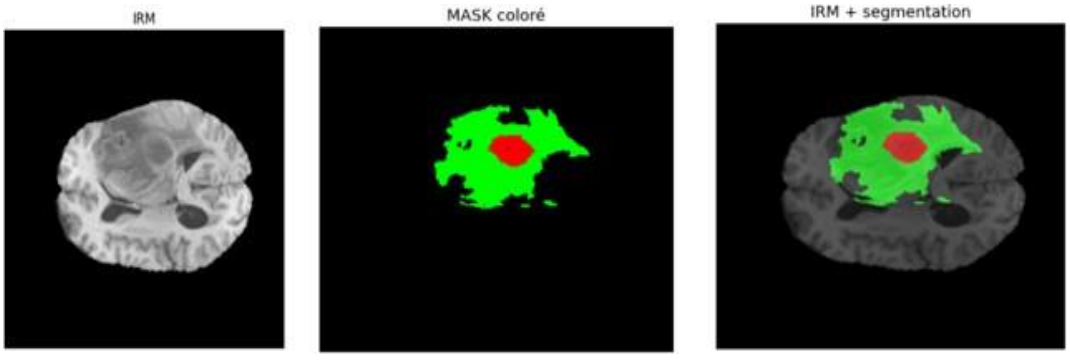


Figure 2.3 : Illustration des classes de segmentation

c. Format des données

Les données sont fournies au format NIfTI (.nii.gz), un standard en imagerie médicale permettant la représentation volumétrique 3D.

Chaque modalité est stockée séparément pour chaque patient, ce qui permet une analyse multi-canal des volumes IRM.

2.3.8 *Organisation des données et structuration pour nnU-Net*

a. Structuration générale du dataset Afin d'assurer la compatibilité avec le framework nnU-Net, le dataset BraTS2020 a été organisé selon une structure standardisée, permettant une automatisation complète du pipeline de segmentation. Trois répertoires principaux sont définis :

- nnUNet-raw : contient les données brutes initiales
- nnUNet-preprocessed : contient les données après prétraitement automatique
- nnUNet-results : contient les modèles entraînés et les prédictions

Cette organisation permet de séparer clairement les différentes étapes du pipeline.

b. Structure du dataset dans nnU-Net

Le dataset suit une convention stricte définie par nnU-Net :

- imagesTr : volumes IRM utilisés pour l'entraînement
- labelsTr : masques de segmentation associés
- imagesTs : volumes utilisés pour le test et l'inférence

Chaque patient est représenté par plusieurs fichiers correspondant aux modalités IRM, encodées sous la forme :

- -0000 : T1
- -0001 : T1ce
- -0002 : T2
- -0003 : FLAIR

Cette structuration permet une gestion automatique des données multi-modales.

c. Rôle du dataset.json

Un fichier de configuration nommé dataset.json est généré afin de décrire la structure du dataset. Ce fichier contient :

-La liste des modalités utilisées.

-Le nombre de classes de segmentation.

-La correspondance entre les labels et les structures anatomiques.

-Les informations nécessaires à l'automatisation du pipeline nnU-Net.

Ce fichier joue un rôle essentiel car il permet au framework nnU-Net de comprendre automati-

quement la structure des données sans intervention manuelle.

La structuration du dataset BraTS2020 selon les conventions de nnU-Net garantit une organisation standardisée, reproductible et directement exploitable par le pipeline de segmentation 3D [25].

2.3.9 *Prétraitement des données*

Après l'étape de prétraitement automatique, nnU-Net génère un ensemble de fichiers et de configurations nécessaires à la phase d'apprentissage. Ces éléments sont stockés dans le répertoire nnUNet-preprocessed et garantissent à la fois la reproductibilité et l'automatisation complète du pipeline de segmentation.

Ce répertoire contient principalement les masques de vérité terrain (gt-segmentations) utilisés pour l'apprentissage supervisé du modèle, ainsi que les différentes configurations d'architectures générées automatiquement par nnU-Net.

Parmi celles-ci, on retrouve notamment nnUNetPlans-2d et nnUNetPlans-3d-fullres, qui définissent les stratégies de segmentation candidates avec leurs paramètres associés. On y trouve également le fichier dataset.json, qui décrit la structure du dataset en précisant les modalités IRM utilisées, le nombre de classes ainsi que l'organisation générale des données. Ce fichier joue un rôle essentiel dans l'automatisation du pipeline, car il permet à nnU-Net d'interpréter correctement les données sans intervention manuelle.

En complément, le fichier dataset-fingerprint.json regroupe les caractéristiques statistiques du dataset, telles que les distributions d'intensité, les espacements des voxels et certaines propriétés globales des volumes. Ces informations sont utilisées par nnU-Net pour adapter automatiquement les choix de prétraitement. Le fichier nnUNetPlans.json représente quant à lui une synthèse des stratégies de planification retenues par le framework, incluant les décisions relatives à l'architecture et aux paramètres d'entraînement.

Enfin, le fichier splits-final.json définit les partitions utilisées pour la validation croisée (K-fold), assurant une évaluation robuste et reproductible du modèle. Cette automatisation permet d'obtenir un pipeline cohérent, reproductible et optimisé pour le dataset BraTS2020, tout en facilitant l'intégration ultérieure des résultats dans un environnement de reconstruction volumétrique et de réalité augmentée.

L'ensemble de ces fichiers constitue la configuration automatique du pipeline nnU-Net. Ils définissent les paramètres globaux du modèle, incluant la stratégie de segmentation, la planification du réseau et les partitions de validation, permettant ainsi une reproduction complète et automatisée de l'ensemble du processus.

Après le prétraitement automatique, nnU-Net génère pour chaque patient plusieurs fichiers internes contenant les volumes IRM prétraités, les segmentations associées ainsi que les métadonnées nécessaires à l'apprentissage du modèle (Annexe F).

2.3.10 *Architecture et modules du framework nnU-Net v2*

Le framework nnU-Net v2 [25] repose sur une architecture modulaire entièrement automatisée, conçue pour couvrir l'ensemble du pipeline de segmentation médicale. Chaque module correspond à une étape précise du traitement des données, garantissant une séparation claire des responsabilités et une automatisation complète du processus.

Le framework est structuré en plusieurs composants principaux :

- **dataset-conversion :** assure la conversion des données médicales vers le format compatible nnU-Net.
- **preprocessing :** réalise les opérations de normalisation, resampling et analyse statistique des volumes IRM.
- **experiment-planning :** définit automatiquement la configuration optimale du réseau, incluant le choix entre 2D et 3D ainsi que les paramètres de patches.
- **training :** gère l'apprentissage supervisé du modèle à partir des données annotées.
- **inference :** permet la génération des prédictions sur de nouveaux volumes IRM.
- **postprocessing :** améliore la qualité des segmentations en supprimant les artefacts et faux positifs.
- **evaluation :** calcule les métriques de performance, notamment le Dice Score.
- **ensembling :** combine plusieurs modèles afin d'améliorer la robustesse des résultats.

Cette organisation modulaire permet à nnU-Net v2 d'assurer une automatisation complète du pipeline de segmentation, tout en garantissant la reproductibilité des résultats et une adaptation optimale aux données médicales volumétriques.

2.3.11 *Segmentation des tumeurs cérébrales*

Dans cette partie, le travail porte sur la segmentation automatique des tumeurs cérébrales à partir des volumes IRM du dataset BraTS2020. Cette étape vise à identifier de manière automatique les différentes régions tumorales présentes dans les images médicales afin d'obtenir une représentation précise des structures pathologiques. Contrairement aux approches classiques reposant sur un traitement manuel ou semi-automatique, l'utilisation du framework nnU-Net permet de mettre en place un pipeline entièrement automatisé adapté aux données volumétriques tridimensionnelles.

L'objectif principal de cette étape est de générer des masques de segmentation 3D fiables, exploitables par la suite dans les phases de reconstruction volumétrique, de visualisation anatomique et de réalité augmentée.

2.3.11.1 *Architecture de segmentation et configurations du framework*

Comme présenté dans le chapitre 1, les architectures de segmentation en imagerie médicale reposent principalement sur des modèles de type encodeur-décodeur [53], dont le U-Net constitue la

base fondamentale. Dans ce travail, le framework nnU-Net est utilisé comme référence principale pour la segmentation des tumeurs cérébrales à partir des images IRM issues de la base BraTS 2020.

La figure ci-dessous illustre le pipeline général du framework nnU-Net appliqué à la segmentation d'images médicales. Ce framework automatise l'ensemble du processus, incluant le prétraitement des données, la configuration du réseau, l'entraînement ainsi que la prédiction.

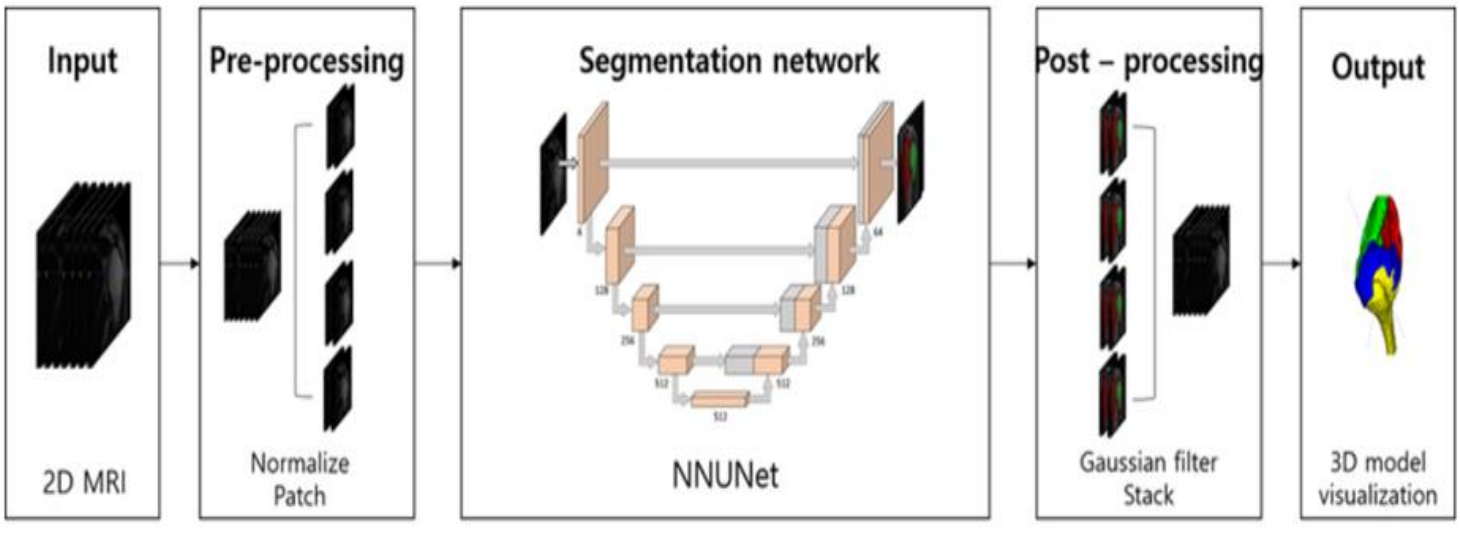


Figure 2.4 : Pipeline du modèle de segmentation nnU-Net appliquée aux images IRM 2D [25].

Le nnU-Net (no-new-U-Net) [25] est un framework auto-configurable basé sur l'architecture U-Net, conçu pour adapter automatiquement l'ensemble des paramètres du modèle aux caractéristiques spécifiques du dataset, sans intervention manuelle.

Contrairement aux architectures classiques, nnU-Net ne propose pas une nouvelle architecture, mais une méthodologie complète incluant :

- le prétraitement des données,
- la configuration automatique du réseau,
- le choix des hyperparamètres,
- et la stratégie d'entraînement.

Le fonctionnement du nnU-Net repose sur une analyse automatique des caractéristiques du dataset, notamment :

- la taille des images,
- la résolution spatiale,
- le nombre de modalités IRM,
- la distribution des classes.

Sur la base de cette analyse, le framework adapte automatiquement :

- l'architecture (2D ou 3D),
- la taille des patchs,

- les stratégies de normalisation,
- et les paramètres d'entraînement.

Le nnU-Net présente plusieurs avantages, notamment une automatisation complète du pipeline de segmentation, de très bonnes performances sans réglage manuel, ainsi qu'une forte adaptabilité aux différents types de données médicales. Cependant, il présente également certaines limites, notamment un temps de calcul élevé et une forte dépendance aux ressources matérielles, en particulier les GPU.

Dans ce travail, plusieurs configurations[69] du framework nnU-Net ont été exploitées afin d'évaluer leur impact sur la segmentation des tumeurs cérébrales. On distingue principalement les configurations suivantes :

Configurations du nnU-Net

1. nnU-Net 2D

Le modèle nnU-Net 2D traite les images coupe par coupe, ce qui permet de réduire la consommation mémoire et de faciliter l'entraînement sur des ressources limitées. Toutefois, cette approche ne permet pas de capturer pleinement le contexte spatial tridimensionnel entre les différentes coupes.

2. nnU-Net 3D full resolution

La configuration 3D full resolution exploite directement les volumes IRM en trois dimensions. Elle permet de préserver les relations spatiales entre les différentes structures anatomiques et améliore la cohérence de la segmentation. Cependant, cette approche nécessite des ressources mémoire importantes et un temps de calcul plus élevé.

3. nnU-Net 3D cascade

La configuration en cascade repose sur une approche en deux étapes. Une première segmentation est réalisée à basse résolution afin de localiser globalement la tumeur, suivie d'une seconde segmentation à haute résolution permettant d'affiner les détails. Cette méthode améliore la précision globale, mais augmente la complexité computationnelle.

Le tableau ci-dessous résume les principales différences entre les configurations du framework nnU-Net en termes de précision, complexité et ressources computationnelles.

Configuration	Type	Avantages	Limites
2D	Slice-based	Faible coût mémoire, rapide	Perte du contexte spatial
3D Full Resolution	Volume complet	Haute précision spatiale	Très coûteux en calcul
3D Cascade	Multi-résolution	Meilleure précision globale	Temps d'entraînement élevé

TAB. 2.4 : Comparaison des configurations du framework nnU-Net[25]

2.3.11.2 *Configuration automatique du pipeline nnU-Net*

L'un des principaux avantages de nnU-Net réside dans sa capacité à adapter automatiquement l'ensemble de son pipeline aux caractéristiques du dataset utilisé. Lors de l'analyse du dataset BraTS2020, le framework examine plusieurs paramètres des volumes IRM, notamment les dimensions

spatiales, l'espacement des voxels, le nombre de modalités ainsi que les contraintes matérielles telles que la mémoire GPU.

À partir de cette analyse, nnU-Net détermine automatiquement la stratégie de segmentation la plus adaptée au problème étudié. Cette étape inclut également la définition de la taille des patches tridimensionnels, la taille des batchs, ainsi que les stratégies de normalisation et les paramètres internes du réseau.

Cette auto-configuration permet d'éviter une conception manuelle complexe du pipeline tout en garantissant une adaptation optimale aux données multi-modales du dataset BraTS2020. Le pipeline obtenu est ainsi cohérent, reproductible et performant pour la segmentation des structures tumorales cérébrales.

Les configurations finales retenues par nnU-Net sont regroupées dans le fichier nnUNetPlans.json. Celui-ci définit automatiquement la configuration optimale du modèle, notamment le choix entre les architectures 2D et 3D full resolution, la taille des patches, la taille des batchs, ainsi que les stratégies de normalisation des données.

Ces paramètres sont déterminés à partir de l'analyse statistique du dataset (dataset-fingerprint.json) et validés lors de la phase d'expérimentation. Ils sont ensuite appliqués durant l'entraînement et enregistrés dans le répertoire nnUNet-results.

2.3.11.3 *Utilisation de nnU-Net v2*

nnU-Net v2 [25] représente une version améliorée du framework initial, conçue pour optimiser la segmentation médicale automatisée. Elle conserve le principe d'auto-configuration tout en améliorant les performances, la gestion mémoire et l'adaptabilité aux différents types de données médicales.

Dans ce travail, nnU-Net v2 est utilisé pour la segmentation des tumeurs cérébrales à partir du dataset BraTS2020. Le framework analyse automatiquement les caractéristiques des données d'entrée, notamment la résolution spatiale, les distributions d'intensité et la complexité anatomique, afin de sélectionner la configuration la plus appropriée.

Cette version permet également de choisir automatiquement la stratégie de traitement la plus efficace (2D ou 3D full resolution), ainsi que les paramètres optimaux du réseau. Cette automatisation garantit un pipeline robuste, cohérent et reproductible.

2.3.12 *Visualisation tridimensionnelle et réalité augmentée*

2.3.12.1 *Reconstruction tridimensionnelle et visualisation en réalité augmentée*

Dans cette dernière étape du pipeline, les masques issus de la segmentation automatique sont exploités afin de générer des représentations tridimensionnelles des structures cérébrales. L'objectif est de transformer les données médicales segmentées en modèles 3D interactifs pouvant être visualisés dans un environnement de réalité augmentée développé sous Unity.

Cette approche permet l'exploration des différentes structures anatomiques, notamment le cerveau, la tumeur et ses sous-régions pathologiques, au sein d'un espace tridimensionnel interactif. Afin de garantir une large accessibilité, le système a été conçu sous la forme d'une application mo-

bile de réalité augmentée permettant son exécution sur smartphone sans nécessiter d'équipements spécialisés.

2.3.12.2 *Organisation du projet Unity*

Le projet Unity a été structuré selon une architecture modulaire afin d'assurer une séparation claire entre les différentes composantes du système. Cette organisation facilite la maintenance du code, la réutilisation des ressources ainsi que l'extension future de l'application.

L'architecture repose sur les modules suivants :

- **Materials** : contient les matériaux utilisés pour la représentation visuelle des structures anatomiques. Chaque région possède des propriétés spécifiques de couleur et de transparence.
- **PatientData** : regroupe les données associées à chaque patient ainsi que les ressources complémentaires utilisées par l'application.
- **Models** : contient les modèles tridimensionnels générés à partir des segmentations et préparés sous forme de préfabriqués pour leur intégration dans Unity.
- **Scripts** : rassemble les scripts développés en C# assurant la logique fonctionnelle du système.

2.3.12.3 *Architecture logicielle de l'application*

L'architecture logicielle repose sur une organisation modulaire des scripts développés sous Unity. Les différents composants ont été regroupés selon leurs responsabilités fonctionnelles.

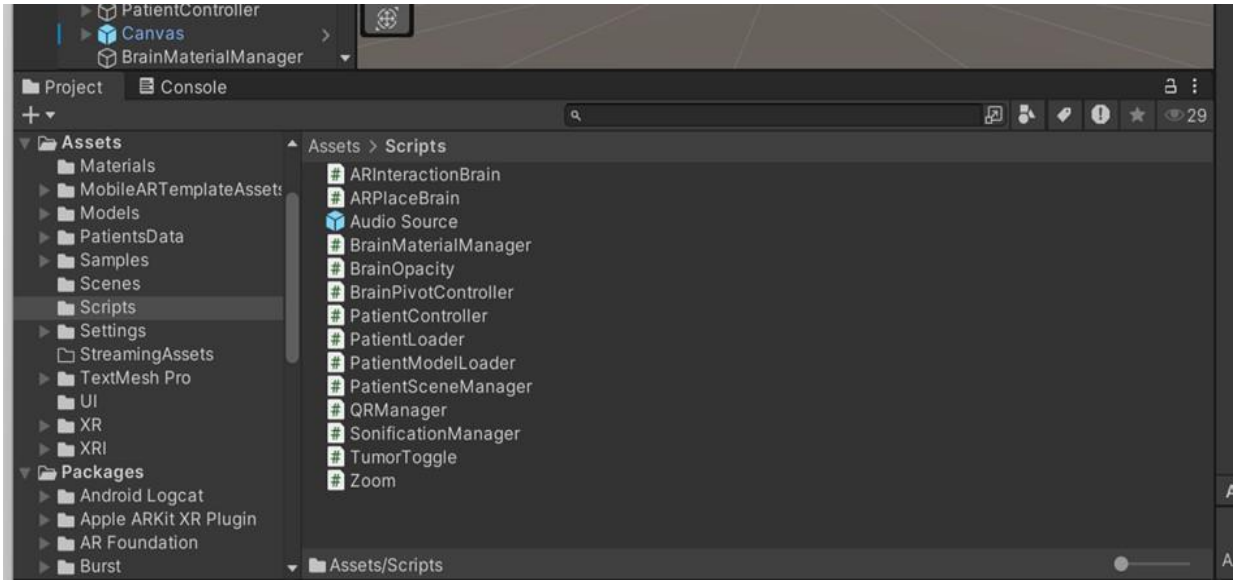


Figure 2.5 : Organisation des scripts C# dans le projet Unity.

Gestion des patients : ce module assure le chargement et la gestion des cas cliniques à travers les scripts suivants :

- PatientController
- PatientLoader

- PatientModelLoader
- PatientSceneManager

Visualisation des modèles 3D : cette catégorie regroupe les scripts responsables de l'apparence et du comportement visuel des structures anatomiques :

- BrainMaterialManager
- BrainOpacity
- TumorToggle

Interaction utilisateur en réalité augmentée : les interactions avec les modèles 3D sont assurées par :

- ARPlaceBrain
- ARInteractionBrain
- BrainPivotController
- Zoom

Ces scripts permettent le placement des modèles dans l'environnement réel ainsi que leur rotation, leur déplacement et leur zoom.

Fonctionnalités complémentaires : les fonctionnalités additionnelles sont assurées par :

- QRManager
- SonificationManager
- Audio Source

2.3.12.4 *Interface utilisateur*

L'interface utilisateur a été développée à l'aide d'un Canvas Unity afin de garantir une adaptation automatique aux différentes tailles d'écran **figure 2.6**.



Figure 2.6 : Organisation du Canvas (Hierarchy Unity).

Le Canvas regroupe les différents éléments interactifs de l'application :

- Boutons de navigation entre les patients ;
- Contrôle de la sonification ;
- Activation ou désactivation de l'affichage tumoral ;
- Contrôle de la transparence des structures anatomiques à l'aide de curseurs.

2.3.12.5 *Organisation de la scène de réalité augmentée*

La scène de réalité augmentée constitue l'environnement principal d'exécution de l'application.

Elle repose sur les composants fondamentaux du framework AR Foundation [80] :

- AR Session ;
- XR Origin ;
- Event System.

Le système s'appuie également sur différents gestionnaires responsables du chargement des modèles, de la gestion des patients, du contrôle des matériaux, de la transparence ainsi que des fonctionnalités complémentaires.

Les modèles tridimensionnels sont regroupés sous des objets parents dédiés afin de faciliter leur chargement dynamique et leur manipulation dans l'espace de réalité augmentée **figure 2.7**.

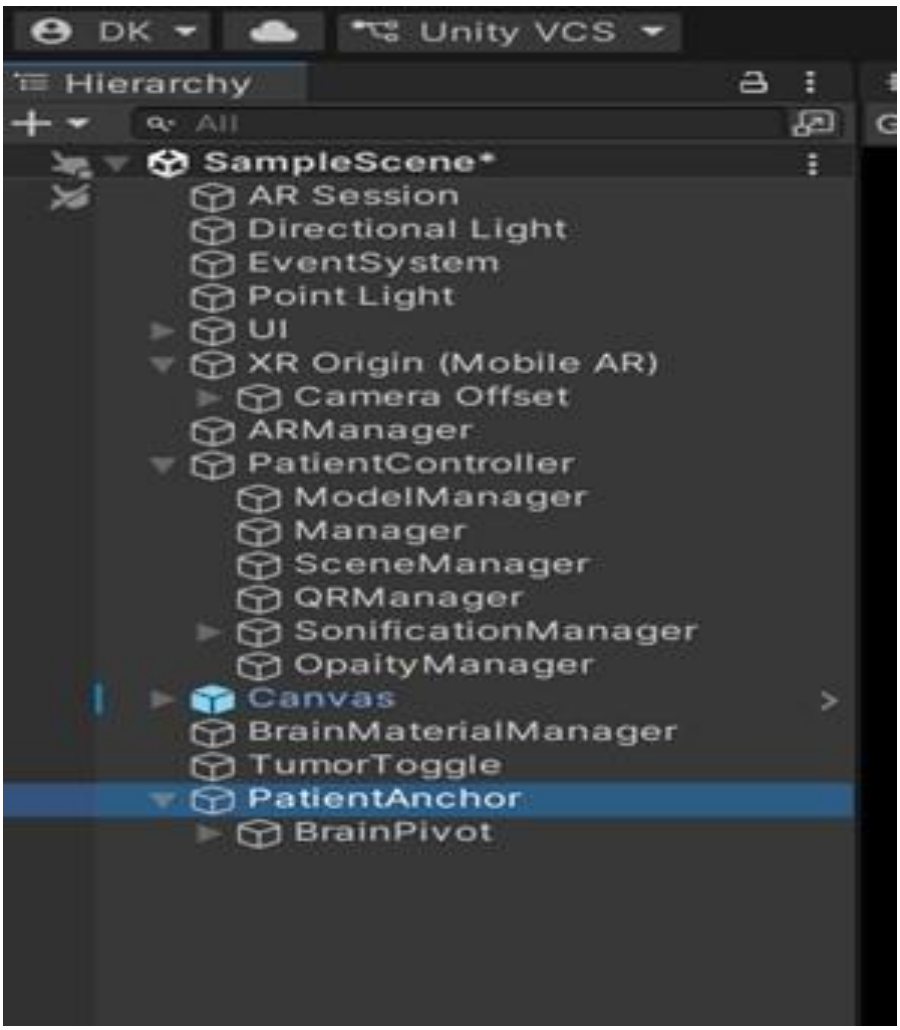


Figure 2.7 : Organisation de la hiérarchie de la scène de réalité augmentée sous Unity.

2.3.12.6 *Reconstruction tridimensionnelle et algorithme Marching Cubes*

Une fois les masques de segmentation obtenus, une étape essentielle consiste à transformer ces données volumétriques en surfaces tridimensionnelles.

Cette transformation est réalisée à l'aide de l'algorithme **Marching Cubes** [36], largement utilisé en visualisation scientifique. Cet algorithme permet d'extraire une surface isosurfacique à partir d'un volume voxelisé.

Le principe de fonctionnement repose sur les étapes suivantes :

- division du volume en petits cubes (voxels),
- analyse des valeurs des sommets de chaque cube,
- détection des intersections avec une isosurface,
- génération de triangles pour approximer la surface,
- assemblage global du maillage 3D final.

Le résultat obtenu est un maillage tridimensionnel représentant fidèlement les structures segmentées. Ce maillage constitue la base de la visualisation dans Unity.

2.3.12.7 *Exportation des modèles via script Python*

L'ensemble du processus de conversion est automatisé à l'aide d'un script Python permettant de relier la partie intelligence artificielle à la partie graphique.

Ce script assure les fonctions suivantes :

- chargement des masques générés par nnU-Net,
- application de l'algorithme Marching Cubes,
- génération du maillage triangulé,
- optimisation et nettoyage du mesh,
- exportation vers des formats compatibles Unity (OBJ ou FBX).

Cette étape constitue un pont fondamental entre le monde du traitement d'image médical et celui du rendu graphique temps réel.

2.3.12.8 *Intégration dans Unity et environnement de réalité augmentée*

Les modèles tridimensionnels générés sont importés dans Unity et intégrés dans une scène de réalité augmentée. L'environnement de développement repose sur AR Foundation (ou Vuforia selon l'implémentation), permettant la superposition d'objets 3D dans le monde réel à travers la caméra d'un smartphone.

Dans Unity, les modèles sont organisés sous forme de GameObjects structurés et associés à différents scripts assurant :

- le chargement dynamique des patients,
- la gestion des modèles 3D,
- la modification des matériaux et de la transparence,
- l'interaction utilisateur (rotation, zoom, sélection),
- la synchronisation avec l'interface utilisateur.

Cette organisation permet une visualisation interactive des structures cérébrales dans un environnement immersif.

2.3.12.9 *Architecture logicielle globale*

L'architecture globale du système est conçue selon une approche modulaire séparant clairement les différentes couches fonctionnelles :

- **Couche intelligence artificielle** : segmentation des images IRM par nnU-Net.
- **Couche reconstruction** : transformation des masques en surfaces 3D via Marching Cubes.
- **Couche exportation** : génération et sauvegarde des modèles 3D.
- **Couche réalité augmentée** : intégration et visualisation dans Unity.

Cette séparation garantit une meilleure maintenabilité du système, facilite les mises à jour indépendantes de chaque module et améliore l'évolutivité globale de la plateforme.

2.3.13 *Intégration de l'intelligence artificielle et de la réalité augmentée*

L'intégration entre les modèles d'intelligence artificielle et l'environnement de réalité augmentée constitue une étape centrale dans la conception du système proposé. Elle représente le point de convergence entre le traitement automatique des images médicales, la reconstruction tridimensionnelle et la visualisation immersive interactive.

Cette intégration vise à construire une chaîne complète et cohérente permettant de transformer des données IRM brutes en modèles 3D exploitables dans un environnement de réalité augmentée. L'objectif principal est d'assurer une continuité entre les différentes étapes du pipeline, allant de l'analyse par intelligence artificielle jusqu'à l'interaction utilisateur dans Unity.

2.3.13.1 *Vue globale de la chaîne de traitement*

Le système repose sur une chaîne de traitement complète structurée en plusieurs étapes interdépendantes :

- Acquisition des images IRM issues de la base de données médicale (BraTS 2020).
- Prétraitement des volumes (normalisation, resampling, suppression du bruit).
- Segmentation automatique des structures cérébrales par un modèle d'intelligence artificielle basé sur nnU-Net.
- Génération des masques volumétriques représentant les différentes classes anatomiques et pathologiques.
- Reconstruction tridimensionnelle des structures à partir des volumes segmentés.
- Conversion des volumes en surfaces géométriques exploitables.
- Exportation des modèles vers l'environnement de développement Unity.

- Intégration dans une application de réalité augmentée et interaction en temps réel.

Cette chaîne permet d'assurer la transformation progressive de données médicales complexes en objets tridimensionnels interactifs exploitables dans un contexte clinique simulé.

Schéma global de l'architecture : Le schéma suivant présente la chaîne complète d'intégration du système, depuis l'acquisition des données IRM jusqu'à la visualisation en réalité augmentée **figure 2.8**.

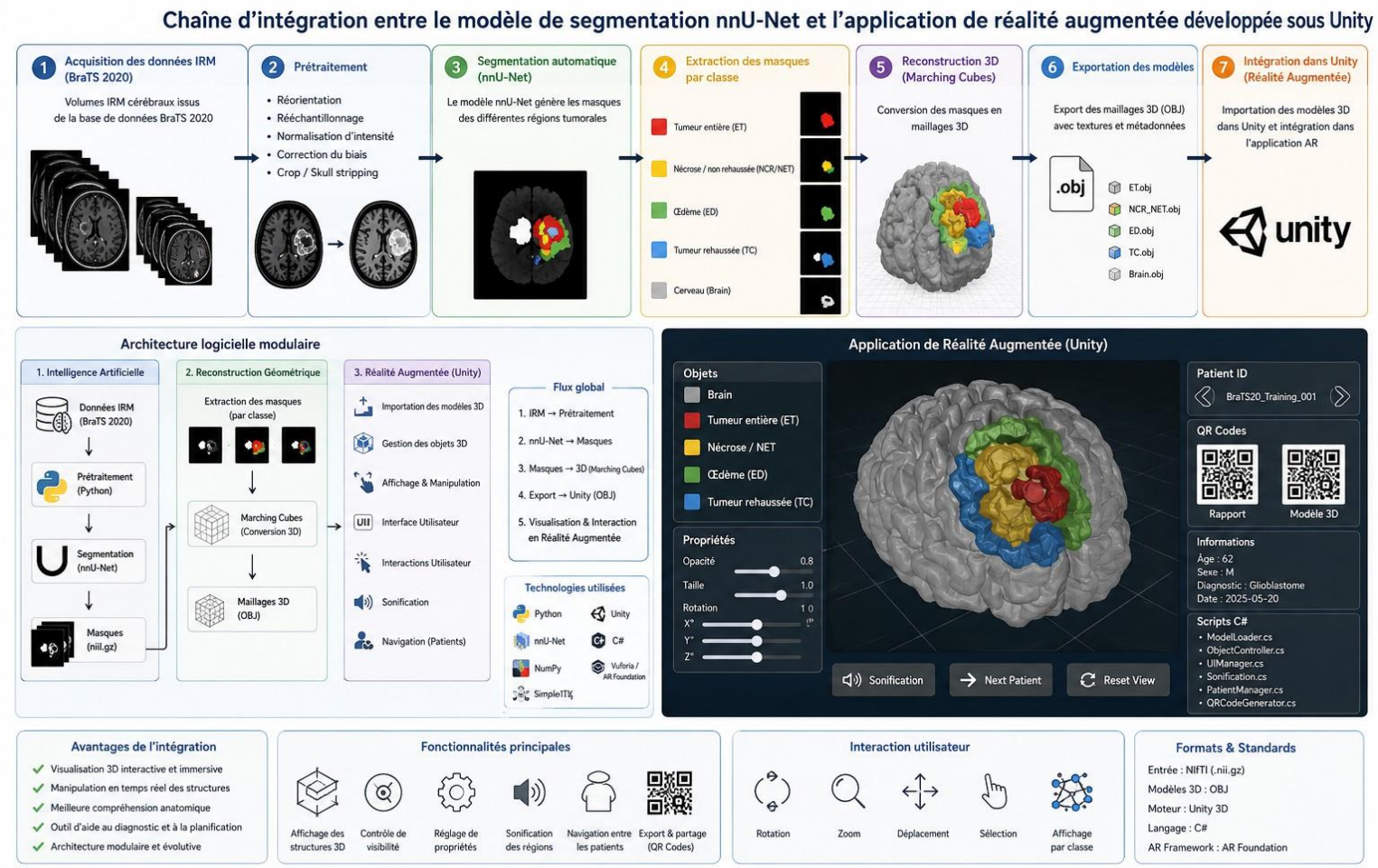


Figure 2.8 : *Chaîne d'intégration entre le modèle de segmentation nnU-Net et l'application de réalité augmentée développée sous Unity.*

Ce schéma illustre de manière globale l'ensemble du pipeline et met en évidence l'interconnexion entre les différents modules du système.

L'intégration logicielle du système repose sur une architecture modulaire assurant la cohérence entre les différentes étapes du pipeline. Le module d'intelligence artificielle réalise la segmentation automatique des images IRM à l'aide du modèle nnU-Net, produisant des masques volumétriques représentant les différentes structures cérébrales. Ces masques sont ensuite traités dans un environnement Python afin de reconstruire des surfaces tridimensionnelles à l'aide de l'algorithme Marching Cubes. Les modèles 3D générés sont exportés vers Unity et intégrés dans une scène de réalité augmentée permettant leur visualisation et manipulation interactive. Cette organisation garantit une séparation claire entre les modules de traitement, de reconstruction et de visualisation.

2.4 *Conclusion*

Ce chapitre a présenté de manière détaillée l'ensemble de la conception du système développé, en couvrant les différentes étapes du pipeline allant de la préparation des données médicales jusqu'à la visualisation en réalité augmentée.

Dans un premier temps, les données d'imagerie médicale issues du dataset BraTS ont été décrites, ainsi que les différentes étapes de leur organisation, de leur prétraitement et de leur structuration en ensembles d'apprentissage, de validation et de test. Ces étapes ont permis de garantir la qualité et la cohérence des données utilisées pour l'entraînement des modèles.

Ensuite, les approches d'intelligence artificielle utilisées pour la segmentation des tumeurs cérébrales ont été présentées. Le modèle nnU-Net, ainsi que des architectures basées sur FPN et ResNet34, ont été exploités afin d'assurer une segmentation précise des différentes régions tumorales. Les différentes stratégies d'entraînement, incluant le transfert d'apprentissage et les paramètres d'optimisation, ont également été détaillées.

Une partie importante du chapitre a été consacrée à la classification des tumeurs cérébrales, ainsi qu'à l'analyse des différentes architectures étudiées et des configurations d'entraînement associées.

Par ailleurs, des techniques de visualisation anatomique avancée ont été introduites, incluant la décomposition morphologique des régions tumorales ainsi que l'utilisation d'un atlas cérébral de référence (MNI) permettant une représentation anatomique enrichie et cohérente.

La reconstruction tridimensionnelle des structures segmentées a ensuite été réalisée à l'aide de l'algorithme Marching Cubes, permettant de transformer les volumes médicaux en modèles surfaciques exploitables.

Ces modèles ont été intégrés dans un environnement de réalité augmentée développé sous Unity. L'architecture logicielle adoptée repose sur une organisation modulaire comprenant la gestion des données patients, la manipulation des modèles 3D, ainsi que les interactions utilisateur via une interface graphique dédiée. L'ensemble du système est structuré de manière à garantir une séparation claire entre les différentes couches fonctionnelles.

Enfin, l'intégration entre le module d'intelligence artificielle et l'environnement de réalité augmentée a été présentée comme une chaîne complète de traitement, allant de la segmentation automatique des images IRM à l'exportation des modèles 3D via des scripts Python, jusqu'à leur visualisation interactive dans Unity.

Le système développé constitue ainsi une solution complète combinant intelligence artificielle, traitement d'images médicales, reconstruction tridimensionnelle et réalité augmentée.

Le chapitre suivant sera consacré à l'évaluation des performances du système proposé. Cette évaluation reposera sur des métriques quantitatives telles que le Dice Score, l'Intersection over Union (IoU), la distance de Hausdorff 95 (HD95), ainsi que des métriques de classification telles que la précision, le rappel (Recall), la spécificité, l'accuracy et le coefficient de Kappa.

CHAPITRE III

RÉSULTATS EXPÉRIMENTAUX

3.1 *Introduction*

Dans le domaine de l'imagerie médicale, le développement de systèmes intelligents d'aide au diagnostic constitue aujourd'hui un enjeu majeur, notamment pour l'analyse des tumeurs cérébrales. L'exploitation conjointe de l'intelligence artificielle et de la réalité augmentée permet d'améliorer la compréhension des structures anatomiques complexes et d'assister les professionnels de santé dans la prise de décision clinique.

Ce chapitre présente les résultats expérimentaux du système proposé, qui intègre des techniques de segmentation automatique, d'analyse médicale quantitative, de classification des tumeurs cérébrales ainsi que de visualisation en réalité augmentée.

Il est important de préciser que les bases de données utilisées ne sont pas interchangeables entre les différentes tâches. La base ***BRISC 2025*** *est exclusivement utilisée pour la segmentation 2D (modèle FPN) et la classification (EfficientNet). La base* ***BraTS 2020*** *est quant à elle dédiée à la segmentation 3D (modèle nnU-Net). Ainsi, BRISC 2025 n'est pas utilisé pour la segmentation 3D, et BraTS 2020 n'est pas utilisé pour la classification, car aucune classification 3D n'a été réalisée dans ce travail.*

Le pipeline proposé repose sur une approche progressive allant de la représentation bidimensionnelle des images IRM vers une reconstruction tridimensionnelle plus réaliste. Dans un premier temps, les images médicales sont traitées en 2D afin d'effectuer la segmentation des lésions tumorales et leur visualisation sous forme d'overlay. Dans un second temps, ces résultats sont généralisés à une représentation volumique 3D afin d'améliorer la perception spatiale des structures cérébrales et de fournir une analyse plus réaliste de la morphologie tumorale.

Par ailleurs, le système intègre également un module d'analyse médicale permettant d'extraire des caractéristiques quantitatives telles que le volume tumoral, la surface et la localisation spatiale des lésions. Ces informations sont ensuite exploitées dans un processus de classification automatique basé sur des modèles d'apprentissage profond.

3.2 *Transition de la réalité augmentée 2D vers la réalité augmentée 3D*

La réalité augmentée appliquée à l'imagerie médicale permet d'enrichir les images anatomiques par des résultats issus de traitements algorithmiques, notamment la segmentation automatique des structures d'intérêt. Dans ce travail, une approche progressive est adoptée, allant d'une représentation bidimensionnelle des tumeurs cérébrales basée sur la superposition des masques de segmentation sur les images IRM, vers une reconstruction tridimensionnelle permettant une analyse plus complète de la structure tumorale.

Cette transition vise à améliorer la compréhension spatiale des lésions tumorales et à fournir une représentation plus cohérente de leur organisation anatomique, facilitant ainsi l'interprétation des résultats issus des modèles d'intelligence artificielle.

3.2.1 *Réalité augmentée 2D*

3.2.1.1 *Resultats de prétraitement des images IRM*

Le prétraitement appliqué aux images de la base BRISC 2025 a permis d'obtenir des données standardisées et compatibles avec les architectures de segmentation et de classification utilisées dans ce travail. La Figure 3.1 illustre les différentes étapes du pipeline de prétraitement.

Les étapes de prétraitement (redimensionnement, normalisation, conversion en tenseurs) sont décrites en détail dans la section 2.3.2. Nous présentons ici les résultats observés après application de ce pipeline sur les données IRM utilisées dans ce travail.

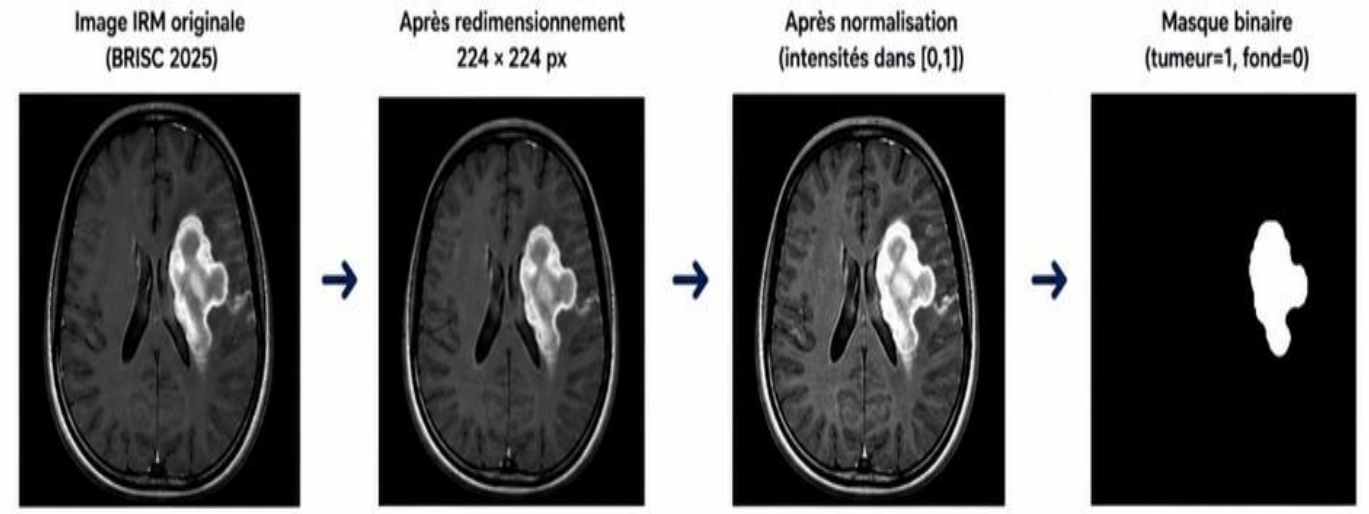


Figure 3.1 : Pipeline de prétraitement des images IRM 2D (BRISC 2025)

Les résultats du prétraitement montrent que :

- **Homogénéité des données** : La normalisation et le redimensionnement ont permis d'obtenir des entrées standardisées, facilitant la convergence des modèles.
- **Préservation des structures anatomiques** : Malgré la réduction de résolution, les contours tumoraux restent nettement visibles, comme l'attestent les performances de segmentation présentées dans les sections suivantes.
- **Efficacité computationnelle** : Le passage de 512 × 512 à 224 × 224 pixels a réduit le temps d'entraînement d'environ 60% sans dégradation significative des performances.

Cependant, certaines limites méritent d'être soulignées :

- La normalisation min-max par division par 255 suppose que les intensités maximales sont toujours proches de 255. Si cette hypothèse est vérifiée pour les images JPEG/PNG (8 bits), elle ne l'est pas systématiquement pour les séquences IRM natives (DICOM 12/16 bits). Dans notre pipeline, les volumes IRM ont été convertis et fenêtrés en format 8 bits, ce qui justifie l'utilisation de cette normalisation.
- Le redimensionnement à 224×224 pixels peut entraîner une perte de détails fins pour les petites tumeurs (moins de 10 mm).

Validation quantitative

Le Tableau 3.1 récapitule les caractéristiques des données après prétraitement.

Paramètre	Valeur
Taille des images	224×224 pixels
Intervalle d'intensités (images)	$[0, 1]$
Valeurs des masques	$\{0, 1\}$
Format des tenseurs	(C, H, W) avec $C = 3$
Nombre d'images traitées	5 995

TAB. 3.1 : Caractéristiques des données après prétraitement

La description détaillée de la base de données est disponible à la section 2.3.1.

En conclusion, le pipeline de prétraitement mis en place a permis de standardiser efficacement les données, garantissant des entrées de qualité pour les modèles de segmentation et de classification des tumeurs cérébrales. Les résultats quantitatifs de segmentation sont présentés dans la section suivante.

3.2.1.2 *Résultats de segmentation 2D*

Une comparaison a été réalisée entre quatre architectures de segmentation largement utilisées dans l'état de l'art de l'imagerie médicale : UNet, UNet++, FPN et DeepLabV3+.

L'objectif principal est d'évaluer leur capacité à segmenter les tumeurs cérébrales à partir des images IRM de la base BRISC 2025 dans un cadre d'évaluation unifié.

Une étude a été menée afin d'analyser l'influence du nombre d'epochs sur les performances des différentes architectures. Plusieurs configurations ont été testées (10, 20, 25 et 30 epochs), et les résultats ont été évalués à l'aide des métriques Dice Score et IoU, tout en prenant en compte le temps d'entraînement.

Présentation des résultats pour différents nombres d'époques

2***Architecture**	**Dice Score / IoU**			
	10 epochs	**20 epochs**	**25 epochs**	**30 epochs**
UNet	0.8318 / 0.72	0.8373 / 0.73	0.84 / 0.74	0.84 / 0.74
UNet++	0.8271 / 0.72	0.8500 / 0.74	0.85 / 0.75	0.84 / 0.74
FPN	0.8405 / 0.7341	0.8377 / 0.7564	0.8474 / 0.7718	0.8430 / 0.7645
DeepLabV3+	0.8302 / 0.72	0.8427 / 0.73	0.84 / 0.74	0.8500 / 0.74

TAB. 3.2 : Performances des architectures de segmentation en fonction du nombre d'époques

Après 10 epochs, les résultats montrent des performances relativement proches entre les différents modèles. L'architecture UNet atteint un Dice Score de 0.8318 avec un temps d'entraînement d'environ 11.41 minutes. L'architecture UNet++ obtient un Dice Score légèrement inférieur (0.8271), mais avec un temps d'apprentissage beaucoup plus élevé en raison de sa structure dense et de sa complexité architecturale.

L'architecture FPN présente les meilleures performances globales avec un Dice Score de 0.8405 et un IoU de 0.7341, tout en conservant le temps d'entraînement le plus faible parmi les architectures

évaluées. L'architecture DeepLabV3+ obtient également des résultats satisfaisants avec un Dice Score de 0.8302, proches de celles du modèle UNet.

Ces premiers résultats indiquent que l'architecture FPN offre un bon compromis entre précision de segmentation et coût computationnel.

Après 20 epochs, une amélioration générale des performances est observée. Le modèle UNet atteint un Dice Score de 0.8373, tandis que UNet++ progresse jusqu'à 0.8500. L'architecture DeepLabV3+ améliore également ses performances avec un Dice Score de 0.8427. L'architecture FPN conserve des résultats stables avec un Dice Score de 0.8377 et un IoU de 0.7564, confirmant sa robustesse durant l'apprentissage.

L'augmentation du nombre d'epochs jusqu'à 25 permet encore une légère amélioration pour certaines architectures. Le modèle FPN atteint un Dice Score de 0.8474 avec un IoU de 0.7718, tandis que UNet++ conserve des performances élevées. Ces résultats indiquent que certaines architectures bénéficient d'un entraînement plus long pour améliorer la précision des contours segmentés.

Cependant, après 30 epochs, les résultats montrent que l'amélioration des performances n'est pas systématique. Certaines architectures commencent à stagner, tandis que d'autres présentent une légère baisse des métriques de segmentation. L'architecture UNet++ voit notamment son Dice diminuer par rapport aux résultats obtenus à 20 et 25 epochs, ce qui peut indiquer un début de surapprentissage (overfitting). À l'inverse, l'architecture DeepLabV3+ continue d'améliorer légèrement ses performances et atteint un Dice Score de 0.8500. L'architecture FPN conserve quant à lui des performances stables avec un Dice Score de 0.8430 et un IoU de 0.7645, sans variation importante malgré l'augmentation du nombre d'epochs.

Ces observations montrent que l'augmentation du nombre d'epochs n'entraîne pas systématiquement une amélioration des performances. Au-delà d'un certain seuil, certaines architectures atteignent une phase de saturation où les gains deviennent faibles, voire négatifs. Ce comportement peut être lié au surapprentissage, notamment lorsque le modèle commence à mémoriser les données d'entraînement au détriment de sa capacité de généralisation.

Par ailleurs, les différences observées entre les architectures s'expliquent également par leurs caractéristiques internes. L'architecture UNet++ possède une structure dense favorisant la fusion multi-niveaux des caractéristiques, mais cette complexité augmente fortement le temps d'apprentissage. L'architecture DeepLabV3+ exploite les convolutions dilatées pour améliorer la perception contextuelle, tandis que l'architecture FPN repose sur une stratégie pyramidale multi-échelles permettant une bonne détection des structures tumorales de tailles variables.

Parmi les architectures évaluées, le modèle FPN présente le meilleur compromis entre stabilité des performances, précision de segmentation et temps d'apprentissage. Ces résultats justifient ainsi son choix pour les étapes suivantes d'analyse et d'optimisation du système proposé.

Analyse détaillée et optimisation du modèle FPN

Le modèle FPN fournit en sortie une carte de probabilités représentant la présence des régions tumorales. Afin d'obtenir un masque binaire final, un seuil de binarisation est appliqué. Ce seuil joue un rôle critique dans la qualité de la segmentation, car il influence directement le compromis entre détection des lésions et suppression du bruit.

Ainsi, une étude de sensibilité a été réalisée afin de déterminer le seuil optimal parmi les valeurs suivantes : 0.4, 0.5, 0.6 et 0.7.

1) Optimisation du seuil de binarisation

2***Architecture**	**Dice Score / IoU**			
	Seuil = 0.4	**Seuil = 0.5**	**Seuil = 0.6**	**Seuil = 0.7**
UNet	0.8220 / 0.7408	0.8323 / 0.7573	0.8180 / 0.7404	0.8151 / 0.7363
UNet++	0.8116 / 0.7372	0.8285 / 0.7525	0.8162 / 0.7448	0.8279 / 0.7351
FPN	0.8331 / 0.7520	0.8316 / 0.7504	0.8310 / 0.7519	0.8097 / 0.7313
DeepLabV3+	0.8288 / 0.7455	0.8155 / 0.7317	0.8165 / 0.7386	0.8209 / 0.7383

TAB. 3.3 : Performance des architectures en fonction du seuil de binarisation

Seuil = 0.4

Le seuil 0.4 permet une détection large des régions tumorales, ce qui améliore la sensibilité du modèle. Cependant, cette sensibilité élevée entraîne également la détection de zones non tumorales (faux positifs), ce qui peut limiter la précision réelle de segmentation.

Seuil = 0.5

À ce seuil, on observe un meilleur équilibre entre détection des tumeurs et réduction du bruit. Les performances restent stables pour toutes les architectures.

Cependant, par rapport au seuil 0.4, une légère variation des métriques est observée, ce qui s'explique par une réduction des pixels considérés comme tumoraux, supprimant certaines régions périphériques faibles.

Seuil = 0.6

À ce seuil plus élevé, le modèle devient plus strict dans la classification des pixels tumoraux. Cela entraîne une diminution du Dice Score pour certaines architectures, due à la suppression de zones tumorales faibles ou mal contrastées. Ce comportement traduit une baisse du rappel (recall), où certaines régions pathologiques ne sont plus détectées.

Seuil = 0.7

Le seuil 0.7 entraîne une diminution globale des performances, notamment de FPN. Cela s'explique par une sur-sévérité du seuil, qui élimine des régions tumorales à faible probabilité. Ainsi, le modèle privilégie la précision mais perd en sensibilité, ce qui dégrade le Dice Score global.

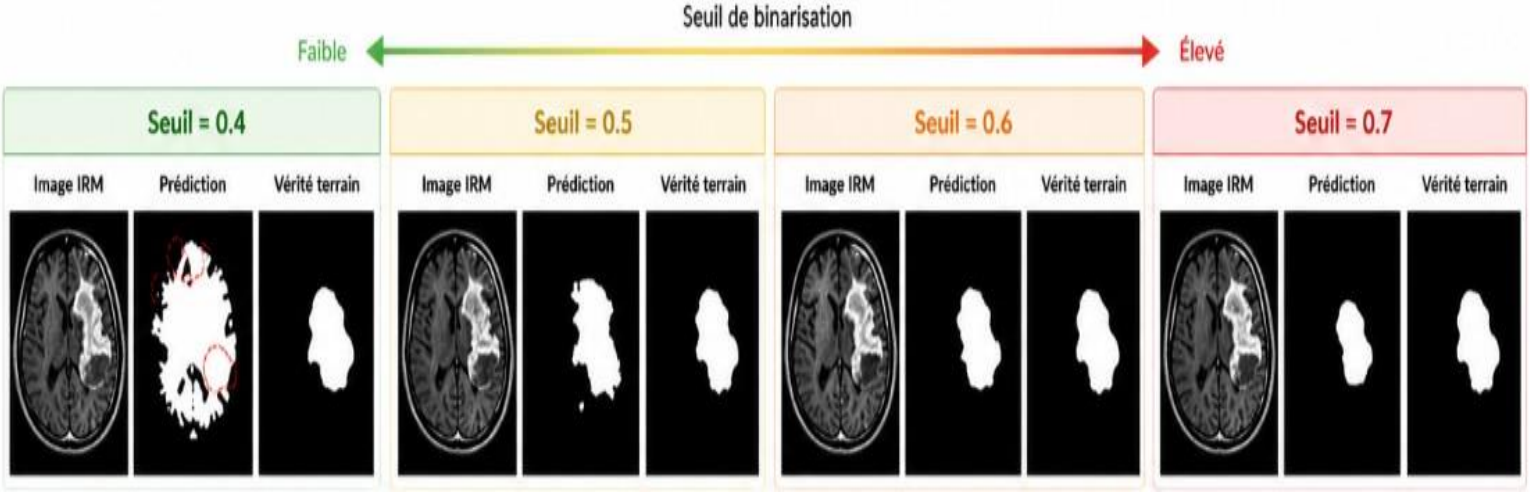


Figure 3.2 : Résumé de l'étude de l'impact du seuil de binarisation sur la segmentation des tumeurs

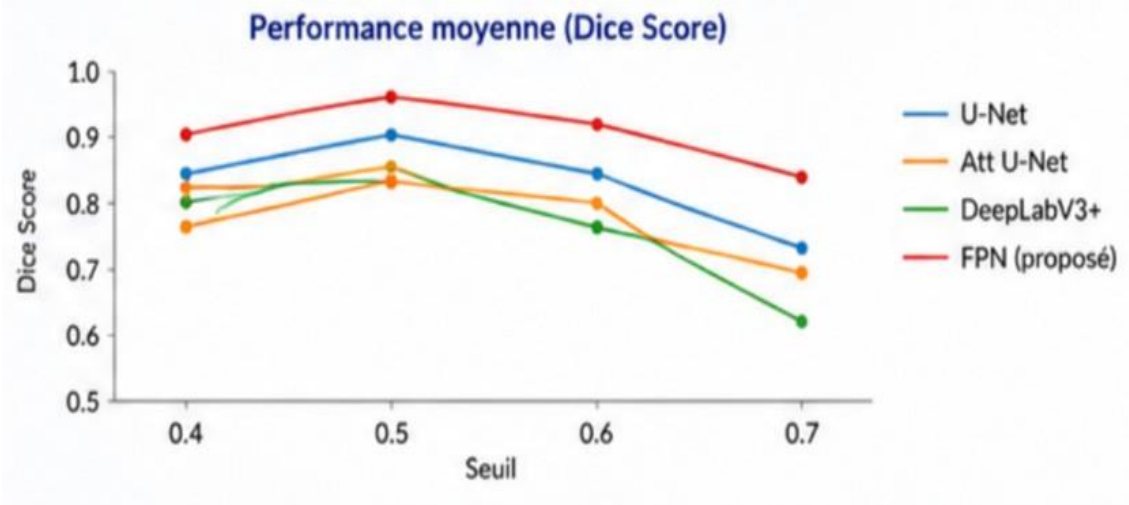


Figure 3.3 : Performance moyenne (Dice Score)

De manière générale, les résultats obtenus montrent que les seuils faibles, tels que 0.4, favorisent une forte sensibilité du modèle mais introduisent un niveau de bruit plus élevé dans les prédictions, en raison de la détection de régions non pertinentes. À l'inverse, les seuils élevés, notamment 0.7, améliorent légèrement la précision mais entraînent une perte d'information en supprimant certaines régions tumorales à faible probabilité, ce qui dégrade la qualité globale de la segmentation. Les seuils intermédiaires, compris entre 0.5 et 0.6, permettent d'obtenir un meilleur équilibre entre sensibilité et précision, en maintenant une détection cohérente des structures tumorales tout en limitant les faux positifs **(figure 3.2).**

Dans ce contexte, le modèle FPN se distingue par la stabilité de ses performances sur l'ensemble des seuils testés, ce qui confirme sa robustesse face à la variation du paramètre de binarisation et sa capacité à maintenir des résultats cohérents indépendamment du seuil choisi.

Par ailleurs, le seuil de 0.5 a été retenu comme compromis optimal, car il permet d'atteindre un équilibre satisfaisant entre sensibilité et précision. Ce choix permet de conserver les régions tumorales pertinentes tout en limitant l'introduction de faux positifs, tout en garantissant une stabilité des performances sur l'ensemble des architectures évaluées. Il présente également des valeurs cohérentes en termes de Dice Score et d'IoU, ce qui justifie son adoption pour la suite du pipeline expérimental.

Ainsi, le seuil de 0.5 permet d'obtenir une segmentation à la fois équilibrée, stable et robuste pour le modèle FPN ainsi que pour les autres architectures étudiées, et est donc retenu comme paramètre optimal dans ce travail **(figure 3.3)**.

2) Optimisation du modèle FPN (hyperparamètres)

Après avoir fixé le seuil de binarisation à 0.5, une phase d'optimisation plus approfondie du modèle FPN a été menée. Cette étape vise à analyser l'influence conjointe du batch size et du nombre d'époques sur les performances de segmentation des tumeurs cérébrales.

L'objectif est d'identifier la configuration permettant d'obtenir le meilleur compromis entre précision de segmentation, stabilité de l'apprentissage et capacité de généralisation du modèle.

Le tableau 3.4 résume les performances obtenues pour différentes configurations expérimentales du modèle FPN, en faisant varier le batch size et le nombre d'époques. Les résultats sont évalués à l'aide du Dice Score, métrique principale utilisée pour mesurer la qualité de segmentation.

Batch size	**Époques**	**Dice Score**
2	10 - 30	[0.8167 - 0.8415]
4	10 - 30	[0.8134 - 0.8533]
6	25 - 60	[0.8719 - 0.8843] (pic 0.8871)
8	25 - 100	[0.8743 - 0.8895]

TAB. 3.4 : Optimisation du modèle FPN : influence du batch size et du nombre d'époques sur le Dice Score

L'analyse des résultats montre une évolution progressive des performances en fonction des paramètres d'entraînement.

- Effet du batch size

Les batch sizes faibles (2 et 4) présentent des performances globalement plus limitées, avec des Dice Scores compris entre 0.81 et 0.85. Cela s'explique par une estimation du gradient plus bruitée, entraînant une convergence moins stable du modèle.

En revanche, l'augmentation du batch size à 6 puis 8 améliore significativement les performances, atteignant des valeurs supérieures à 0.88. Cette amélioration est liée à une meilleure stabilité des mises à jour du gradient, permettant une optimisation plus efficace du modèle.

- Effet du nombre d'époques

L'augmentation du nombre d'époques améliore les performances jusqu'à un certain seuil. Par exemple, avec un batch size de 6, le Dice Score augmente jusqu'à 0.8871 à 30 époques, puis se stabilise autour de 0.8817–0.8843. Cette stagnation s'explique par une saturation du modèle, où l'apprentissage n'apporte plus d'amélioration significative.

- Batch size 8

Le batch size 8 présente la meilleure dynamique d'apprentissage. Les performances augmentent progressivement de 0.8743 à 25 époques jusqu'à atteindre un maximum de 0.8895 pour environ 100 époques.

Cette configuration montre une meilleure capacité de généralisation et une convergence plus stable par rapport aux autres paramètres.

De manière globale, les résultats montrent que :

- Les batch sizes faibles limitent la performance du modèle
- Les batch sizes élevés améliorent la stabilité et la convergence
- L'augmentation des époques est bénéfique uniquement jusqu'à un seuil
- Au-delà, le modèle atteint une zone de saturation

Le meilleur compromis est obtenu avec un batch size de 8 et environ 100 époques, donnant un Dice Score maximal de 0.8895, ce qui représente la meilleure configuration du modèle FPN dans cette étude.

Cette phase d'optimisation a permis d'identifier les paramètres les plus adaptés au modèle FPN. Les résultats confirment que la performance du modèle dépend fortement du choix du batch size et du nombre d'époques, et qu'un réglage approprié permet d'améliorer significativement la qualité de segmentation des tumeurs cérébrales.

Afin de compléter l'évaluation quantitative du modèle FPN, une analyse qualitative des résultats de segmentation a également été réalisée. Les figures suivantes présentent des exemples de prédictions obtenues par le modèle, en comparant l'image IRM originale, le masque de vérité terrain (Ground Truth) et le masque prédit **(figure 3.4).**

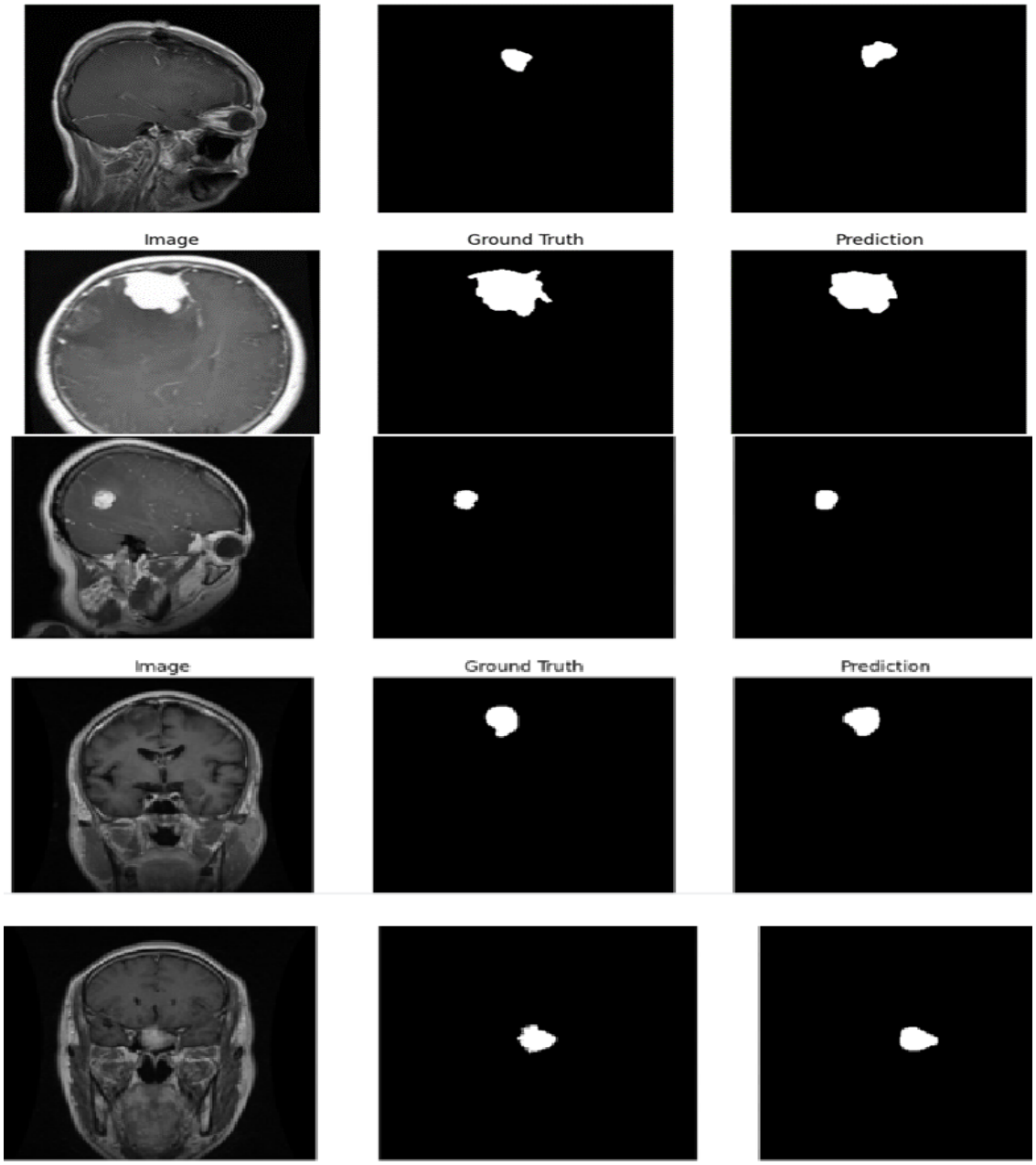


Figure 3.4 : Exemple de segmentation obtenue avec le modèle FPN (Image IRM – Ground Truth – Prediction)

Les résultats visuels montrent une bonne correspondance entre les régions tumorales prédites et les masques de référence. Le modèle FPN parvient à détecter correctement les principales structures tumorales tout en conservant des contours relativement précis.

Afin d'analyser le comportement du modèle durant l'apprentissage, les courbes d'évolution du Dice Score ont également été étudiées au cours des différentes époques d'entraînement **(figure 3.5).** La courbe montre une augmentation progressive du Dice Score au cours des premières époques, traduisant une amélioration continue des performances du modèle. Une phase de stabilisation est ensuite observée, indiquant la convergence progressive du processus d'apprentissage.

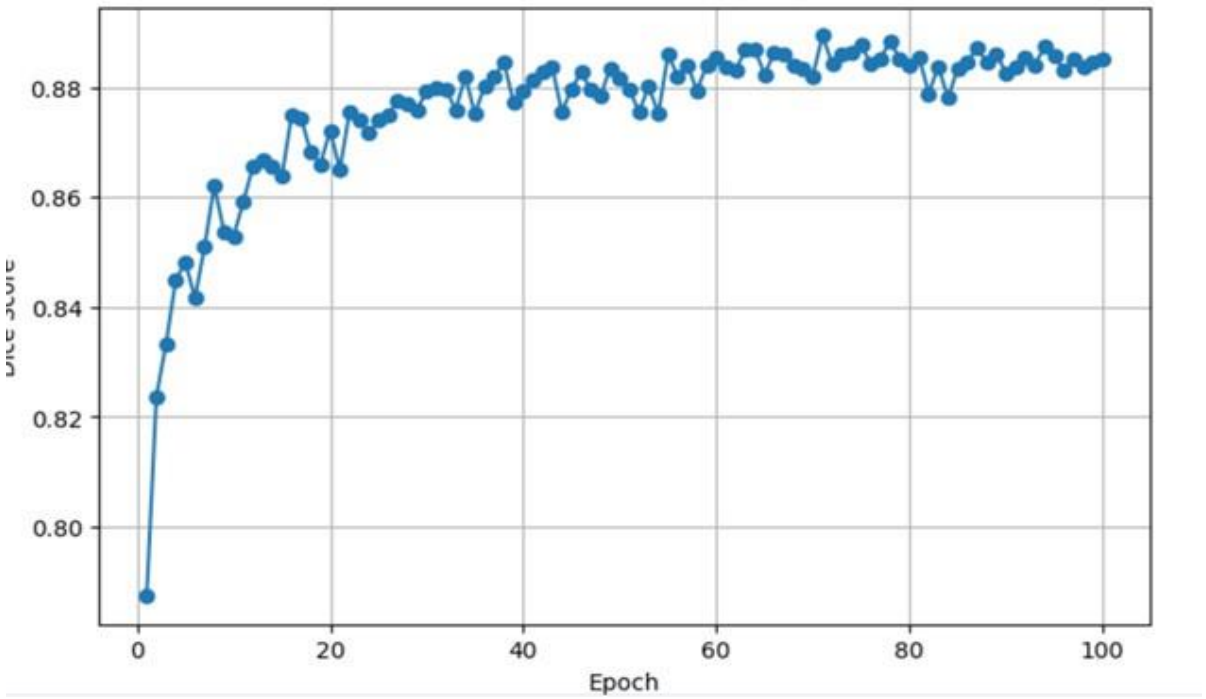


Figure 3.5 : Courbe d'apprentissage du modèle FPN pour un batch size de 8

3.2.2 *Visualisation anatomique et représentation enrichie des résultats*

Cette section présente l'évaluation du modèle FPN sur le jeu de test, à travers une analyse quantitative basée sur le Dice Score ainsi qu'une analyse qualitative des cartes de segmentation obtenues.

L'objectif principal de cette phase est de fournir une représentation enrichie des tumeurs cérébrales en combinant les résultats de segmentation avec des informations anatomiques issues d'un atlas cérébral de référence, tout en intégrant une décomposition morphologique des lésions.

3.2.2.1 *Analyse globale des résultats de segmentation*

L'ensemble des opérations réalisées dans cette section permet de construire une représentation complète et interprétable des résultats de segmentation. En combinant la prédiction du modèle, la décomposition morphologique et l'atlas cérébral, le système fournit une visualisation enrichie des tumeurs cérébrales.

Cette approche améliore significativement :

- L'interprétation qualitative des résultats.
- La compréhension des structures tumorales.
- La liaison entre intelligence artificielle et anatomie cérébrale.

Elle constitue également une étape intermédiaire importante vers la reconstruction et la réalité augmentée tridimensionnelle abordée dans la suite du travail.

La figure suivante **(Figure 3.6)** illustre un exemple complet de visualisation des résultats de segmentation. Elle regroupe l'image IRM originale, le masque de vérité terrain (GT), la prédiction du modèle, la superposition de la segmentation sur l'image IRM, la décomposition de la tumeur en trois régions (centre, bord et œdème), ainsi que la fusion entre la segmentation et l'atlas cérébral de référence.

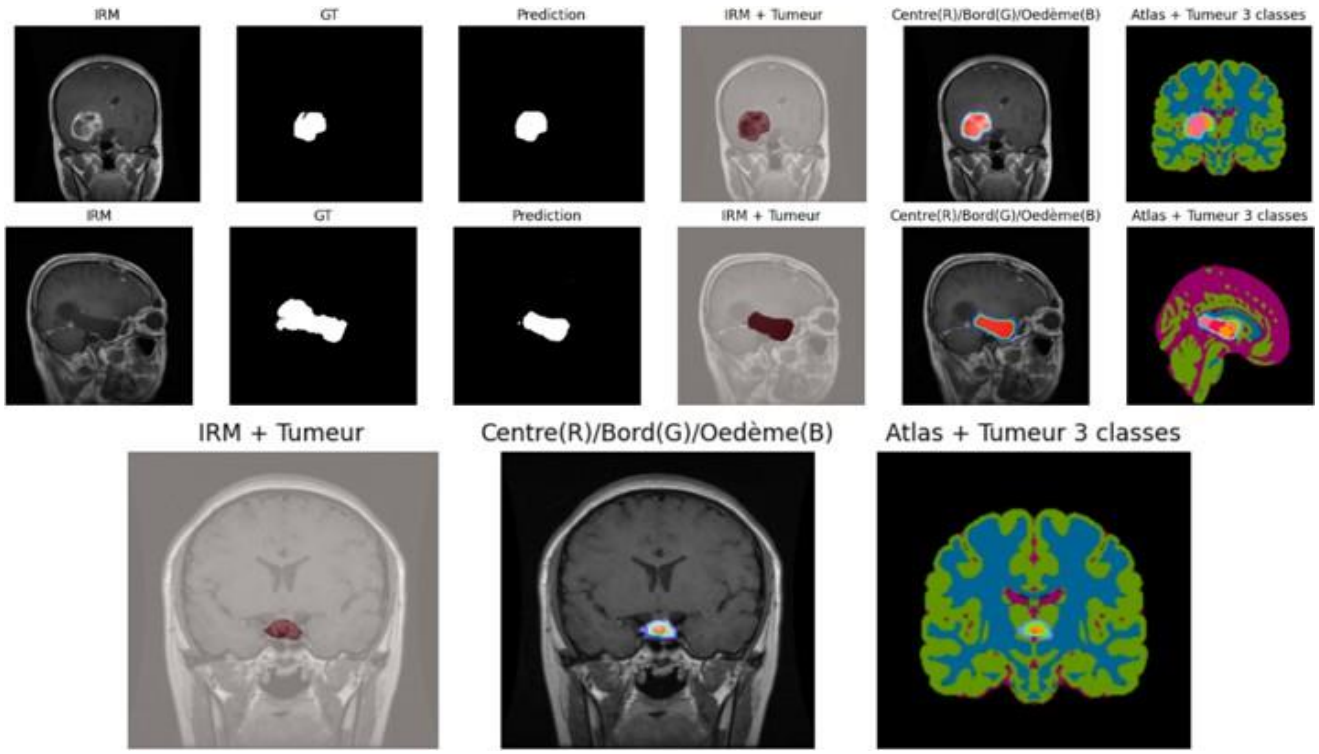


Figure 3.6 : Visualisation complète des résultats de segmentation : IRM, GT, prédiction, superposition, décomposition tumorale et fusion avec atlas cérébral

Afin d'évaluer la stabilité des performances du modèle, la figure suivante **(Figure 3.7**) présente la courbe du Dice Score calculé image par image sur l'ensemble du jeu de test. Cette représentation permet d'analyser la variabilité des performances selon les cas cliniques.

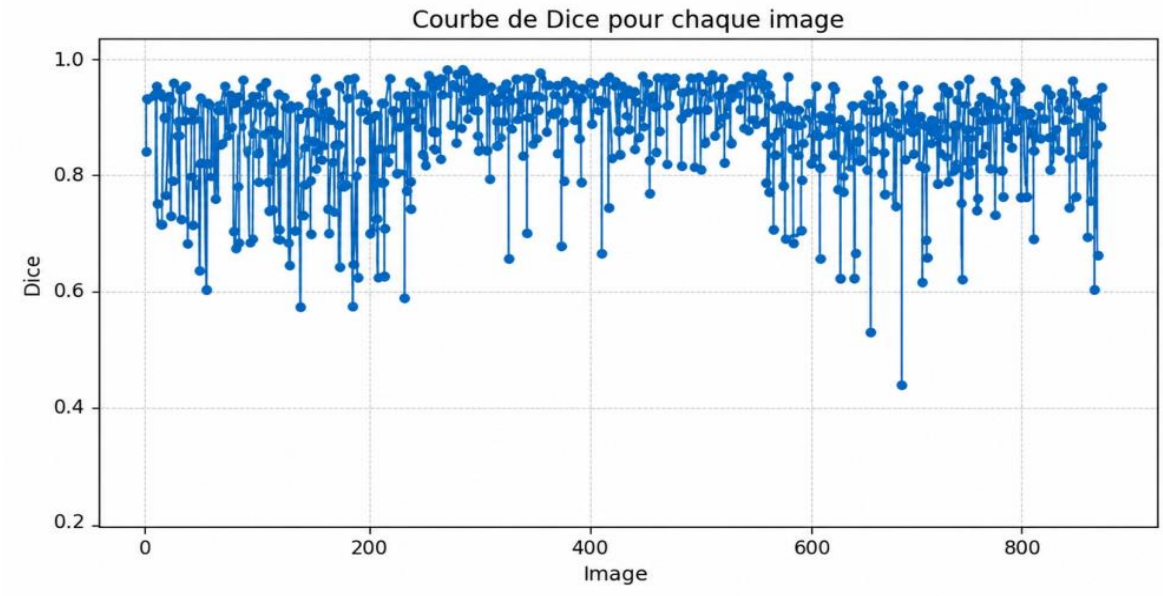


Figure 3.7 : Courbe du Dice Score par image sur le jeu de test

Les résultats obtenus confirment que le modèle FPN produit des segmentations globalement cohérentes avec les annotations de référence, tout en maintenant une performance stable sur l'ensemble des images du jeu de test.

Les résultats obtenus dans la partie précédente ont permis de valider l'efficacité du modèle FPN pour la segmentation des tumeurs cérébrales en 2D, ainsi que l'intérêt de la visualisation enrichie intégrant les structures anatomiques issues des atlas. Toutefois, cette approche reste limitée par le traitement coupe par coupe (slice-by-slice), ce qui ne permet pas de représenter entièrement la continuité spatiale des structures tumorales dans un volume cérébral complet.

Dans ce contexte, la suite du travail vise à dépasser ces limitations en introduisant une approche de reconstruction tridimensionnelle et de réalité augmentée 3D, permettant d'exploiter les informations volumétriques des images IRM.

3.2.3 *Réalité augmentée 3D*

3.2.3.1 *Résultats du prétraitement automatique par nnU-Net*

La Figure 3.8 illustre l'ensemble des transformations appliquées automatiquement par nnU-Net à une image IRM de la modalité FLAIR issue du dataset BraTS2020.

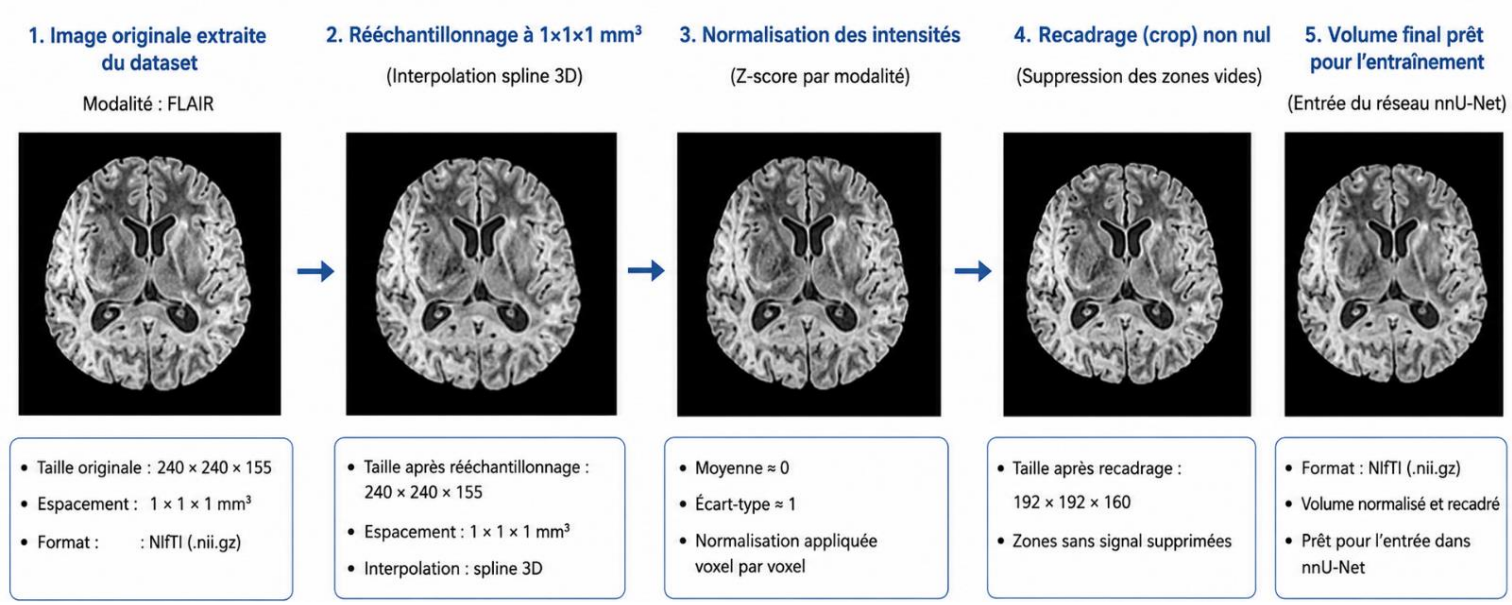


Figure 3.8 : Prétraitement automatique avec nnU Net sur une image

Étapes du pipeline de transformation

Le framework nnU-Net a appliqué séquentiellement quatre opérations :

1. **Rééchantillonnage** : interpolation spline 3D pour homogénéiser la résolution spatiale à 1 × 1 × 1 mmS.
2. **Normalisation des intensités** : standardisation par Z-score appliquée indépendamment par modalité (moyenne = 0, écart-type = 1).
3. **Recadrage non nul (cropping)** : suppression automatique des zones périphériques sans signal.
4. **Conversion au format interne** : transformation vers le format compressé .b2nd (nnU-Net).

Observations sur la méthodologie développée dans ce travail

Le recadrage non nul a permis d'éliminer les zones sans signal, réduisant la taille du volume de 240×240×155 à 192×192×160 voxels, ce qui a diminué la charge computationnelle sans perte d'information anatomique. La normalisation Z-score a harmonisé les intensités voxel par voxel, facilitant l'apprentissage du modèle.

3.2.3.2 *Segmentation des tumeurs cérébrales*

Entraînement du modèle de segmentation

L'apprentissage du modèle repose sur une approche supervisée utilisant les volumes IRM annotés du dataset BraTS2020. Durant cette phase, le réseau apprend progressivement à associer les structures anatomiques présentes dans les images aux différentes classes tumorales définies dans les masques de vérité terrain.

L'entraînement s'effectue de manière itérative sur plusieurs époques. À chaque étape, le modèle traite des mini-lots de données (batches), ce qui permet d'optimiser l'utilisation de la mémoire GPU et d'accélérer le processus d'apprentissage. Les performances sont évaluées à l'aide du coefficient de Dice, une métrique standard en segmentation médicale permettant de mesurer la similarité entre les prédictions et les annotations réelles. Une moyenne mobile du Dice est également utilisée afin de stabiliser l'évaluation et de réduire les fluctuations liées aux batches.

Tout au long du processus d'apprentissage, le framework nnU-Net affiche des métriques détaillées permettant de suivre la convergence du modèle, notamment la loss, le learning rate, ainsi que le pseudo-Dice Score pour chaque classe tumorale (Annexe F).

Ces logs montrent que le modèle améliore progressivement ses performances au fil des époques, avec une diminution de la loss et une stabilisation du pseudo-Dice, ce qui indique une bonne convergence du réseau.

Validation croisée (K-Fold Cross-Validation)

Afin d'assurer une évaluation robuste, nnU-Net utilise une validation croisée de type K-Fold. Dans cette approche, le dataset est divisé en plusieurs sous-ensembles appelés folds. À chaque itération, une partie des données est utilisée pour la validation tandis que le reste sert à l'apprentissage du modèle.

Cette méthode permet d'évaluer les performances sur différentes répartitions des données et de réduire le risque de surapprentissage. Elle améliore également la capacité de généralisation du modèle en garantissant que les résultats ne dépendent pas d'un sous-ensemble spécifique de patients.

Afin d'évaluer la robustesse du modèle, nnU-Net utilise une validation croisée de type K-Fold. Cette approche génère automatiquement une structure organisée contenant les résultats de chaque fold ainsi que les configurations utilisées pour l'entraînement (Figure 3.9).

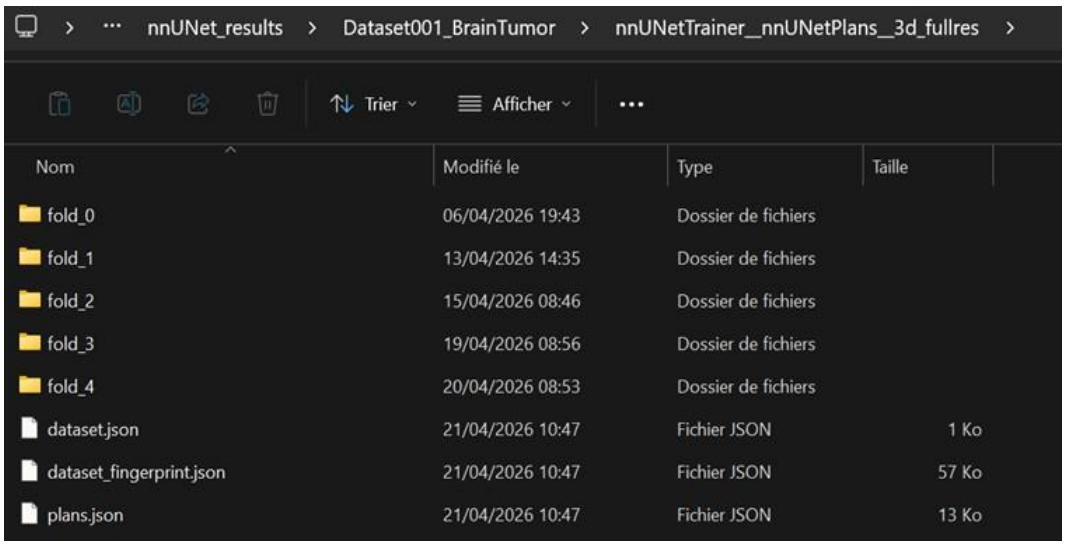


Figure 3.9 : Structure des résultats de validation croisée (nnU-Net – K-Fold)

Cette organisation permet d'évaluer le modèle sur plusieurs partitions du dataset, garantissant ainsi une meilleure généralisation et une réduction du risque de surapprentissage.

Pour chaque fold de validation croisée, nnU-Net génère automatiquement un ensemble de fichiers permettant le suivi complet de l'entraînement et de l'évaluation du modèle. Ces fichiers incluent notamment les checkpoints sauvegardés, les journaux d'apprentissage ainsi que les courbes d'évolution des performances (Figure 3.19).

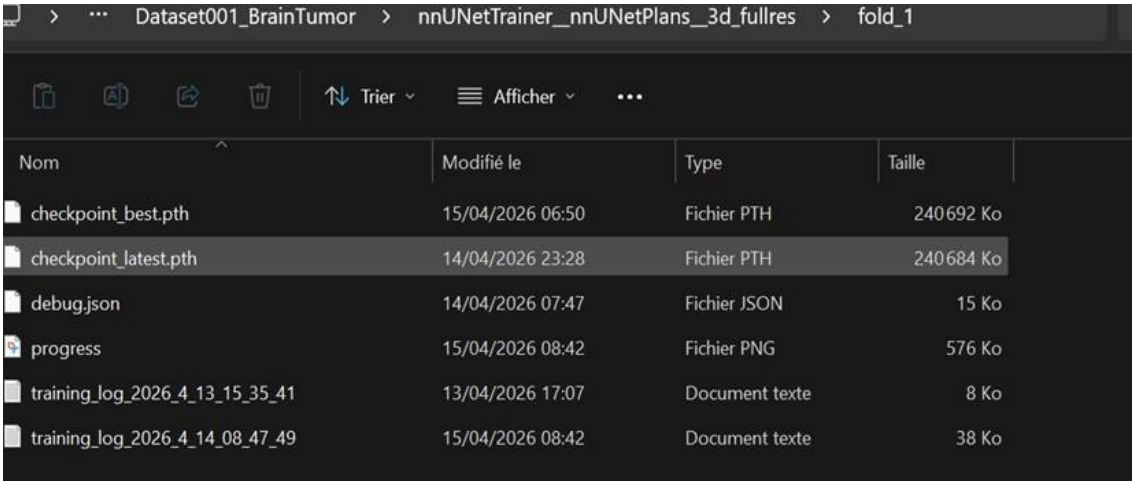


Figure 3.10 : Contenu du dossier fold_1 généré après l'entraînement du modèle nnU-Net

Les fichiers checkpoint_best.pth et checkpoint_latest.pth correspondent respectivement au meilleur modèle sauvegardé durant l'entraînement et au dernier état du modèle. Le fichier progress.png contient les courbes d'évolution des performances, tandis que les fichiers training_log enregistrent les informations détaillées liées au processus d'apprentissage.

Sélection automatique du meilleur modèle

Au cours de l'entraînement, nnU-Net surveille en continu les performances du modèle sur les données de validation. Lorsqu'une amélioration est détectée, le framework sauvegarde automatiquement l'état du modèle sous forme de checkpoint.

Le modèle final retenu correspond à celui présentant les meilleures performances en termes de

Dice Score sur la validation. Cette stratégie permet une sélection automatique, fiable et reproductible du modèle optimal sans intervention manuelle.

Afin de sélectionner la configuration la plus performante pour la segmentation des tumeurs cérébrales, une comparaison quantitative a été réalisée entre les architectures nnU-Net 2D et nnU-Net 3D full resolution à partir des Dice Scores obtenus sur les différents folds de validation croisée (Figure 3.11).

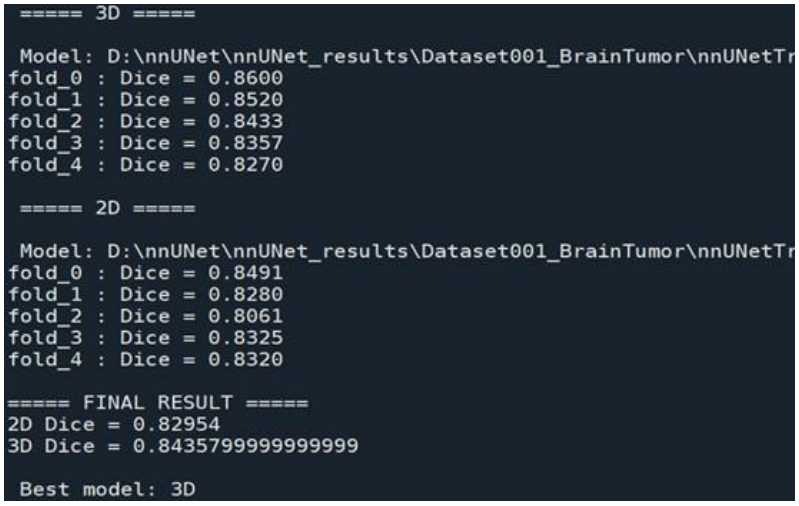


Figure 3.11 : Comparaison des Dice Scores obtenus avec les configurations nnU-Net 2D et 3D

Les résultats montrent que la configuration nnU-Net 3D full resolution obtient les meilleures performances globales avec un Dice moyen de 0.8436, contre 0.8295 pour la configuration 2D.

Cette différence s'explique principalement par la capacité du modèle 3D à exploiter les relations spatiales entre les différentes coupes IRM volumétriques. Contrairement à l'approche 2D, qui traite chaque coupe indépendamment, la segmentation 3D permet une meilleure prise en compte de la continuité anatomique des structures tumorales.

L'architecture 3D offre ainsi une représentation plus fidèle des volumes tumoraux complexes, ce qui améliore la précision de segmentation, notamment au niveau des régions actives et des noyaux tumoraux.

En raison de ces performances supérieures, la configuration nnU-Net 3D full resolution a été retenue comme modèle final pour les étapes de prédiction, de reconstruction volumétrique et d'intégration en réalité augmentée.

Génération des prédictions volumétriques

Une fois l'entraînement terminé, le modèle est utilisé en phase d'inférence pour générer des prédictions sur de nouveaux volumes IRM. Les résultats obtenus sont des masques de segmentation tridimensionnels, dans lesquels chaque voxel est associé à une classe anatomique ou pathologique.

Ces volumes segmentés sont sauvegardés au format NIfTI (.nii.gz), ce qui permet leur exploitation dans les étapes de visualisation 3D, de reconstruction volumétrique et d'applications en réalité augmentée médicale.

Afin d'évaluer qualitativement les performances du modèle, une visualisation des résultats de segmentation a été réalisée sur un exemple du dataset BraTS2020. La Figure 3.12 illustre la comparaison entre l'image IRM (modalité FLAIR), le masque de vérité terrain (GT), la prédiction du modèle nnU-Net, ainsi que leur superposition permettant d'analyser la concordance spatiale entre les régions tumorales réelles et prédites.

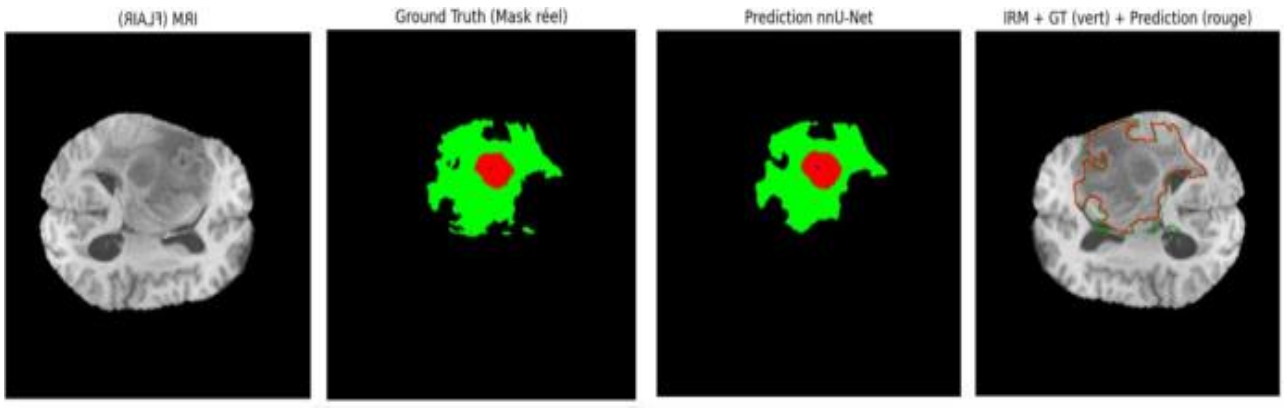


Figure 3.12 : Visualisation qualitative des résultats de segmentation sur un cas BraTS2020

Cette visualisation montre une bonne cohérence globale entre les annotations de référence et les prédictions du modèle, notamment au niveau des régions tumorales principales.

Analyse des performances du modèle

L'évaluation des performances repose principalement sur le Dice Score. Les résultats montrent une évolution stable et progressive durant l'entraînement, traduisant une bonne capacité du modèle à apprendre les structures tumorales.

Les segmentations obtenues sont globalement cohérentes pour les différentes classes tumorales, notamment les régions d'œdème, les noyaux tumoraux et les zones actives.

nnU-Net génère automatiquement des courbes de suivi incluant le Pseudo Dice, la loss, le learning rate ainsi que le temps d'exécution pour chaque époque d'entraînement (Figure 3.13).

Définition du Pseudo Dice : Le Pseudo Dice est une estimation du coefficient de Dice calculée sur chaque mini-lot (batch) pendant l'entraînement, contrairement au Dice final calculé sur l'ensemble de la validation. Il permet un suivi plus fréquent des performances. Les valeurs détaillées et les courbes d'évolution des Pseudo Dice pour les sous-classes tumorales sont présentées en Annexe F . La configuration détaillée des paramètres d'entraînement **(Figure 53)** est également disponible en **Annexe F.**

L'entraînement du modèle nnU-Net v2 a été réalisé sur le dataset Dataset001-BrainTumor en configuration 2D avec le fold 1 de la validation croisée. Le modèle utilise une architecture Plain-ConvUNet composée de 7 niveaux (stages) avec un nombre de caractéristiques par niveau de [32, 64, 128, 256, 512, 512, 512]. La taille des patches d'entrée est fixée à 256 × 256 pixels, avec une taille médiane des images de 240 × 240 pixels et un espacement de [1.0, 1.0]. La taille du batch est de 49. Les convolutions utilisent des noyaux de taille 3 × 3 et des strides [[1,1], [2,2], [2,2], [2,2], [2,2], [2,2], [2,2]]. La normalisation appliquée aux images est la Z-Score Normalization, tandis que le réseau utilise une couche de normalisation InstanceNorm2d et la fonction d'activation LeakyReLU. Le taux d'apprentissage initial (learning rate) est fixé à 0,01. La base de données est répartie en 294 cas d'entraînement et 74 cas de validation.

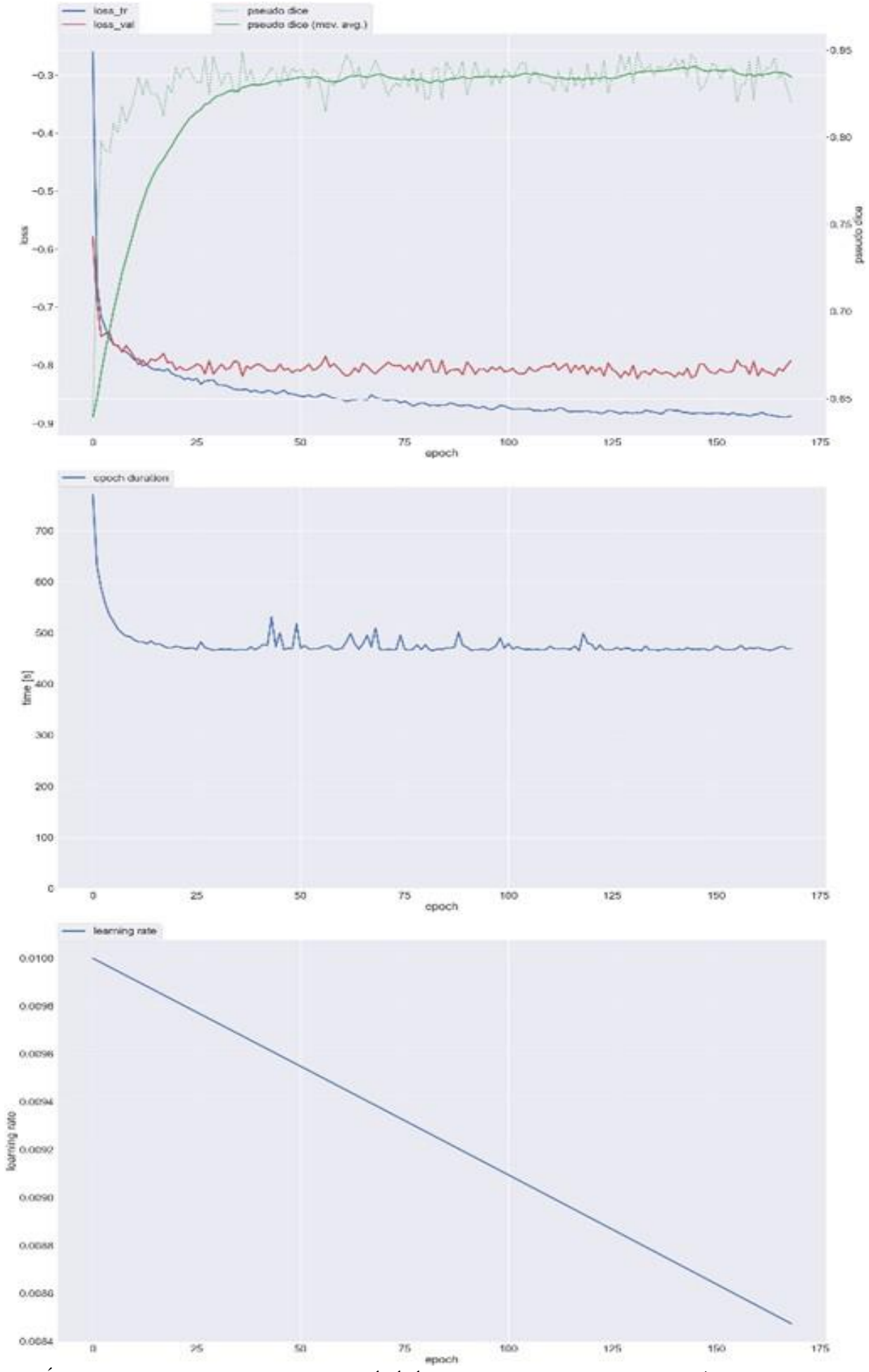


Figure 3.13 : Évolution du Dice Score, de la loss1e1t1du learning rate durant l'entraînement (Fold 0 – configuration 3D full resolution)

L'évolution du Pseudo Dice ainsi que de sa moyenne mobile (EMA Pseudo Dice) met en évidence une convergence progressive du modèle durant l'apprentissage. La stabilisation des courbes vers les dernières époques indique une amélioration de la qualité de segmentation et une bonne stabilité de l'entraînement.

3.2.4 *Post-traitement, analyse quantitative et génération des modèles 3D pour la réalité augmentée*

Après la phase de segmentation réalisée avec nnU-Net v2, un pipeline de post-traitement avancé a été développé afin d'exploiter les résultats obtenus dans un environnement de visualisation tridimensionnelle et de réalité augmentée. Ce pipeline assure plusieurs fonctions complémentaires, notamment le chargement des volumes IRM segmentés, le calcul des métriques quantitatives, la reconstruction 3D des structures anatomiques ainsi que l'exportation des modèles vers l'environnement Unity.

L'objectif principal de cette étape est de transformer les segmentations volumétriques produites par nnU-Net en représentations anatomiques exploitables dans un environnement immersif interactif. Le pipeline permet également d'assurer l'alignement spatial précis des modèles, l'évaluation automatique des performances de segmentation ainsi que la génération de différents supports d'analyse, tels que les superpositions visuelles, les rapports numériques, les QR codes et les signaux audio.

Cette approche établit ainsi une liaison complète entre la segmentation médicale automatique, la reconstruction tridimensionnelle et les applications de réalité augmentée.

Organisation des résultats par patient

Après la phase de segmentation, le pipeline génère automatiquement un sous-dossier dédié à chaque patient traité. Cette organisation permet de centraliser l'ensemble des résultats produits par le système, incluant les données volumétriques, les métriques quantitatives, les modèles tridimensionnels ainsi que les fichiers destinés à l'intégration dans Unity.

Chaque dossier patient contient plusieurs fichiers correspondant aux différentes étapes du pipeline de post-traitement et de reconstruction 3D. Cette structuration facilite l'analyse des résultats, la visualisation anatomique ainsi que l'exploitation des données dans un environnement immersif de réalité augmentée (Figure 3.14).

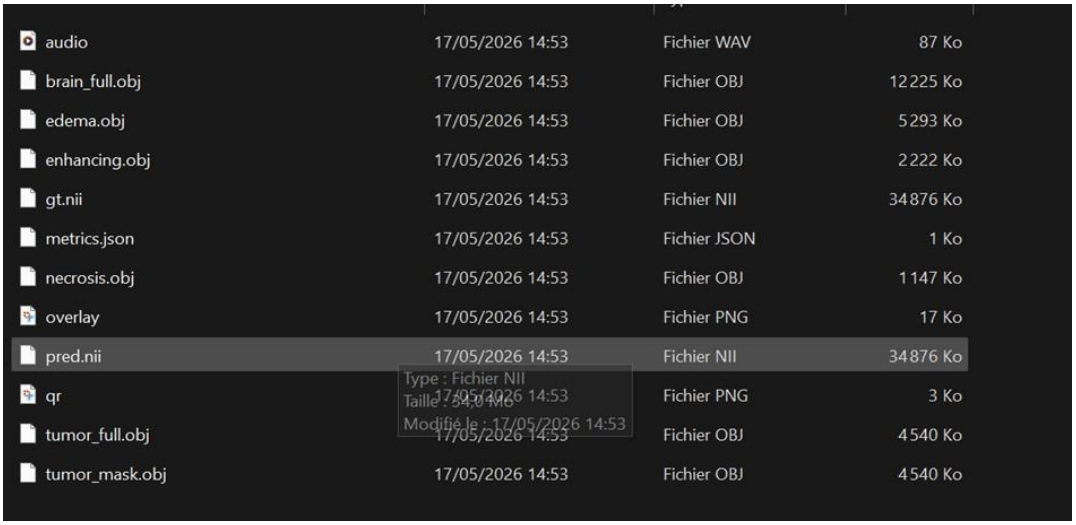


Figure 3.14 : Structure des résultats générés automatiquement pour un patient

Afin de clarifier la structure des résultats générés pour chaque patient, le Tableau 3.5 ci-dessous présente l'ensemble des fichiers produits par le pipeline de post-traitement. Ces fichiers regroupent à la fois les données de segmentation, les reconstructions tridimensionnelles, les visualisations qualitatives ainsi que les indicateurs quantitatifs permettant l'évaluation des performances du modèle.

Fichier	**Description**
pred.nii	segmentation prédite
gt.nii	vérité terrain
brain_full.obj	reconstruction 3D du cerveau
tumor_full.obj	reconstruction complète de la tumeur
edema.obj	région d'œdème
necrosis.obj	région nécrotique
overlay.png	superposition GT/prédiction
metrics.json	métriques quantitatives
qr.png	QR code du rapport
audio.wav	représentation sonore de la tumeur

TAB. 3.5 : Description des fichiers générés par patient

3.2.5 *Visualisation qualitative par superposition*

Afin d'évaluer qualitativement les performances du modèle, une visualisation par superposition a été réalisée sur une coupe axiale du volume IRM. Cette représentation permet de comparer la vérité terrain (ground truth) et la segmentation prédite, mettant en évidence la concordance spatiale entre les régions tumorales réelles et celles détectées automatiquement (Figure 3.14). On observe une bonne concordance globale entre la vérité terrain et la segmentation prédite, ce qui confirme la capacité du modèle à localiser correctement les régions tumorales.

3.2.6 *Évaluation quantitative des performances*

Bien que l'évaluation quantitative ait été réalisée sur l'intégralité des patients, le Tableau 3.6 présente cinq cas représentatifs permettant d'apprécier la diversité des résultats obtenus.



Figure 3.15 : Superposition entre la vérité terrain et la segmentation prédite (overlay)

Patient Volume (mmS)	DICE	IoU	Recall	Spécificité	Accuracy	HD95
000 184318.0	0.91614	0.84525	0.85636	0.999680	0.9962776	1
001 63659.0	0.90425	0.82523	0.88165	0.9994830	0.9985986	1
002 29778.0	0.81715	0.690834	0.816754	0.99938	0.998779	2.23606
003 161598.0	0.89739	0.8138820	0.8447136	0.999206	0.996038	1.414213
004 15262.0	0.78576	0.647123	0.6658926	0.999928	0.9991067	2.449489

TAB. 3.6 : Métriques d'évaluation pour cinq patients du dataset

L'analyse des résultats sur les cinq patients montre une performance **acceptable mais perfectible** du modèle. Les valeurs du Dice varient entre 0,78 et 0,91, avec de meilleures performances pour les patients 000 et 001.

Le patient 004 présente le score le plus bas (0,785), valeur inférieure au seuil de 0,8 généralement considéré comme acceptable en segmentation médicale, **en raison probablement d'une tumeur de petite taille (15 262 mmS).**

L'IoU et le recall confirment la cohérence des prédictions. La spécificité reste proche de 1 pour tous les cas, indiquant une faible présence de faux positifs. Les valeurs HD95 (1 à 2,45 mm) montrent un alignement spatial acceptable.

Ainsi, ce pipeline permet une exploitation complète des résultats de segmentation, de l'évaluation quantitative à la visualisation 3D et la réalité augmentée.

Les Pseudo Dice par sous-classe tumorale (nécrose, œdème, tumeur rehaussante) sont suivis pendant l'entraînement. Pour la configuration 2D, les valeurs observées en fin d'entraînement sont respectivement d'environ 0,80 pour la nécrose, 0,82 pour l'œdème et 0,87 pour la tumeur rehaussante. Pour la configuration 3D, elles atteignent environ 0,85 pour la nécrose, 0,86 pour l'œdème et 0,88

pour la tumeur rehaussante. Les valeurs détaillées et les courbes d'évolution sont présentées en **Annexe F**. Les mesures quantitatives de la reconstruction 3D ont été calculées pour 7 patients . Le Tableau 3.7 présente pour chaque structure anatomique le nombre moyen de sommets, le nombre moyen de polygones et le temps moyen de génération du maillage.

Structure	Sommets	Polygones	Temps (s)
Cerveau complet	113 323	226 626	0,22
Tumeur	19 080	38 183	0,03
Œdème	20 407	40 946	0,04
Nécrose	4 615	9 243	0,02
Tumeur rehaussée	9 798	19 641	0,02

TAB. 3.7 : Mesures quantitatives moyennes de la reconstruction 3D

3.2.7 *Reconstruction tridimensionnelle et visualisation en réalité augmentée*

Dans cette dernière étape du pipeline, les résultats issus de la segmentation automatique sont exploités afin de générer une visualisation immersive des structures cérébrales. Cette phase vise à transformer les données médicales reconstruites en représentations tridimensionnelles interactives, intégrées dans un environnement de réalité augmentée développé sous Unity.

L'objectif principal est de permettre une exploration intuitive des structures anatomiques (cerveau, tumeur et sous-régions pathologiques) dans un espace 3D interactif. Cette approche facilite l'analyse visuelle et améliore la compréhension des résultats de segmentation en offrant une visualisation plus proche du contexte clinique réel.

Afin de garantir l'accessibilité du système, le choix a été orienté vers une application mobile en réalité augmentée, permettant l'exécution sur smartphone sans nécessiter de dispositifs spécialisés tels que des casques ou lunettes AR. Cette stratégie assure une large compatibilité tout en conservant un niveau d'interaction avancé.

3.2.7.1 *Interface utilisateur*

L'interface utilisateur de l'application a été développée sous Unity à l'aide d'un Canvas afin d'assurer une interaction fluide entre l'utilisateur et le système en réalité augmentée. Elle est conçue pour s'adapter automatiquement aux différentes tailles d'écran grâce au mode Scale With Screen Size, garantissant une disposition cohérente des éléments sur les appareils mobiles.

La Figure 3.16 présente l'interface utilisateur en scène, montrant l'ensemble des éléments interactifs (boutons de navigation, sliders de contrôle de transparence, affichage des informations patient) permettant une exploration complète des structures tumorales en réalité augmentée.



Figure 3.16 : Interface utilisateur en scène

3.2.7.2 *Résultat final de visualisation*

La phase de visualisation finale représente l'étape d'intégration globale du système de réalité augmentée. Elle constitue l'aboutissement du pipeline complet du projet, allant du traitement des données médicales et de la segmentation automatique jusqu'à la reconstruction tridimensionnelle et la visualisation interactive dans un environnement immersif.

Dans cette étape, les différentes structures anatomiques reconstruites sont organisées de manière hiérarchique et visualisées avec une différenciation par couleurs afin de faciliter leur interprétation. Ainsi, chaque composante est représentée selon un code couleur spécifique permettant une distinction claire entre les structures principales et les sous-régions tumorales. Cette organisation colorimétrique permet une meilleure lisibilité des structures et facilite l'analyse visuelle des relations spatiales entre les différentes régions cérébrales.

Une première expérimentation a été réalisée dans l'environnement Unity afin de valider l'affichage simultané de ces structures segmentées. Comme illustré dans la figure suivante, le cerveau complet ainsi que la tumeur sont affichés ensemble dans la scène, permettant une visualisation globale des résultats de segmentation dans un environnement contrôlé.

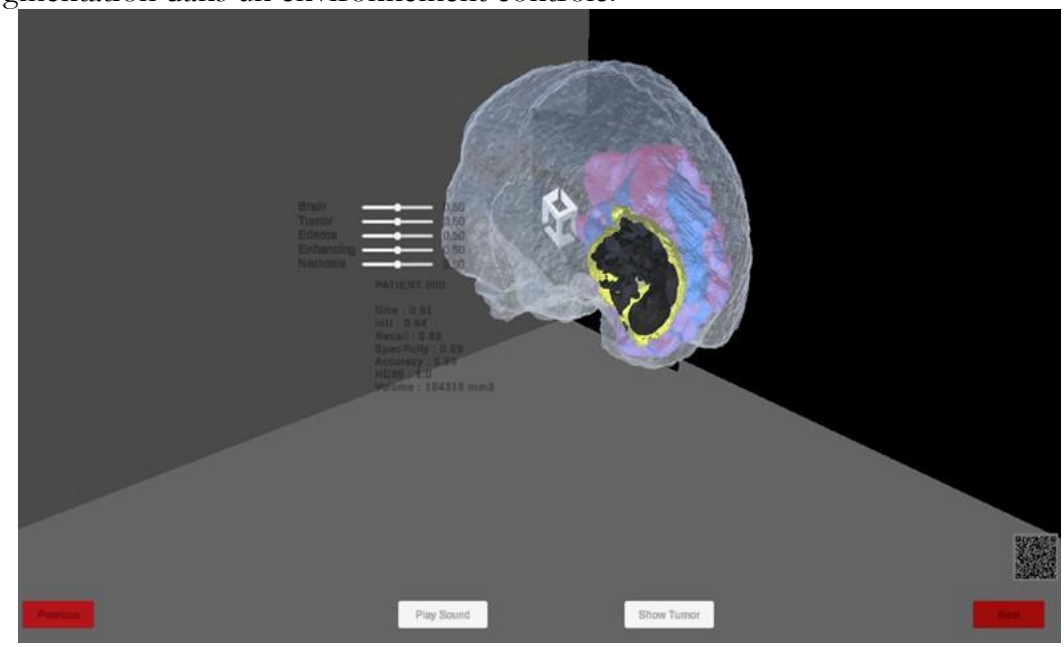


Figure 3.17 : Visualisation du cerveau et de la tumeur dans l'environnement Unity

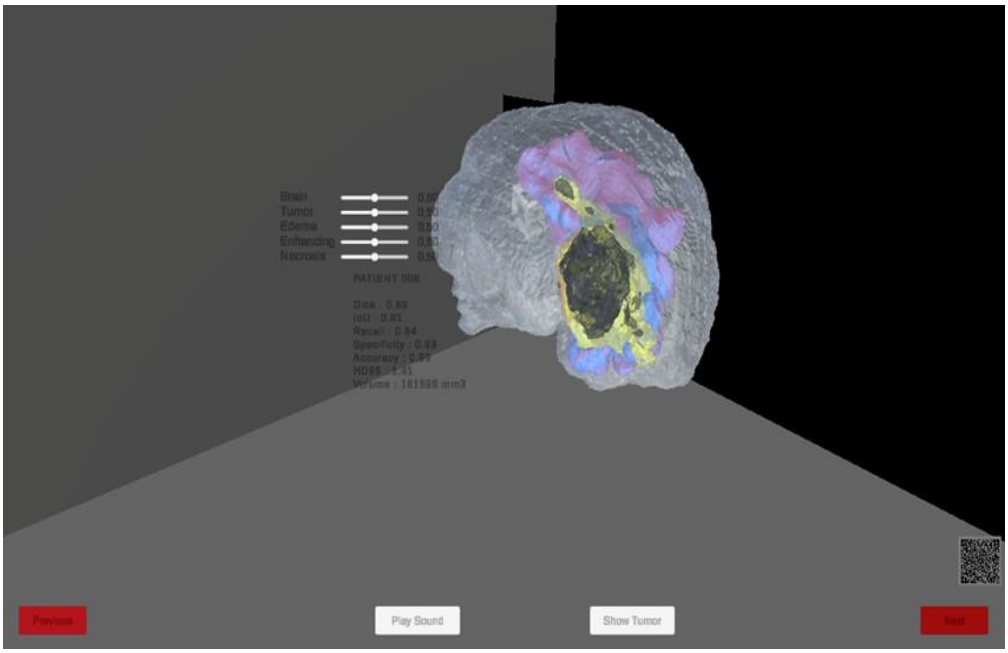


Figure 3.18 : Navigation entre les patients et interaction avec les modèles 3D

Dans cette configuration, les modèles peuvent être manipulés en temps réel. L'utilisateur peut effectuer des opérations de rotation et de zoom afin d'explorer les structures cérébrales sous différents angles. Afin de valider le bon fonctionnement des interactions, une transition vers un autre cas patient a été réalisée à travers les boutons Next et Previous. Comme illustré dans la figure suivante, le passage entre les patients met automatiquement à jour les modèles affichés ainsi que les informations associées, confirmant le bon fonctionnement du système de navigation.

Une autre étape importante de la validation consiste à isoler certaines structures anatomiques. Ainsi, une visualisation centrée uniquement sur la tumeur a été réalisée afin de vérifier les fonctionnalités de contrôle d'affichage. Comme présenté dans la figure suivante, seule la structure tumorale est affichée, permettant une analyse plus précise des régions pathologiques.

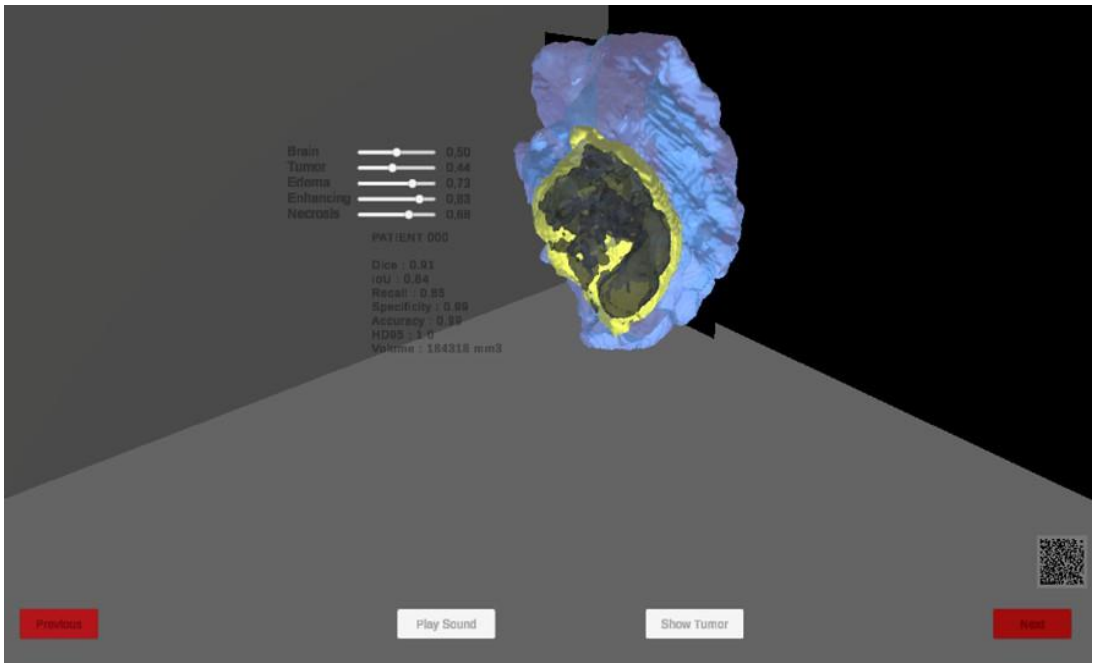


Figure 3.19 : Visualisation de la tumeur seule

Dans ce contexte, la fonctionnalité de sonification a été activée, permettant la génération d'un signal audio associé au patient sélectionné. Elle est déclenchée via le bouton Play Sound et assure une représentation audio des données médicales, enrichissant ainsi l'analyse multimodale du système.

La méthode de sonification repose sur les principes suivants :

- Le signal audio est généré à partir des métriques de segmentation (score Dice et volume tumoral).
- La fréquence du son varie proportionnellement au score Dice (plus le score est élevé, plus le son est aigu), permettant une interprétation auditive intuitive.
- L'amplitude du signal est adaptée pour une écoute confortable, avec une durée fixe d'une seconde.
- Le fichier audio est sauvegardé au format WAV et associé au patient correspondant.
- Cette approche offre une perception alternative des données médicales, complémentaire à la visualisation.

Bien que cette fonctionnalité soit opérationnelle, aucune évaluation clinique ou perceptive n'a été réalisée dans le cadre de ce travail. Elle constitue une démonstration technique exploratoire, ouvrant des perspectives pour une analyse multimodale future.

Par ailleurs, le système intègre un contrôle avancé de la visibilité des structures anatomiques à travers des sliders d'opacité. Ces derniers permettent d'ajuster dynamiquement la transparence du cerveau, de la tumeur et des différentes sous-régions pathologiques. Cette fonctionnalité permet une exploration progressive des structures et facilite leur interprétation.

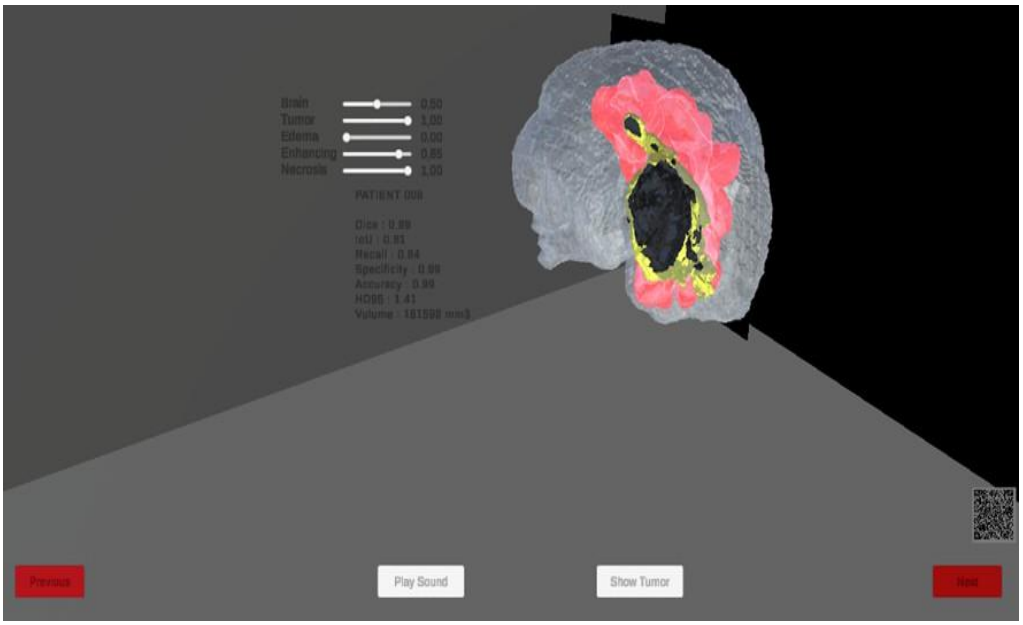


Figure 3.20 : Contrôle de l'opacité des structures anatomiques à l'aide des sliders

En complément, une interaction supplémentaire permet l'affichage ou la disparition d'un code QR associé à chaque patient. Lorsqu'il est activé, le QR code apparaît en grand format et peut être utilisé pour accéder aux informations du patient ou aux résultats associés. Un simple clic permet ensuite de le masquer, assurant une interaction fluide et intuitive.

L'interface utilisateur affiche également les informations relatives à chaque patient, positionnées de manière cohérente avec les éléments de contrôle afin de garantir une lecture claire des données. Les sliders sont disposés de manière ergonomique, tandis que les informations textuelles sont affichées en bas de l'interface pour une meilleure lisibilité.

Sur le plan technique, l'application a été générée via le processus de Build sous Unity, permettant la création d'un exécutable installé directement sur le téléphone. Une fois lancée, l'application utilise la caméra du dispositif pour superposer les modèles tridimensionnels et la scène Unity dans l'environnement réel. Cela permet une visualisation immersive où les structures cérébrales sont projetées dans le monde réel, reproduisant un contexte proche de la visualisation préopératoire.

Enfin, les tests techniques réalisés dans l'environnement Unity confirment le bon fonctionnement global du système de réalité augmentée. Les interactions (rotation, zoom, déplacement, contrôle de transparence) sont fluides et réactives. La navigation entre les patients est cohérente : le passage d'un cas à l'autre met à jour automatiquement les modèles 3D, les métriques, le QR code et la sonification, sans latence perceptible.

La qualité de la visualisation tridimensionnelle est satisfaisante : les maillages 3D (cerveau, tumeur, œdème, nécrose) sont affichés avec un niveau de détail suffisant pour une interprétation anatomique. Ces résultats ont été obtenus dans des conditions de tracking stable. Une dégradation du tracking pourrait affecter la précision de la superposition des modèles virtuels sur l'environnement réel.

3.2.7.3 *Analyse comparative réalité augmentée 2D vs réalité augmentée 3D et discussion*

L'exploitation de la réalité augmentée en 3D met en évidence une amélioration significative de la cohérence spatiale dans la représentation des structures cérébrales segmentées. Toutefois, cette approche introduit certaines limites liées à la complexité de reconstruction des volumes et à la sensibilité aux erreurs de segmentation initiale. En effet, toute imprécision au niveau voxel peut être amplifiée lors de la reconstruction tridimensionnelle, ce qui peut affecter la fidélité globale du modèle. De plus, l'intégration en environnement immersif impose des contraintes supplémentaires en termes de performance et de stabilité d'affichage, notamment lors de la manipulation en temps réel des volumes. Malgré ces contraintes, la représentation 3D reste plus informative sur le plan anatomique et constitue une base plus riche pour les analyses avancées et les applications intelligentes.

La comparaison entre les deux niveaux de réalité augmentée met en évidence une évolution progressive entre une représentation locale et une représentation globale des structures cérébrales. Les résultats montrent que la version 2D, bien que stable et facile à interpréter, reste fortement dépendante de la qualité des coupes individuelles, ce qui limite la cohérence globale des données utilisées en aval. À l'inverse, la représentation 3D permet une meilleure intégration des informations spatiales, mais elle rend également plus visible l'impact des erreurs issues de la segmentation initiale. Cette observation suggère que la qualité de la segmentation ne se limite pas à l'étape de visualisation, mais peut également influencer les tâches secondaires, notamment la classification. En effet, les différences observées entre les performances obtenues sur les données originales et celles reconstruites à partir des segmentations indiquent que la structuration des données en amont joue un rôle déterminant dans la robustesse des modèles de classification.

Dans ce contexte, la problématique suivante se pose : dans quelle mesure la qualité de la segmentation influence-t-elle les performances des modèles de classification des tumeurs cérébrales ?

3.2.8 *De la segmentation à la classification : impact de l'enrichissement des données*

Dans cette partie, l'analyse porte sur l'exploitation des résultats issus du pipeline de segmentation dans un cadre de classification des tumeurs cérébrales. L'objectif est d'évaluer l'impact de la représentation des données sur les performances des modèles de machine learning. L'étude repose sur une représentation en deux dimensions, correspondant aux images IRM enrichies par les résultats de segmentation, utilisées comme entrée pour la tâche de classification. Cette analyse permet d'évaluer l'influence de la qualité de la segmentation sur la robustesse et la généralisation des modèles de classification.

3.2.8.1 *Benchmarking des architectures de classification*

Dans cette première étape expérimentale, une étude comparative a été réalisée afin d'évaluer les performances de quatre architectures de deep learning largement utilisées dans les tâches de classification d'images médicales, à savoir ResNet50, DenseNet121, MobileNet et EfficientNet-B0. Le choix de ces architectures n'est pas arbitraire, mais repose sur leur complémentarité en termes de profondeur, de mécanismes d'extraction de caractéristiques et de compromis entre performance et complexité computationnelle.

En effet, ResNet50 est une architecture de référence dans le domaine de la vision par ordinateur, introduisant les connexions résiduelles permettant de résoudre le problème de disparition du gradient dans les réseaux profonds. Elle est souvent utilisée comme modèle baseline pour évaluer la capacité d'apprentissage des réseaux profonds sur des données complexes.

DenseNet121, quant à lui, repose sur un principe de connexions denses entre les couches, favorisant la réutilisation des caractéristiques et une meilleure propagation de l'information. Cette architecture est particulièrement adaptée aux images médicales, où les détails fins et les textures jouent un rôle important dans la discrimination des classes.

MobileNet a été sélectionné pour sa légèreté et son efficacité computationnelle. Grâce à l'utilisation de convolutions séparables en profondeur, il permet de réduire considérablement le nombre de paramètres tout en maintenant de bonnes performances, ce qui le rend pertinent dans des contextes où les ressources sont limitées.

Enfin, EfficientNet-B0 représente une architecture récente et optimisée, basée sur une stratégie de scaling composé (compound scaling), qui ajuste simultanément la profondeur, la largeur et la résolution du réseau. Cette approche permet d'obtenir un excellent compromis entre précision et efficacité, ce qui en fait un modèle particulièrement performant pour les tâches de classification en imagerie médicale.

Ainsi, ces quatre architectures ont été choisies afin de couvrir différents paradigmes de conception des réseaux de neurones convolutionnels, allant des modèles classiques aux architectures les plus récentes et optimisées.

Toutes les architectures ont été entraînées dans des conditions identiques afin d'assurer une comparaison équitable :

- Optimizer : Adam

- Learning rate : 0.0005
- Batch size : 8
- Image size : 224 × 224

Les performances ont été évaluées sur plusieurs configurations d'entraînement en utilisant trois métriques principales : Accuracy, F1-score et coefficient de Kappa, permettant ainsi d'analyser à la fois la performance globale, l'équilibre entre les classes et la robustesse des prédictions.

3.2.8.2 *Résultats sans réalité augmentée*

2***Modèle**	**Accuracy / F1-score / Kappa**			
	5 epochs	**10 epochs**	**20 epochs**	**25 epochs**
ResNet	0,907 / 0,908 / 0,873	0,942 / 0,941 / 0,920	0,887 / 0,895 / 0,846	0,871 / 0,870 / 0,823
EfficientNet	0,984 / 0,986 / 0,978	0,964 / 0,965 / 0,951	0,995 / 0,995 / 0,993	0,992 / 0,992 / 0,989
DenseNet	0,951 / 0,950 / 0,933	0,965 / 0,967 / 0,952	0,983 / 0,985 / 0,977	0,997 / 0,976 / 0,969
MobileNet	0,970 / 0,972 / 0,959	0,976 / 0,977 / 0,967	0,972 / 0,971 / 0,962	0,969 / 0,966 / 0,958

TAB. 3.8 : Performances des architectures de classification sans réalité augmentée

Résultats à 5 epochs

À 5 époques, les modèles présentent un apprentissage initial rapide mais encore non stabilisé.

EfficientNet-B0 obtient les meilleures performances (Accuracy = 0.984, F1 = 0.986), ce qui s'explique par sa capacité à extraire efficacement les caractéristiques multi-échelles grâce à son architecture optimisée (compound scaling). Cela lui permet de converger rapidement dès les premières itérations.

MobileNet arrive en deuxième position (Accuracy = 0.970), ce qui est dû à sa structure légère basée sur des convolutions séparables, favorisant un apprentissage rapide mais moins profond.

DenseNet atteint des performances intermédiaires (0.951), grâce à ses connexions denses qui facilitent la propagation des informations, mais qui peuvent introduire une certaine redondance dans les features.

ResNet présente les performances les plus faibles (0.907), ce qui peut s'expliquer par la profondeur du réseau nécessitant davantage d'epochs pour exploiter pleinement les connexions résiduelles.

EfficientNet montre une convergence plus rapide et une meilleure capacité d'extraction dès le début de l'entraînement.

Résultats à 10 epochs

À 10 époques, une amélioration globale est observée pour toutes les architectures.

EfficientNet reste performant (Accuracy = 0.964) mais montre une légère variation, ce qui est normal dans la phase d'ajustement des poids.

MobileNet progresse fortement (0.976), grâce à sa capacité d'apprentissage rapide sur des structures simples.

DenseNet améliore ses performances (0.965), ce qui s'explique par sa capacité à réutiliser les features extraites des couches précédentes.

ResNet progresse également (0.942), mais reste limité par une convergence plus lente.

Cette étape confirme une amélioration globale des modèles avec l'augmentation des époques, avec une stabilité relative d'EfficientNet.

Résultats à 20 epochs

À 20 époques, les modèles atteignent leur performance maximale.

EfficientNet-B0 obtient les meilleurs résultats globaux (Accuracy = 0.995, F1 = 0.995, Kappa = 0.993), ce qui confirme sa forte capacité de généralisation. Cela est dû à son architecture optimisée qui équilibre profondeur, largeur et résolution.

DenseNet atteint également de très bonnes performances (0.983), grâce à la forte circulation des informations entre couches.

MobileNet reste performant (0.972), mais montre une capacité limitée à exploiter des représentations complexes.

ResNet obtient une performance plus faible (0.887), ce qui peut être lié à une saturation ou une adaptation moins efficace aux caractéristiques des données médicales.

Le point optimal d'apprentissage est atteint à 20 époques, où EfficientNet domine clairement.

Résultats à 25 epochs

À 25 époques, on observe un début de surapprentissage pour certaines architectures.

EfficientNet reste stable (0.992), ce qui confirme sa robustesse.

DenseNet et MobileNet montrent une légère baisse des performances, indiquant une sensibilité au surapprentissage.

ResNet continue de décroître (0.871), ce qui confirme sa faible stabilité sur ce dataset.

L'augmentation des époques au-delà de 20 n'améliore plus les performances et introduit un léger surapprentissage.

Les résultats expérimentaux montrent que les performances augmentent progressivement jusqu'à 20 époques, puis se stabilisent ou diminuent légèrement.

Parmi les architectures testées, EfficientNet-B0 se distingue comme le modèle le plus performant et le plus stable grâce à sa capacité d'extraction multi-échelle et sa bonne généralisation sur les données médicales.

3.2.8.3 *Résultats avec réalité augmentée*

2***Modèle**	**Accuracy / F1-score / Kappa**			
	5 epochs	**10 epochs**	**20 epochs**	**25 epochs**
ResNet	0,944 / 0,952 / 0,923	0,939 / 0,942 / 0,917	0,926 / 0,917 / 0,900	0,973 / 0,974 / 0,963
EfficientNet	0,961 / 0,965 / 0,947	0,978 / 0,978 / 0,970	0,984 / 0,985 / 0,978	0,984 / 0,985 / 0,978
DenseNet	0,966 / 0,968 / 0,954	0,867 / 0,884 / 0,818	0,944 / 0,949 / 0,923	0,980 / 0,981 / 0,973
MobileNet	0,960 / 0,963 / 0,945	0,977 / 0,979 / 0,969	0,981 / 0,982 / 0,974	0,969 / 0,972 / 0,958

TAB. 3.9 : Performances des architectures de classification avec réalité augmentée

Résultats à 5 epochs

À ce stade initial, les modèles montrent déjà de bonnes performances malgré la complexité supplémentaire introduite par la fusion des données. On observe que DenseNet et MobileNet présentent légèrement de meilleures performances que les autres architectures.

Cela peut s'expliquer par leurs mécanismes internes :

- DenseNet121 réutilise les caractéristiques à travers ses connexions denses, ce qui lui permet d'exploiter efficacement les informations supplémentaires issues de la segmentation.
- MobileNet, grâce à ses convolutions légères et efficaces, apprend rapidement les structures globales même avec des données enrichies.
- EfficientNet reste stable mais n'exploite pas encore pleinement les informations supplémentaires à ce stade.
- ResNet50 montre des performances plus faibles, ce qui peut être lié à une adaptation plus lente aux données fusionnées.

Résultats à 10 epochs

Une évolution globale des performances est observée, notamment pour EfficientNet et MobileNet, qui atteignent des valeurs élevées et stables.

Cependant, DenseNet présente une chute importante des performances. Cela peut être expliqué par :

1. Une sensibilité élevée aux données bruitées introduites par la segmentation automatique,

2. Une instabilité dans la propagation des caractéristiques denses, surtout lorsque la distribution des données change.

EfficientNet, au contraire, reste robuste grâce à sa capacité de scaling équilibré (profondeur, largeur, résolution), ce qui lui permet de mieux généraliser sur des données complexes.

Résultats à 20 epochs

À 20 époques, les modèles atteignent globalement leur meilleure convergence.

- EfficientNet obtient les meilleures performances globales, confirmant sa robustesse.
- MobileNet se rapproche fortement d'EfficientNet, grâce à sa capacité à extraire efficacement les caractéristiques essentielles.
- DenseNet se stabilise après la chute observée à 10 epochs, ce qui indique une adaptation progressive aux données enrichies.
- ResNet reste en retrait, ce qui montre une difficulté à exploiter pleinement les informations de segmentation.

Cette étape représente le point optimal d'apprentissage, où le compromis entre généralisation et surapprentissage est le plus équilibré.

Résultats à 25 epochs

À 25 époques, on observe une stabilisation générale des performances.

- EfficientNet reste stable, confirmant sa robustesse.
- DenseNet rattrape progressivement ses performances après son instabilité initiale, grâce à sa capacité de réutilisation des features.
- ResNet s'améliore légèrement mais reste dépendant d'un apprentissage plus long.
- MobileNet montre une légère baisse, indiquant un début de saturation de sa capacité de représentation.

Les résultats expérimentaux montrent que la base de données enrichie par segmentation améliore la richesse des informations disponibles pour la classification, mais augmente également la complexité d'apprentissage.

Cette complexité explique :

- les fluctuations observées pour certaines architectures (DenseNet, ResNet),
- la stabilité remarquable d'EfficientNet,
- et la performance rapide mais limitée de MobileNet.

3.2.8.4 *Sélection du meilleur modèle*

Ainsi, la segmentation agit comme un facteur d'enrichissement des données, mais son impact dépend fortement de l'architecture.

Après l'évaluation des différentes architectures de classification (ResNet50, DenseNet121, MobileNet et EfficientNet-B0) sur les deux bases de données (originale et enrichie par segmentation), une analyse globale des performances a été réalisée afin d'identifier le modèle le plus performant et le plus stable.

Les résultats expérimentaux montrent de manière constante que l'architecture EfficientNet-B0 surpasse les autres modèles en termes d'Accuracy, de F1-score et de coefficient de Kappa, et ce dans toutes les configurations d'entraînement (5, 10, 20 et 25 epochs), ainsi que sur les deux bases de données.

Cette supériorité s'explique par la conception spécifique d'EfficientNet-B0, basée sur le compound scaling, qui permet un équilibre optimal entre la profondeur, la largeur et la résolution du réseau. Cette architecture est particulièrement adaptée aux données médicales, car elle permet une extraction efficace des caractéristiques multi-échelles tout en limitant la complexité computationnelle.

Par ailleurs, les résultats montrent que le meilleur compromis entre performance et généralisation est atteint autour de 20 epochs pour l'ensemble des architectures. Au-delà de ce seuil, les performances tendent à se stabiliser, voire à diminuer légèrement pour certaines architectures, indiquant un début de surapprentissage.

Il est également important de noter que le même modèle optimal (EfficientNet-B0) est retenu pour les deux expériences. Cela montre que son efficacité est indépendante du type de représentation des données, qu'il s'agisse d'images IRM originales ou d'images enrichies par segmentation.

Ainsi, EfficientNet-B0 est sélectionné comme modèle final en raison de sa robustesse, de sa stabilité et de sa capacité de généralisation sur les deux types de données.

3.2.8.5 *Analyse du modèle retenu (EfficientNet-B0)*

- Analyse des courbes accuracy / loss

Afin d'analyser plus en détail le comportement du modèle sélectionné, une étude du processus d'apprentissage d'EfficientNet-B0 a été réalisée à travers les courbes d'accuracy et de loss pour les deux expériences.

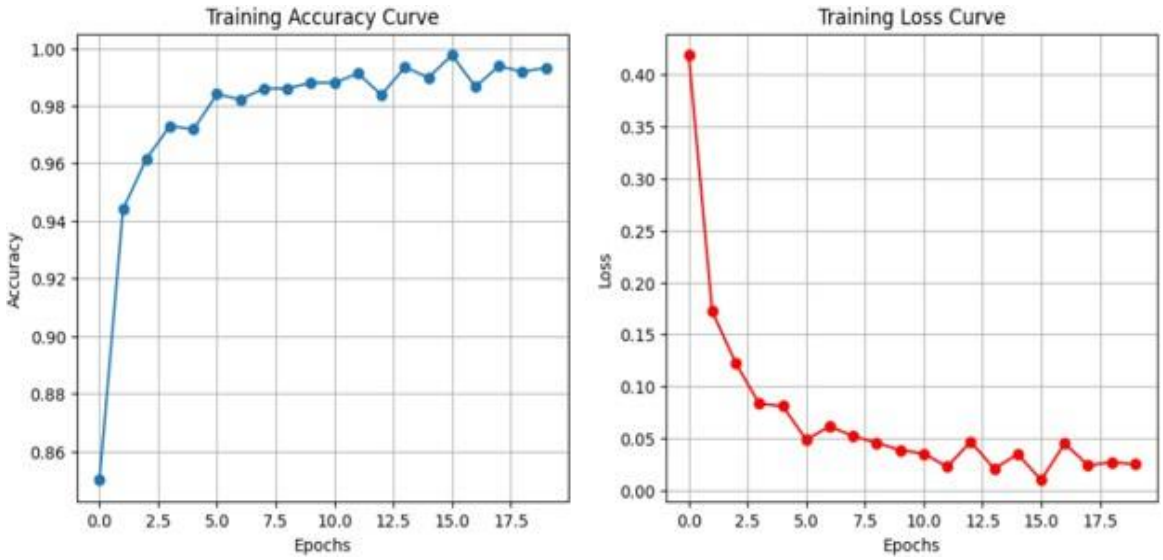


Figure 3.21 : Courbe EfficientNet sur base de données originale

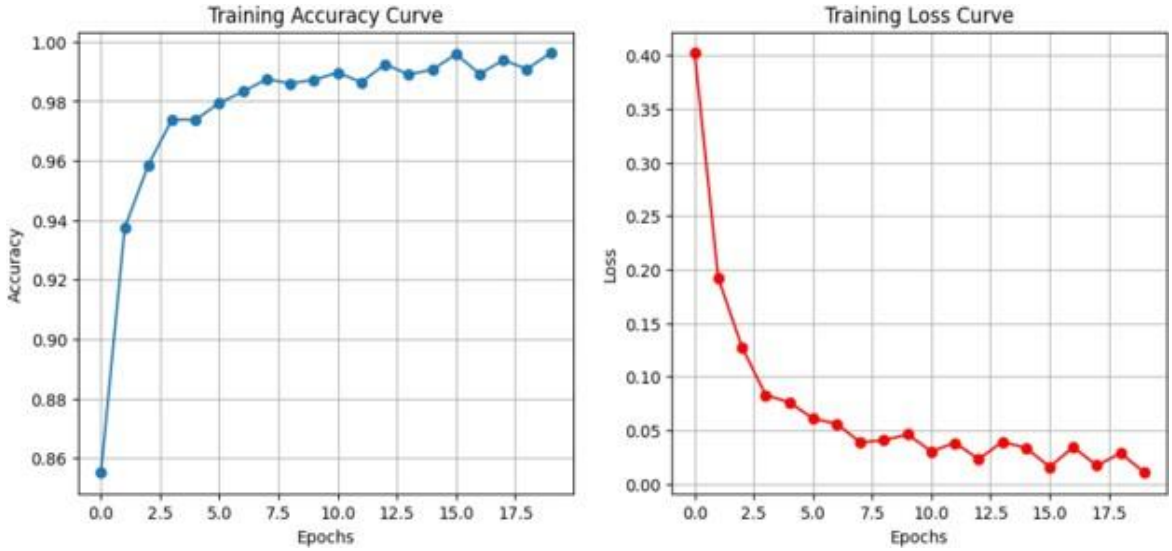


Figure 3.22 : Courbe EfficientNet sur base de données fusionnée (segmentation)

Les courbes d'apprentissage montrent un comportement globalement stable et cohérent du modèle EfficientNet-B0 dans les deux configurations expérimentales. Une augmentation progressive de l'accuracy est observée au fil des époques, tandis que la loss diminue régulièrement, traduisant une bonne convergence du réseau.

Pour la base de données originale, la convergence apparaît légèrement plus rapide, ce qui peut être attribué à la nature moins complexe des images IRM utilisées comme données d'entrée. Les performances élevées sont atteintes dès les premières époques avant de se stabiliser progressivement.

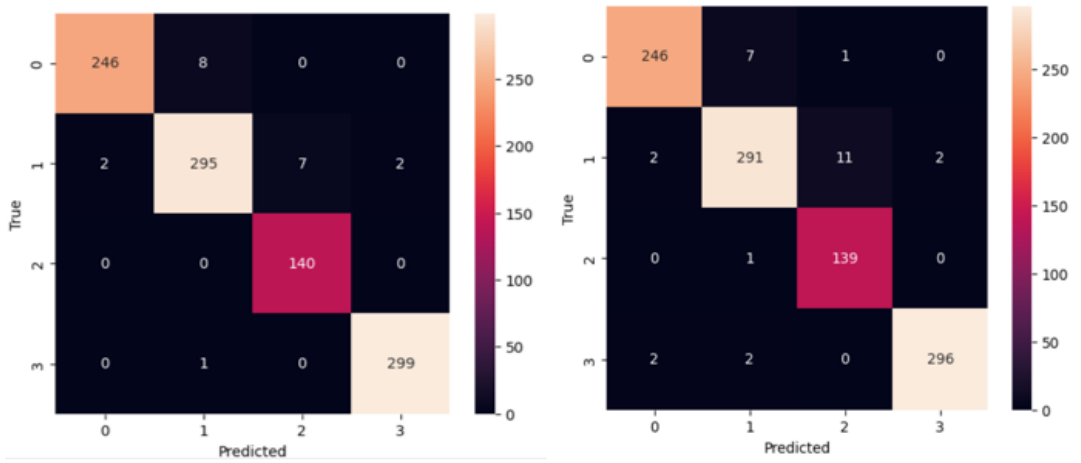
Dans le cas de la base de données fusionnée, l'apprentissage est légèrement plus progressif. Cette différence s'explique par l'intégration d'informations supplémentaires issues de la segmentation tumorale, qui enrichissent la représentation des données mais augmentent également la complexité de la tâche de classification.

Malgré cette complexité supplémentaire, les courbes montrent une convergence stable sans phénomène de surapprentissage important. Dans les deux expériences, une stabilisation des performances est observée autour de 20 époques, confirmant que ce nombre d'époques constitue le meilleur compromis entre qualité d'apprentissage et capacité de généralisation.

Ces résultats renforcent le choix d'EfficientNet-B0 comme modèle final de classification grâce à sa robustesse, sa stabilité d'entraînement et ses excellentes performances sur les deux bases de données étudiées.

- Analyse des matrices de confusion

Pour évaluer plus précisément les performances du modèle EfficientNet-B0, les matrices de confusion obtenues sur la base de données originale et sur la base de données enrichie par segmentation sont présentées ci-dessous. La matrice de la base de données originale est affichée à gauche, tandis que celle de la base de données fusionnée est présentée à droite.



(a) (b)

Figure 3.23 : Matrices de confusion du modèle EfficientNet-B0. (a) Matrice sur la base de données originale. (b) Matrice sur la base de données fusionnée (avec segmentation).

L'analyse de la matrice de confusion sur la base de données fusionnée donne les résultats suivants :

- Classe 0 : 246 correctement classés, avec 7 échantillons confondus avec la classe 1, 1 avec la classe 2, et 0 avec la classe 3.
- Classe 1 : 291 correctement classés, avec 2 erreurs vers la classe 0, 11 vers la classe 2, et 2 vers la classe 3.
- Classe 2 : 139 correctement classés, avec 0 erreur vers la classe 0, 1 vers la classe 1, et 0 vers la classe 3.
- Classe 3 : 296 correctement classés, avec 2 erreurs vers la classe 0, 2 vers la classe 1, et 0 vers la classe 2.

On observe donc une forte concentration des valeurs sur la diagonale principale : 246, 291, 139, 296, ce qui traduit une très bonne capacité de classification globale.

Cependant, quelques confusions apparaissent principalement entre les classes 1 et 2, ce qui peut être expliqué par la proximité visuelle de certaines caractéristiques tumorales après segmentation et fusion des données.

Pour la base de données originale, la matrice de confusion présente les résultats suivants :

- Classe 0 : 246 correctement classés, avec 0 erreur vers les classes 1, 2 et 3.
- Classe 1 : 295 correctement classés, avec 8 erreurs vers la classe 0, 0 vers la classe 2, et 1 vers la classe 3.
- Classe 2 : 140 correctement classés, avec 7 erreurs vers la classe 1, et 0 vers les autres classes.
- Classe 3 : 299 correctement classés, avec 2 erreurs vers la classe 1, et 0 vers les classes 0 et 2.

La diagonale principale reste dominante : 246, 295, 140, 299, ce qui confirme une classification très performante.

On remarque toutefois que certaines confusions apparaissent entre la classe 1 et la classe 0 ou 2, mais elles restent limitées.

La comparaison des deux matrices de confusion montre que le modèle EfficientNet-B0 présente d'excellentes performances dans les deux configurations expérimentales.

La base de données originale présente une classification légèrement plus stable avec moins de confusion globale.

La base de données fusionnée introduit davantage de complexité, ce qui augmente légèrement les confusions, notamment entre les classes 1 et 2.

Cependant, dans les deux cas, la forte dominance des valeurs diagonales confirme la robustesse du modèle et sa capacité à bien généraliser sur les différentes distributions de données.

Ainsi, EfficientNet-B0 conserve des performances élevées et stables malgré la variation de la structure des données.

- Analyse des courbes ROC

Dans le but d'évaluer la capacité discriminante du modèle EfficientNet-B0, les courbes ROC (Receiver Operating Characteristic) ont été tracées pour les deux expériences réalisées, à savoir la base de données originale et la base de données enrichie par segmentation.

Ces courbes permettent de mesurer la capacité du modèle à distinguer correctement les différentes classes en analysant le compromis entre le taux de vrais positifs (TPR) et le taux de faux positifs (FPR).

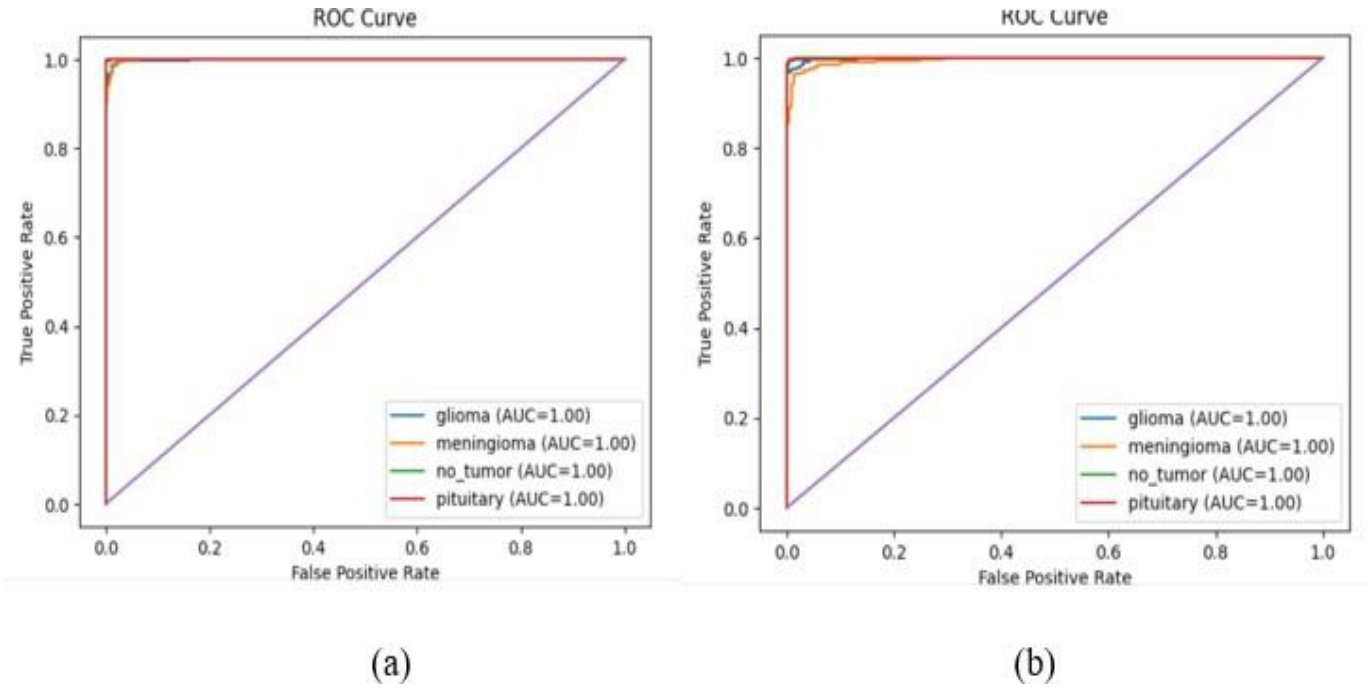


Figure 3.24 : Courbes ROC du modèle EfficientNet-B0 : (a) base de données originale, (b) base de données fusionnée

La courbe ROC obtenue sur la base de données fusionnée montre des performances globalement excellentes, avec des courbes proches du coin supérieur gauche, traduisant une très bonne capacité de classification.

On observe toutefois de très légers écarts principalement pour la classe meningioma, et dans une moindre mesure pour la classe glioma. Ces variations restent faibles et n'ont pas d'impact significatif sur les performances globales du modèle.

Pour la base de données originale, les courbes ROC montrent également des performances très élevées, avec des courbes proches de la zone idéale.

Cependant, de légers écarts sont observés pour les classes meningioma et glioma, avec des fluctuations similaires à celles de la base de données fusionnée.

La comparaison des courbes ROC montre que le modèle conserve une excellente capacité de séparation des classes dans les deux cas.

Dans les deux expériences, les courbes restent très proches de la zone optimale, ce qui confirme une forte capacité discriminative du modèle EfficientNet-B0.

L'ensemble des analyses réalisées à travers les métriques quantitatives, les courbes d'apprentissage, les matrices de confusion et les courbes ROC confirme la supériorité de l'architecture EfficientNet-B0 pour la tâche de classification des tumeurs cérébrales. Le modèle présente des performances élevées, une excellente stabilité d'apprentissage ainsi qu'une forte capacité de généralisation, aussi bien sur la base de données originale que sur la base de données enrichie par segmentation. Ces résultats justifient son adoption comme modèle final dans le cadre de ce travail.

En comparant les résultats obtenus sur la base de données originale et la base de données enrichie par segmentation (réalité augmentée), on observe que les performances des modèles de classification sont globalement légèrement meilleures sur la base de données originale, notamment en termes d'Accuracy, de F1-score et de stabilité globale. En effet, bien que la segmentation permette d'enrichir les images IRM en apportant des informations anatomiques supplémentaires sur les sous-régions tumorales, cette étape introduit également une complexité additionnelle ainsi qu'une sensibilité aux

erreurs de segmentation, ce qui peut expliquer une légère diminution ou instabilité des performances dans certains cas.

Cependant, malgré cette différence quantitative, les deux configurations montrent des performances globalement élevées et cohérentes, avec une forte capacité de discrimination confirmée par les matrices de confusion et les courbes ROC. Les résultats ROC, notamment, montrent une excellente séparation des classes dans les deux cas, avec seulement de légères variations pour certaines classes comme le gliome et le méningiome.

Ainsi, la réponse à la problématique est que la qualité de la segmentation influence effectivement les performances des modèles de classification : une segmentation précise améliore la richesse et la représentativité des données, tandis que les imperfections peuvent introduire du bruit et affecter légèrement les performances. Néanmoins, son principal apport reste l'enrichissement de l'information et l'amélioration de l'interprétabilité des résultats plutôt qu'une amélioration systématique des métriques de classification.

Comparaison avec l'état de l'art

Les comparaisons avec l'état de l'art présentées dans cette partie sont indicatives et dépendent des protocoles expérimentaux, des bases de données et des réglages spécifiques à chaque étude.

Segmentation 2D (FPN) sur BRISC 2025

Méthode	Dice	Année	Objectif du travail
FPN (Notre modèle)	0,8895	2026	Segmentation 2D légère avec ResNet34
U-Net++ (EfficientNet-B1) [62]	0,9055	2026	Maximiser la précision avec encodeur puissant
Ensemble 2.5D [63]	0,798	2025	Tester la fusion de modèles 2.5D

TAB. 3.10 : Comparaison des performances de segmentation 2D sur BRISC 2025

Notre modèle (Dice = 0,8895) est devancé par U-Net++ (0,9055) car U-Net++ utilise un encodeur EfficientNet-B1 plus large (plus de paramètres) et possède des connexions denses qui améliorent la propagation du gradient. Notre objectif était la légèreté, pas la maximisation absolue de la précision. Notre modèle surpasse l'Ensemble 2.5D (0,798) car l'approche 2.5D fusionne des informations de coupes voisines mais introduit du bruit, tandis que notre FPN 2D standard est plus adapté à la base BRISC.

Segmentation 3D (nnU-Net) sur BraTS 2020

Méthode	WT Dice	Année	Objectif du travail
Isensee et al. (nnU-Net 1ère place) [64]	0,8895	2020	Gagner le challenge BraTS 2020
Isensee et al. (nnU-Net original) [64]	0,912	2021	Framework auto-configurable universel
Magadza et al. (Efficient nnU-Net) [65]	0,912	2023	Réduire le coût computationnel
Raza et al. [66]	0,866	2023	Améliorer la segmentation des petites tumeurs
Wang et al. [67]	0,901	2022	Intégrer l'attention mécanisme
Liew et al. [68]	0,893	2023	Optimiser pour la région ET
Notre modèle (70 epochs/fold)	**0,8436**	**2026**	**Validation préliminaire (contraintes matérielles)**

TAB. 3.11 : Comparaison des performances de segmentation 3D sur BraTS 2020

Notre modèle (Dice = 0,8436) est inférieur aux références (0,8895 - 0,912) en raison d'un entraînement limité à 70 époques par fold (contre 1000 dans la littérature), de la validation croisée 5-fold qui réduit les données d'entraînement (80% seulement), des coupures électriques fréquentes au laboratoire (Annexe E), d'un GPU limité (GTX 1650 Ti 4Go) et de l'absence de post-traitement (region-based training, optimisations avancées). Malgré cela, notre score reste comparable à certaines études (Raza et al. 0,866, Liew et al. 0,893).

Nos résultats inférieurs s'expliquent par un entraînement limité (70 epochs/fold contre 1000), la validation croisée 5-fold et des contraintes matérielles (Annexe G).

Classification (EfficientNet) sur BRISC 2025

Méthode	Accuracy (%)	Année	Objectif du travail
EfficientNet-B1 [72]	99,70	2026	Version plus large, plus de paramètres
MobileViT V2-100 [69]	98,90	2025	Architecture légère pour mobiles
ResNet18 [70]	98,70	2025	Architecture classique de référence
EfficientNet-B0 (BRISC benchmark) [71]	98,70	2025	Établir une référence de base
Notre modèle (EfficientNet-B0)	**99,50**	**2026**	**Validation sur BRISC avec optimisation**

TAB. 3.12 : Comparaison des performances de classification sur BRISC 2025

Notre modèle (99,50%) surpasse les références BRISC (98,70-98,90%) car nous avons optimisé les hyperparamètres (learning rate 0,0005, batch size 8), utilisé l'augmentation de données (rotation, flip, bruit) et entraîné jusqu'à convergence (20-25 epochs). Notre modèle est légèrement devancé par EfficientNet-B1 (99,70%) car EfficientNet-B1 a plus de paramètres (7,8M contre 5,3M) et son objectif est la maximisation de la précision, pas la légèreté.

Synthèse des comparaisons

Nos résultats se situent dans la moyenne haute des performances reportées dans la littérature pour la classification (99,50%) et la segmentation 2D (Dice = 0,8895). Pour la segmentation 3D, nos résultats (Dice = 0,8436) sont légèrement inférieurs aux références en raison des contraintes expérimentales détaillées ci-dessus, mais restent acceptables pour une étude préliminaire.

3.3 *Conclusion*

Les résultats expérimentaux présentés dans ce chapitre valident l'approche proposée. En segmentation 2D, le modèle FPN atteint un Dice de 0,8895 sur BRISC 2025, démontrant sa capacité à délimiter les contours tumoraux. En segmentation 3D, nnU-Net obtient un Dice moyen de 0,8436 sur BraTS 2020, avec des scores par fold allant de 0,8270 à 0,8600, confirmant la stabilité du modèle. En classification, EfficientNet-B0 enregistre une accuracy de 99,50%, un F1-score de 0,985 et un Kappa de 0,978, ce qui témoigne d'une excellente capacité discriminante.

La réalité augmentée 2D permet de superposer les masques segmentés, les contours tumoraux et l'atlas MNI sur les images IRM. Cette visualisation facilite l'interprétation clinique pour les radiologues en mettant en évidence la localisation précise des lésions. L'interface développée offre également l'analyse quantitative (Dice, IoU, volume) et la classification automatique avec niveau de confiance, permettant une évaluation complète des résultats.

La réalité augmentée 3D, développée sous Unity, offre une visualisation immersive et interactive

des structures cérébrales. L'utilisateur peut manipuler les modèles (rotation, zoom, transparence), naviguer entre les patients et bénéficier de fonctionnalités complémentaires (QR code, sonification). Les temps de génération des maillages sont de 0,22 seconde pour le cerveau complet et 0,03 seconde pour la tumeur, permettant une expérience fluide en temps réel. Cette approche 3D constitue une avancée significative par rapport à la simple visualisation 2D, car elle restitue la continuité spatiale des structures anatomiques et facilite la planification préopératoire.

L'ensemble de ces résultats démontre l'efficacité du système NeuroVision pour la segmentation, la classification et la visualisation en réalité augmentée des tumeurs cérébrales. La combinaison des modèles d'intelligence artificielle (FPN, nnU-Net, EfficientNet) avec les environnements de réalité augmentée (2D et 3D) offre une solution complète et innovante pour l'aide au diagnostic en neuro-imagerie.

Conclusion générale

Ce travail de fin d'études a porté sur le développement de NeuroVision, une plateforme dédiée à l'analyse des tumeurs cérébrales à partir d'images IRM, combinant des techniques d'intelligence artificielle pour la segmentation et la classification, ainsi que des environnements de réalité augmentée 2D et 3D pour la visualisation des résultats.

Une étude approfondie de l'état de l'art a permis de situer notre travail par rapport aux méthodes existantes et d'identifier les architectures les plus adaptées : FPN pour la segmentation 2D, nnU-Net pour la segmentation 3D, et EfficientNet pour la classification multi-classes.

Nous avons ensuite conçu et développé l'ensemble du pipeline de traitement, incluant le prétraitement des données, l'entraînement des modèles sur les bases BRISC 2025 et BraTS 2020, ainsi que la création des interfaces de visualisation en réalité augmentée sous Unity.

Les expérimentations menées ont démontré la faisabilité et l'efficacité de l'approche proposée. Les résultats obtenus en segmentation 2D (Dice = 0,8895), segmentation 3D (Dice = 0,8436) et classification (Accuracy = 99,50%) sont encourageants et montrent le potentiel du système.

Les principales contributions de ce travail sont la mise en place d'un pipeline complet allant du prétraitement des IRM à la visualisation en réalité augmentée, l'évaluation comparative de plusieurs architectures de deep learning, le développement d'une interface de réalité augmentée 2D avec superposition des masques segmentés, des contours et de l'atlas MNI, ainsi que la création d'un environnement de réalité augmentée 3D sous Unity pour une visualisation immersive des structures cérébrales.

Plusieurs perspectives peuvent être envisagées pour étendre ce travail, notamment l'amélioration des performances de la segmentation 3D, l'extension de la base de données BRISC, l'intégration de la détection automatique des structures critiques, la validation clinique du système et le développement d'une version mobile de l'application.

NeuroVision constitue une première étape prometteuse vers l'intégration de l'intelligence artificielle et de la réalité augmentée dans le diagnostic des tumeurs cérébrales, démontrant la faisabilité d'une solution locale performante et accessible.

Annexes

Environnement matériel et logiciel de développement :

Afin de développer et d'évaluer les différentes composantes du système proposé, plusieurs environnements logiciels ont été utilisés au cours de ce travail. Ces outils ont permis d'assurer le prétraitement des données, l'entraînement des modèles d'intelligence artificielle, l'exécution des algorithmes de segmentation et de classification, ainsi que le développement des interfaces de visualisation.

Les principaux environnements exploités sont Anaconda Prompt, Kaggle, Google Colab, Spyder, Unity, C# et Visual Studio Code. Chacun de ces outils a été utilisé selon les besoins spécifiques des différentes étapes du projet.

1) Anaconda Prompt

Anaconda Prompt est un environnement en ligne de commande intégré à la distribution Anaconda, largement utilisé dans les projets de science des données et d'intelligence artificielle. Il permet d'accéder directement aux fonctionnalités de Conda, un gestionnaire de paquets et d'environnements particulièrement adapté au développement en Python.

Dans ce projet, Anaconda Prompt a principalement été utilisé pour :

- L'installation des bibliothèques nécessaires (PyTorch, TensorFlow, scikit-learn, segmentation-models-pytorch, etc.) ;
- La configuration des environnements de travail isolés pour chaque phase du projet ;
- L'exécution des commandes liées aux modèles de segmentation nnU-Net ;
- Le lancement des scripts d'entraînement et de prétraitement à grande échelle ;
- La gestion des dépendances logicielles et la résolution des conflits de versions.

L'utilisation d'Anaconda a permis de garantir une meilleure stabilité du système et une compatibilité optimale entre les différentes bibliothèques utilisées tout au long du pipeline de traitement.

2) Kaggle

Kaggle est une plateforme collaborative dédiée à la science des données et à l'apprentissage automatique. Elle fournit un environnement cloud intégrant des ressources de calcul GPU (jusqu'à 30 heures par semaine) ainsi que des notebooks interactifs permettant l'exécution des modèles d'intelligence artificielle sans configuration matérielle complexe.

Dans ce travail, Kaggle a été utilisé pour :

- L'entraînement intensif des modèles de segmentation et de classification ;
- L'exploitation des ressources GPU (NVIDIA Tesla P100) pour réduire considérablement le temps de calcul ;

- Le stockage et l'organisation des bases de données médicales (BraTS2020, BRISC 2025) ;
- L'exécution des expériences d'apprentissage profond avec des volumes importants de données ;
- L'évaluation comparative des performances des modèles sur les datasets de validation.

L'utilisation de Kaggle a facilité le traitement des volumes importants de données IRM (plusieurs centaines de gigaoctets) tout en offrant un environnement adapté aux expérimentations en apprentissage profond (DL). La plateforme a également permis une reproductibilité des résultats grâce à la sauvegarde automatique des notebooks et des sorties d'exécution.

3) Google Colab

Google Colab est un environnement de développement basé sur les notebooks Jupyter et hébergé sur le cloud. Il permet l'exécution de programmes Python avec accès à des ressources GPU (NVIDIA Tesla T4 ou K80) et TPU, sans nécessiter d'installation locale.

Dans le cadre de ce projet, Google Colab a été utilisé pour :

- Tester rapidement certaines architectures de réseaux neuronaux (ResNet, DenseNet, EfficientNet, MobileNet) ;
- Exécuter des expériences de segmentation et de classification à petite échelle pour le prototypage ;
- Visualiser les résultats intermédiaires (courbes d'apprentissage, matrices de confusion, cartes de segmentation) ;
- Effectuer des traitements nécessitant une exécution flexible et collaborative (partage de notebooks avec l'équipe).

Google Colab a également permis d'accélérer les phases de prototypage et de validation des modèles grâce à son accessibilité (aucune configuration matérielle requise) et à sa simplicité d'utilisation. L'intégration native avec Google Drive a facilité le stockage et le partage des données et des résultats.

4) Spyder

Spyder (Scientific Python Development Environment) est un environnement de développement intégré (IDE) dédié au langage Python. Il offre une interface adaptée à la programmation scientifique et au traitement d'image, avec des fonctionnalités telles que l'éditeur avec analyse syntaxique, la console IPython intégrée, et l'inspecteur de variables.

Dans ce projet, Spyder a été principalement utilisé pour :

- Le développement et le débogage des scripts Python pour le prétraitement des images ;
- L'implémentation des algorithmes de traitement d'image (filtrage, normalisation, morphologie) ;

- La visualisation interactive des résultats de segmentation (superposition image GT/prédiction) ;
- Le développement de certaines fonctionnalités liées à l'analyse médicale et à la visualisation 2D ;
- L'analyse exploratoire des données et le calcul des métriques d'évaluation.

Son interface interactive a facilité le suivi des variables, l'analyse des résultats et l'exécution progressive du code, permettant ainsi un développement itératif et efficace des algorithmes de traitement.

L'association de ces différents environnements a permis de construire une chaîne de développement complète adaptée aux contraintes du projet. Le tableau ci-dessous récapitule l'utilisation de chaque outil selon les phases du projet.

5) Unity et C#

Unity est un moteur de jeu multiplateforme largement utilisé pour le développement d'applications 2D et 3D, incluant les environnements de réalité virtuelle et augmentée. Il offre un moteur de rendu performant, un système physique réaliste et une interface de développement intuitive.

Dans ce projet, Unity a été utilisé comme moteur principal pour le développement de l'environnement de réalité augmentée et de visualisation 3D. Il a permis la création d'une interface immersive pour la manipulation et l'exploration des structures cérébrales reconstruites à partir des résultats de segmentation.

Le langage C# (C Sharp), langage de programmation orienté objet développé par Microsoft, a été utilisé comme langage principal pour l'écriture des scripts dans Unity. Il a permis de :

- Développer les scripts de contrôle des objets 3D (cerveau, tumeur, œdème, nécrose) ;
- Gérer les interactions utilisateur (rotation, zoom, sélection de structures) via des événements ;
- Implémenter la navigation entre les différentes données patientes (boutons Next/Previous) ;
- Assurer l'intégration et l'affichage dynamique des volumes reconstruits au format OBJ ;
- Synchroniser les modules de classification avec la visualisation en réalité augmentée ;
- Contrôler la transparence des structures anatomiques via des sliders d'opacité ;
- Générer des représentations sonores (sonification) associées aux données du patient.

6) Visual Studio Code

Dans le cadre du développement de l'interface utilisateur du système NeuroVision, l'environnement de travail utilisé est Visual Studio Code (VS Code). Il s'agit d'un éditeur de code léger, gratuit et open-source développé par Microsoft, qui constitue un outil puissant et extensible adapté au développement web et aux applications multiplateformes.

VS Code offre de nombreuses fonctionnalités ayant facilité le développement :

- Colorisation syntaxique pour HTML, CSS et JavaScript ;
- Auto-complétion intelligente (IntelliSense) ;
- Débogage intégré avec points d'arrêt ;
- Intégration native avec Git pour le contrôle de version ;
- Extensions riches pour le développement web (Live Server, Prettier, ESLint).

Le développement de l'interface a été réalisé en utilisant les technologies suivantes :

- **HTML5 (HyperText Markup Language)** : pour la structure des pages web, la définition des éléments d'interface (formulaires, boutons, conteneurs) et l'organisation sémantique du contenu ;
- **CSS3** : pour la mise en forme, le style visuel, la responsivité et l'adaptation aux différentes tailles d'écran ;
- **JavaScript (ES6)** : pour l'interactivité, la manipulation dynamique du DOM, la communication asynchrone avec le backend (API REST) et le traitement des événements utilisateur.

Ces technologies ont permis de concevoir une interface dynamique, interactive et adaptée aux besoins du système médical proposé. L'interface développée offre les fonctionnalités suivantes :

- Affichage en temps réel des résultats de segmentation et de classification ;
- Visualisation des métriques de performance (Accuracy, F1-score, Kappa, Dice) ;
- Navigation intuitive entre les différents patients ;
- Contrôle des paramètres d'affichage (opacité, sélection des structures) ;
- Génération de rapports visuels (matrices de confusion, courbes ROC) ;
- Accès aux informations patient et aux historiques de diagnostic.

Environnement	Phase d'utilisation	Rôle principal
Anaconda Prompt	Configuration & déploiement	Gestion des dépendances, lancement des scripts nnU-Net
Kaggle	Entraînement intensif	Calcul GPU massif sur BraTS2020 et BRISC 2025
Google Colab	Prototypage rapide	Tests d'architectures, visualisations collaboratives
Spyder	Développement local	Implémentation des algorithmes, débogage interactif
Unity & C#	Visualisation 3D	Réalité augmentée, environnement immersif
Visual Studio Code	Interface web	Développement frontend (HTML/JavaScript)

TAB. 3.13 : Récapitulatif des environnements de développement utilisés

Les plateformes cloud telles que Kaggle et Google Colab ont été utilisées pour les calculs intensifs liés à l'apprentissage profond, tandis qu'Anaconda Prompt et Spyder ont servi au développement local, à la gestion des bibliothèques et à l'implémentation des traitements d'image et des outils d'analyse médicale.

La combinaison de ces différents environnements de développement a permis de mener à bien l'ensemble des phases du projet, depuis le prétraitement des données IRM jusqu'à la visualisation en réalité augmentée. Le tableau 3.17 synthétise l'articulation entre les outils utilisés et leurs rôles respectifs dans la chaîne de traitement.

Cette approche modulaire et diversifiée a offert une grande flexibilité, permettant d'exploiter les avantages spécifiques de chaque plateforme : la puissance de calcul cloud pour l'entraînement des modèles profonds, la flexibilité des IDE locaux pour le développement algorithmique, et les capacités immersives d'Unity pour la visualisation 3D et la réalité augmentée.

Mise en place de l'environnement de réalité augmentée

Mise en place de l'environnement de réalité augmentée La mise en place de l'environnement de développement constitue la première étape de l'implémentation du système de réalité augmentée. Elle vise à installer et configurer les outils nécessaires au développement d'une application interactive sous Unity, dédiée à la visualisation de structures cérébrales en trois dimensions.

Dans ce projet, l'environnement a été conçu pour fonctionner sur des appareils mobiles, en particulier les smartphones Android, afin de garantir une solution accessible et facilement déployable. Ce choix permet d'éviter la dépendance à des dispositifs spécialisés tels que les casques de réalité augmentée, tout en conservant une expérience immersive basée sur l'utilisation de la caméra et des capteurs du téléphone.

Sur le plan technique, la réalisation repose sur Unity, choisi pour ses capacités de rendu 3D et son support natif de la réalité augmentée via AR Foundation. Ce framework permet l'intégration des fonctionnalités ARCore, assurant la détection de l'environnement et le suivi spatial en temps réel.

La configuration de cet environnement inclut l'installation de Unity Hub, Unity Editor, ainsi que des packages AR Foundation et ARCore XR Plugin, qui constituent la base du système de réalité augmentée utilisé dans ce projet.

1. Installation de Unity Hub Unity Hub constitue l'outil principal de gestion de l'environnement de développement Unity. Il permet de centraliser les projets, de gérer les versions du moteur Unity ainsi que de faciliter l'installation des différents modules nécessaires au développement de l'application.

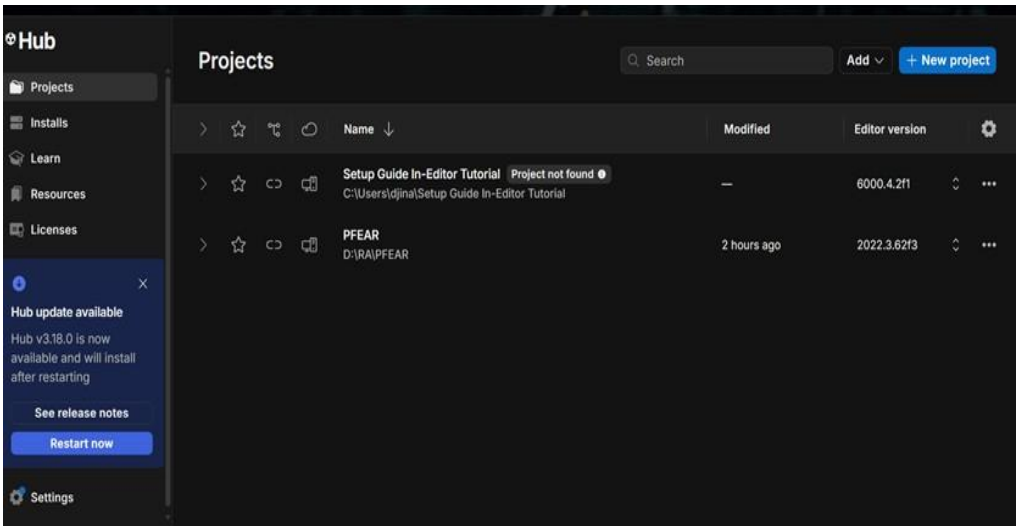


Figure 25 : Interface Unity Hub

2. Installation de Unity Editor L'installation de Unity Editor représente une étape fondamentale dans la mise en place de l'environnement de développement. Une version compatible avec AR Foundation a été sélectionnée afin de garantir le bon fonctionnement des fonctionnalités de réalité augmentée et assurer la stabilité du projet sur les appareils Android ciblés.

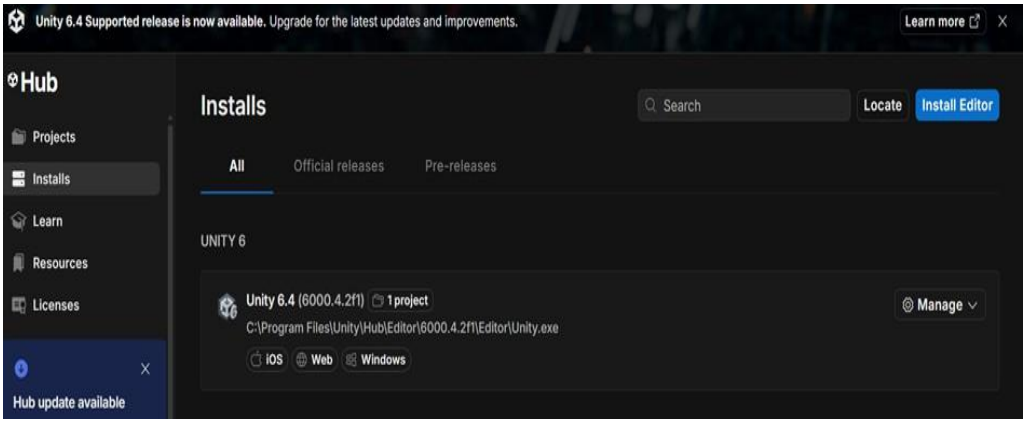


Figure 26 : Version Unity 6000.4.2f1 installée dans Unity Hub

3. Installation Template et création du projet Lors de la création du projet, Unity propose plusieurs templates adaptés à différents types d'applications (2D, 3D, VR, AR, etc.).

Dans ce travail, le template AR Mobile a été sélectionné afin de configurer automatiquement les composants nécessaires à la réalité augmentée sur appareils mobiles.

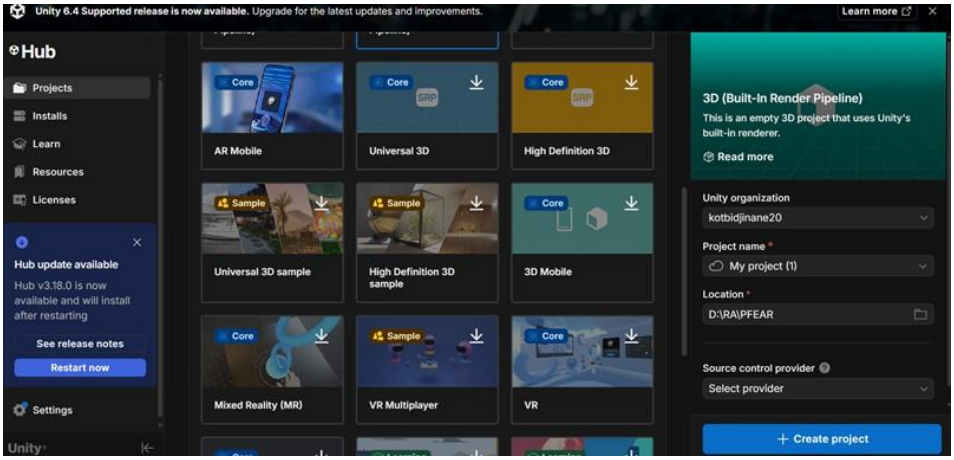


Figure 27 : Sélection du template AR Mobile

Après la sélection du template AR Mobile, le projet a été créé sous le nom PFEAR. Cette étape permet de générer la structure initiale du projet Unity, incluant les dossiers de base et les configurations nécessaires au développement.

4. Installation packages RA Afin d'activer les fonctionnalités de réalité augmentée, plusieurs packages ont été installés via le Package Manager de Unity. Ces composants permettent l'intégration du tracking, de la détection de surfaces et de l'ancrage des objets 3D dans l'environnement réel.

Les packages installés sont :

- AR Foundation : framework principal de développement AR multiplateforme
- ARCore XR Plugin : support des fonctionnalités AR sur Android
- XR Plugin Management : gestion des systèmes XR et des plateformes

Ces éléments assurent la détection de l'environnement réel, le suivi de mouvement ainsi que le positionnement des modèles 3D dans l'espace physique via la caméra du smartphone.

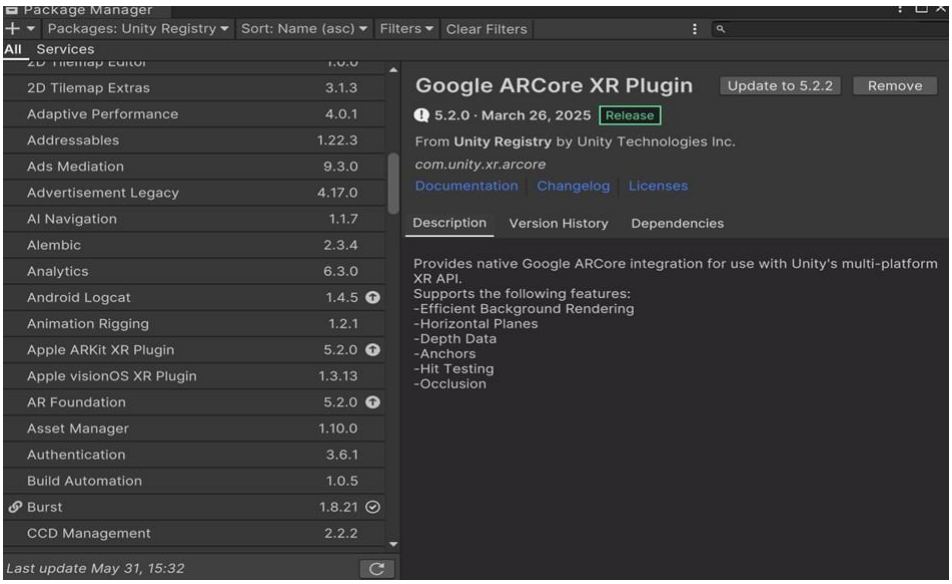


Figure 28 : Packages de réalité augmentée installés via Unity Package Manager

Interface utilisateur du système NeuroVision

Afin de faciliter l'interaction entre les utilisateurs et les différents modules du système NeuroVision, une interface graphique intuitive et moderne a été développée. Cette interface constitue le point d'accès principal aux fonctionnalités d'analyse, de visualisation et d'exploitation des données médicales. Elle a été conçue dans le but d'offrir une expérience utilisateur simple, fluide et adaptée aux différents profils intervenant dans le processus de diagnostic et d'aide à la décision médicale.

L'interface adopte une architecture multi-rôles permettant à chaque utilisateur d'accéder uniquement aux fonctionnalités correspondant à ses besoins. Elle repose sur une conception ergonomique intégrant des éléments visuels modernes, des tableaux de bord interactifs et des outils de visualisation adaptés au domaine de l'imagerie médicale.

L'accès au système s'effectue à travers une page d'authentification permettant de sécuriser l'utilisation de la plateforme. L'utilisateur doit renseigner son nom d'utilisateur ainsi que son mot de passe afin d'accéder aux différentes fonctionnalités proposées par NeuroVision.

La Figure 29 présente la page d'authentification développée pour le système NeuroVision.

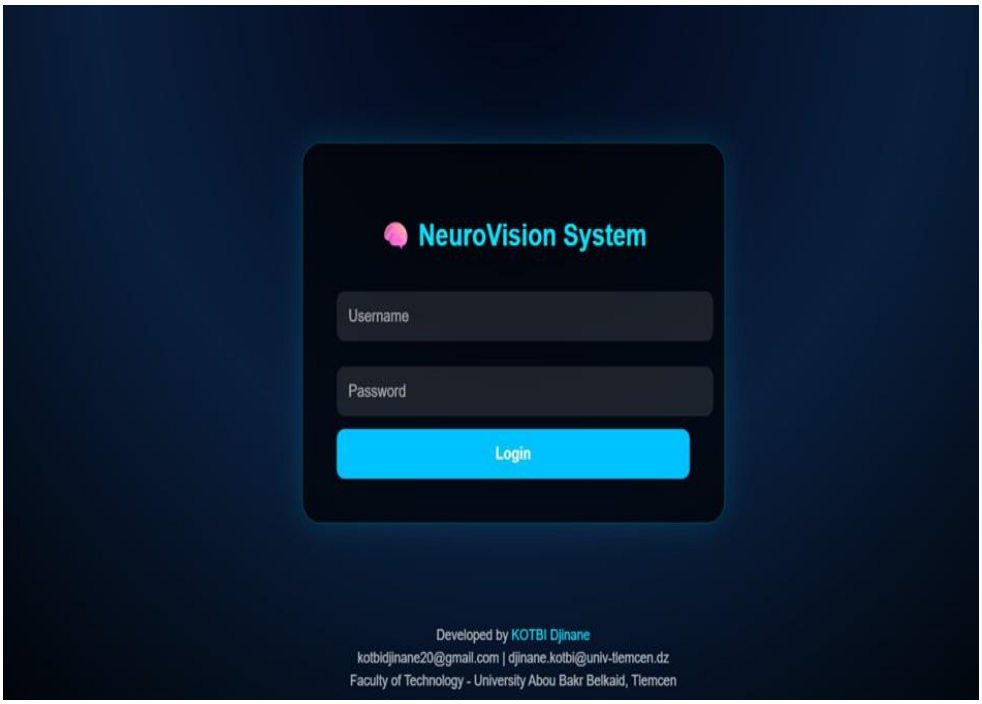


Figure 29 : Interface de connexion de la plateforme NeuroVision

Après validation des informations d'authentification, l'utilisateur est redirigé vers une interface de sélection de rôle. Cette étape permet d'orienter automatiquement l'utilisateur vers l'environnement correspondant à son domaine d'intervention. Trois profils principaux ont été définis au sein du système : le radiologue (Radiologist), le neurochirurgien (Neurosurgeon) et le chercheur (Researcher).

Cette séparation des rôles permet une meilleure organisation des fonctionnalités et garantit une utilisation adaptée aux besoins spécifiques de chaque catégorie d'utilisateur.

La Figure 30 illustre l'interface de sélection des rôles proposée par la plateforme NeuroVision.

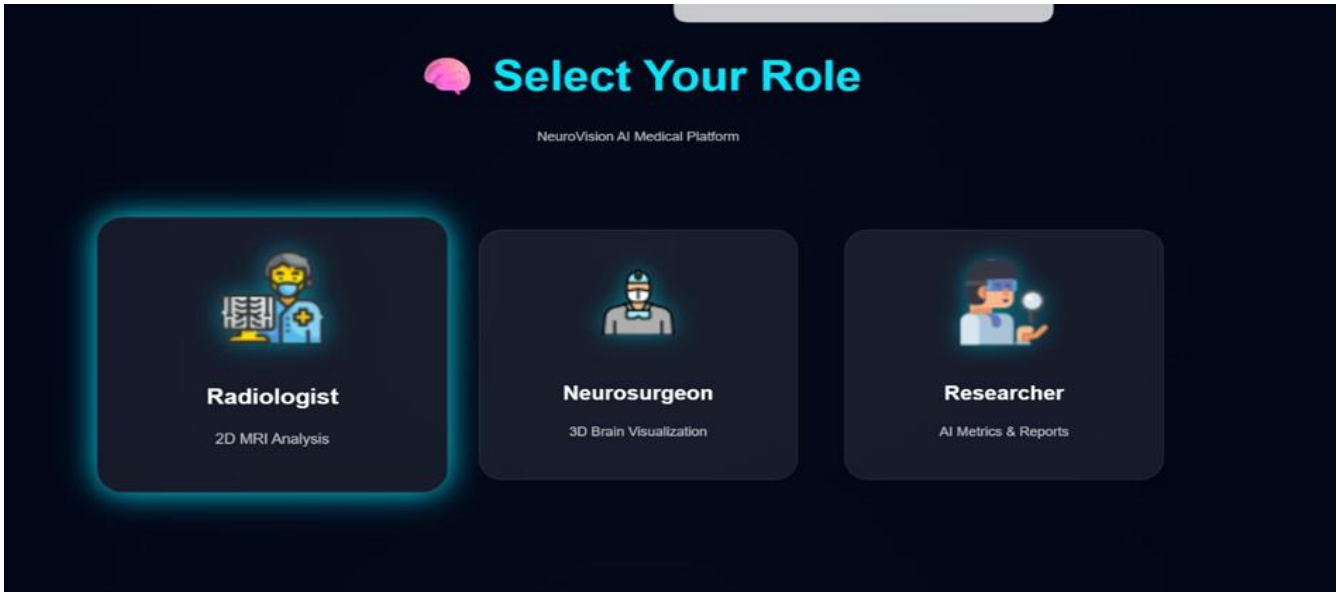


Figure 30 : Interface de sélection des rôles d'utilisateurs de NeuroVision

Interface du module Radiologist

Le module Radiologist AI Viewer a été développé afin d'assister le radiologue dans l'analyse des images IRM cérébrales et dans l'interprétation des résultats générés par les algorithmes d'intelligence artificielle. Cette interface regroupe l'ensemble des outils nécessaires à la visualisation des données médicales, à la segmentation tumorale ainsi qu'à l'évaluation quantitative des résultats.

Dans un premier temps, l'utilisateur sélectionne l'identifiant du patient à partir de la liste disponible. Après avoir choisi le patient souhaité, le bouton Load Patient permet de charger automatiquement les données médicales correspondantes ainsi que les informations nécessaires à l'analyse.

La Figure 31 présente l'interface principale du module Radiologist après le chargement d'un patient.



Figure 31 : Interface principale du module Radiologist AI Viewer

L'interface met à disposition plusieurs fonctionnalités permettant d'explorer progressivement les données médicales. Le bouton Original Image permet d'afficher l'image IRM originale utilisée comme

entrée du système. Cette étape constitue la référence visuelle sur laquelle seront réalisées les différentes analyses.

La Figure 32 illustre un exemple d'image IRM originale chargée dans l'interface.



Figure 32 : Image IRM originale du patient

Le bouton Mask réel permet d'afficher le masque de vérité terrain de la tumeur fourni dans la base de données. Ce masque représente la référence (ground truth) utilisée pour l'évaluation des performances des modèles de segmentation. La Figure 33 présente un exemple de masque réel associé à une image IRM.



Figure 33 : Masque réel (Ground Truth) de la tumeur

Le bouton Predicted Mask permet de visualiser le masque généré automatiquement par le modèle d'intelligence artificielle. Ce masque représente la région tumorale détectée par l'algorithme de segmentation.

La Figure 34 présente le résultat de segmentation obtenu par le modèle FPN, illustrant la prédiction du réseau sur l'image d'entrée.



Figure 34 : Masque prédit par le modèle de segmentation

Le bouton Contour permet de visualiser les contours de la tumeur superposés sur l'image IRM. Cette représentation met en évidence les limites de la lésion et facilite l'analyse de sa forme et de sa localisation. La Figure 36 illustre le résultat de la visualisation des contours tumoraux obtenus à partir de l'image IRM.



Figure 35 : Visualisation des contours de la tumeur superposés sur l'image IRM

Le bouton Overlay MRI permet de superposer le masque tumoral à l'image IRM originale. Cette représentation facilite l'interprétation visuelle des résultats en mettant en évidence la localisation précise de la tumeur sur les structures anatomiques du cerveau. La Figure 36 illustre le résultat de la fusion entre l'image IRM et le masque tumoral.



Figure 36 : Superposition du masque tumoral sur l'image IRM

Le bouton Analysis permet d'effectuer une évaluation quantitative de la segmentation réalisée. Plusieurs métriques sont calculées automatiquement afin de mesurer la qualité des résultats obtenus. Parmi ces indicateurs figurent notamment le coefficient de Dice, l'indice IoU, la distance HD95 ainsi que différentes mesures géométriques liées à la tumeur.

La Figure 37 présente un exemple des résultats d'analyse générés par le système.

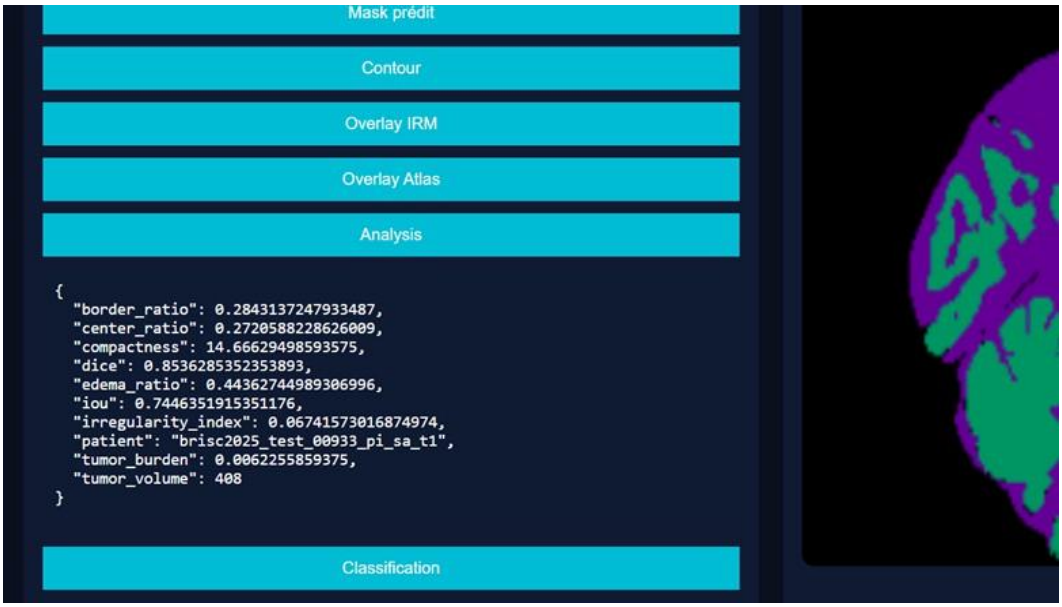


Figure 37 : Résultats de l'analyse quantitative de la segmentation

Les résultats d'analyse obtenus montrent un coefficient de Dice de 0,851 et un indice IoU de 0,741, traduisant une bonne qualité de segmentation malgré quelques différences entre le masque prédit et la vérité terrain. Le volume tumoral estimé est de 354 voxels, correspondant à une charge tumorale d'environ 0,54 % du volume total étudié. Les paramètres morphologiques calculés, notamment la compacité (13,72) et l'indice d'irrégularité (0,034), indiquent une tumeur présentant une forme relativement régulière. Par ailleurs, les ratios de localisation montrent une répartition de la région tumorale entre les zones centrales et périphériques de l'image, accompagnée d'une proportion notable d'œdème autour de la lésion.

Le bouton Atlas Overlay permet de superposer la tumeur segmentée sur un atlas cérébral, tout en affichant les différentes sous-classes de la lésion. Cette visualisation permet de localiser précisément

les régions tumorales dans un référentiel anatomique standard et d'identifier la répartition des sous-régions. La Figure 38 illustre la superposition de la tumeur avec les sous-classes sur l'atlas cérébral.

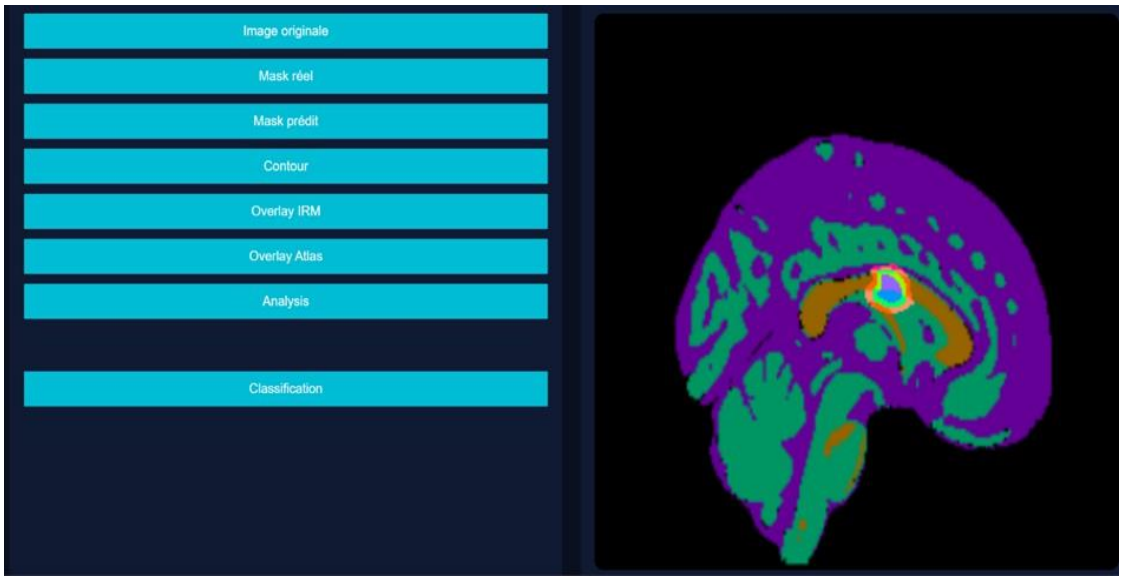


Figure 38 : Superposition de la tumeur et de ses sous-classes sur un atlas cérébral

Enfin, le bouton Classification permet de déterminer automatiquement la catégorie de la tumeur à partir des informations extraites de l'image médicale. Le système fournit la classe prédite ainsi que le niveau de confiance associé à cette prédiction.

La Figure 39 présente un exemple de résultat obtenu après l'étape de classification.

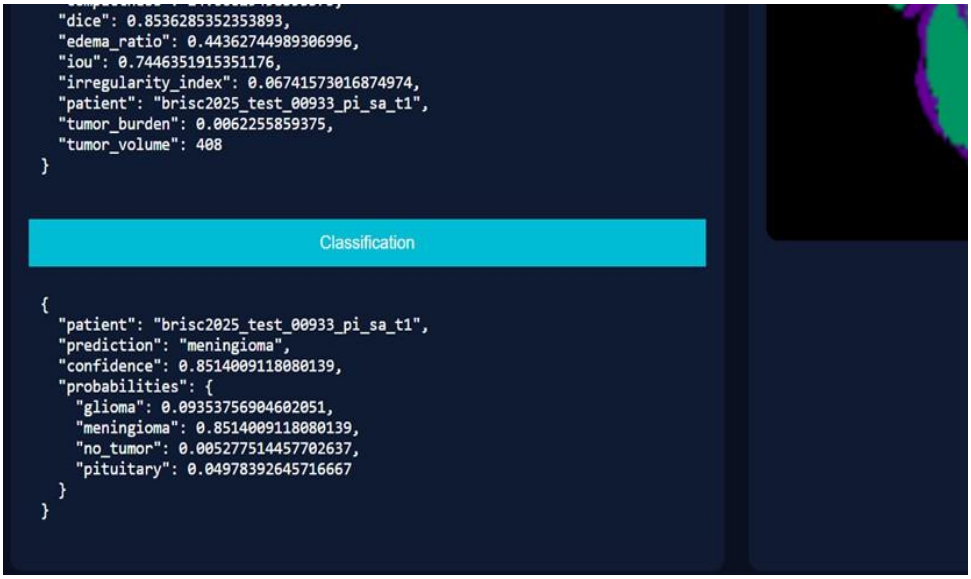


Figure 39 : Résultat de classification de la tumeur cérébrale

Les résultats de classification indiquent que la tumeur est identifiée comme un méningiome (meningioma) avec un niveau de confiance élevé de 85,14 %. Les probabilités associées aux autres classes restent nettement inférieures, avec 9,35 % pour le gliome, 4,98 % pour la tumeur pituitaire et seulement 0,53 % pour l'absence de tumeur. Cette distribution des probabilités confirme la forte certitude du modèle quant à l'appartenance de l'image à la classe méningiome.

Interface du module Neurosurgeon

Le module Neurosurgeon a été développé afin de fournir une visualisation avancée des données médicales destinées à l'assistance préopératoire et à l'analyse tridimensionnelle des tumeurs cérébrales. Contrairement au module Radiologist, ce module repose directement sur le chargement des volumes médicaux au format NIfTI (.nii ou .nii.gz).

Dans un premier temps, l'utilisateur sélectionne le fichier médical à analyser à l'aide du bouton de chargement. Après la sélection du volume IRM, le système charge automatiquement les données afin de préparer les différentes étapes de traitement.

La Figure 40 présente l'interface du module Neurosurgeon après le chargement d'un fichier médical.

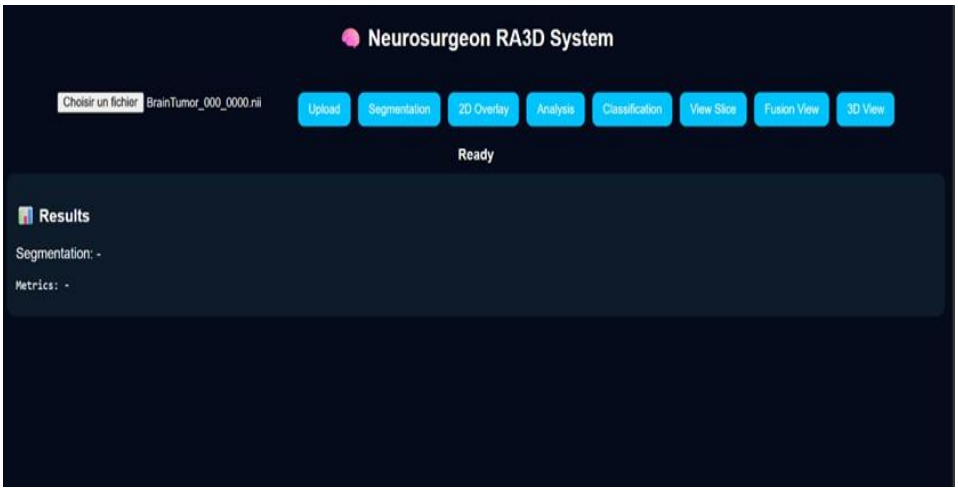


Figure 40 : Chargement d'un volume médical dans le module Neurosurgeon

Une fois les données chargées, l'utilisateur peut lancer l'étape de segmentation tumorale. Le système génère alors automatiquement un masque tumoral qui est enregistré sous forme de fichier NIfTI. Le bouton Open File permet d'accéder directement au fichier de segmentation produit par le système.

Afin de faciliter l'interprétation visuelle des résultats, le système propose également une représentation sous forme d'overlay 2D. Cette visualisation superpose la région tumorale détectée à l'image IRM correspondante afin de mettre en évidence la localisation anatomique de la lésion.

La Figure 41 présente un exemple de superposition entre l'image IRM et le masque tumoral.

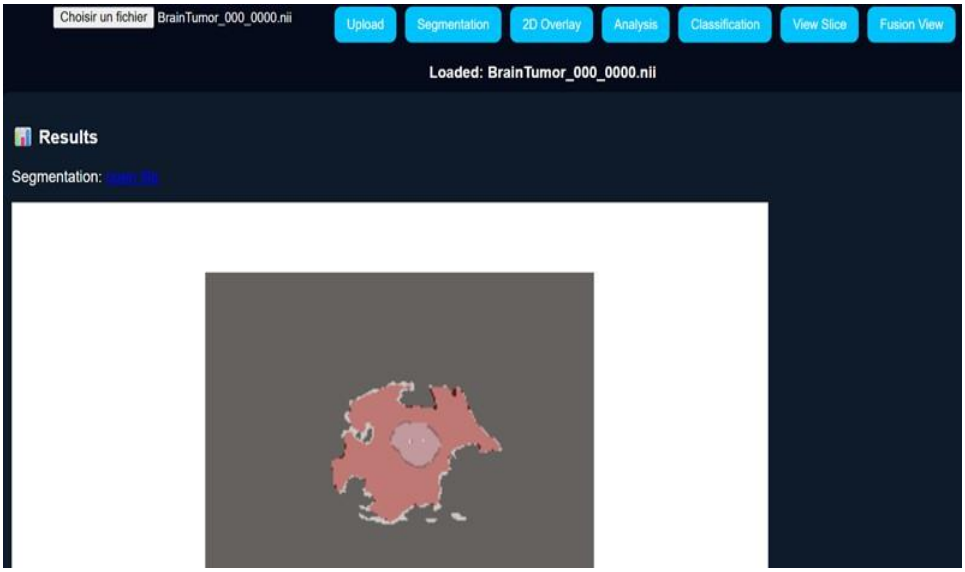


Figure 41 : Visualisation 2D de l'overlay tumoral

Le module permet ensuite d'afficher les résultats d'analyse quantitative générés automatiquement à partir du masque segmenté. Ces résultats regroupent les différents paramètres d'évaluation et les caractéristiques morphologiques de la tumeur.

La Figure 42 présente la fenêtre des résultats d'analyse obtenus après la segmentation.



Figure 42 : Résultats d'analyse du module Neurosurgeon

Enfin, le système réalise une classification automatique de la tumeur afin d'identifier la catégorie la plus probable parmi les classes étudiées. Le résultat de prédiction est affiché directement dans l'interface avec le niveau de confiance associé.

La Figure 43 présente un exemple de résultat obtenu après l'étape de classification.

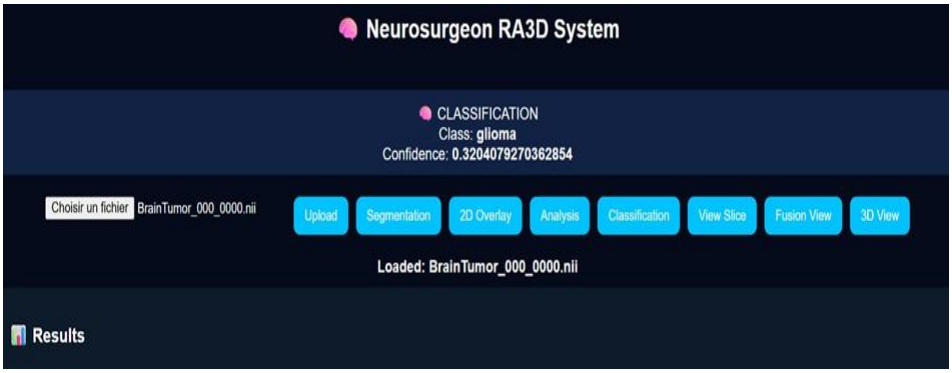


Figure 43 : Résultat de classification de la tumeur

Visualisation des coupes et rendu 3D (Neurosurgeon)

Après l'étape de segmentation et d'analyse, le module Neurosurgeon propose des outils de visualisation avancée permettant une meilleure interprétation des résultats.

Dans un premier temps, la fonctionnalité View Slice permet d'afficher une coupe 2D spécifique du volume IRM du patient. Cette coupe correspond à une tranche représentative de l'examen médical et permet d'observer la structure anatomique ainsi que la région suspecte de manière plus détaillée.

La Figure 44 présente un exemple de coupe IRM affichée via la fonction View Slice.



Figure 44 : Visualisation d'une coupe IRM (View Slice)

Ensuite, le système propose un module de visualisation 3D basé sur Unity, permettant une représentation volumique de l'IRM ainsi que de la tumeur segmentée. Cette étape constitue une extension vers une visualisation immersive proche de la réalité augmentée, bien que le système actuel repose principalement sur un rendu 3D interactif.

Ce module permet de manipuler le modèle de manière intuitive, incluant des fonctionnalités de rotation, zoom et déplacement. L'utilisateur peut ainsi explorer la structure cérébrale sous différents angles afin de mieux comprendre la localisation et l'étendue de la tumeur.

De plus, l'interface 3D intègre des commandes de navigation telles que Next et Previous, permettant de parcourir les différents patients ou cas médicaux sans quitter l'environnement de visualisation.

La Figure 45 illustre l'interface de visualisation 3D sous Unity avec la représentation de la tumeur cérébrale.

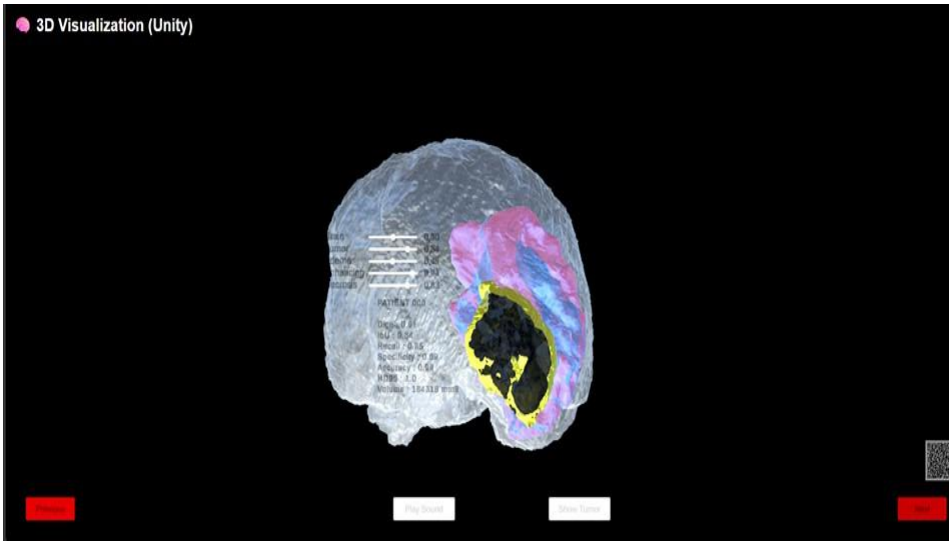


Figure 45 : Visualisation 3D interactive de l'IRM et de la tumeur sous Unity

Ce module offre ainsi une approche interactive et intuitive de l'analyse médicale, facilitant la compréhension des résultats de segmentation et améliorant l'expérience utilisateur pour les spécialistes médicaux.

Interface du module Researcher

Le module Researcher a été conçu afin de permettre l'exploration et la consultation centralisée des données patients. Contrairement aux autres modules orientés traitement et analyse, cette interface est principalement dédiée à la visualisation et à la recherche clinique.

Recherche et profil du patient

Dans cette première étape, l'utilisateur peut effectuer une recherche de patient à partir de son identifiant (Patient ID). Une fois l'identifiant saisi, le système affiche automatiquement le profil complet du patient.

Ce profil comprend les informations essentielles suivantes :

- Identifiant du patient
- Nom et prénom
- Modalité d'examen (MRI Brain)
- Image d'identité du patient

La Figure 46 présente l'interface de recherche et d'affichage du profil patient.

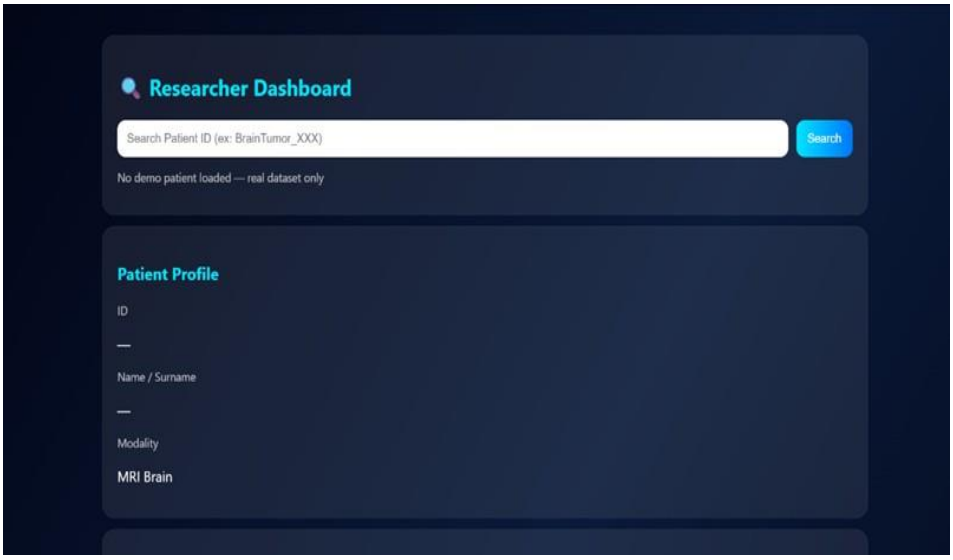


Figure 46 : Interface de recherche et profil patient

Contexte clinique

Après sélection du patient, le système affiche une section dédiée au contexte clinique. Cette partie permet une visualisation structurée des informations médicales sous forme de champs organisés.

Le contexte clinique comprend :

- Raison de l'examen (Why exam)
- Symptômes
- Hypothèses médicales
- Décision clinique finale

Ces informations peuvent être présentées sous forme de zones de texte ou de barres de progression afin de faciliter la lecture et la compréhension rapide du dossier médical.

La Figure 47 illustre la section du contexte clinique du patient.

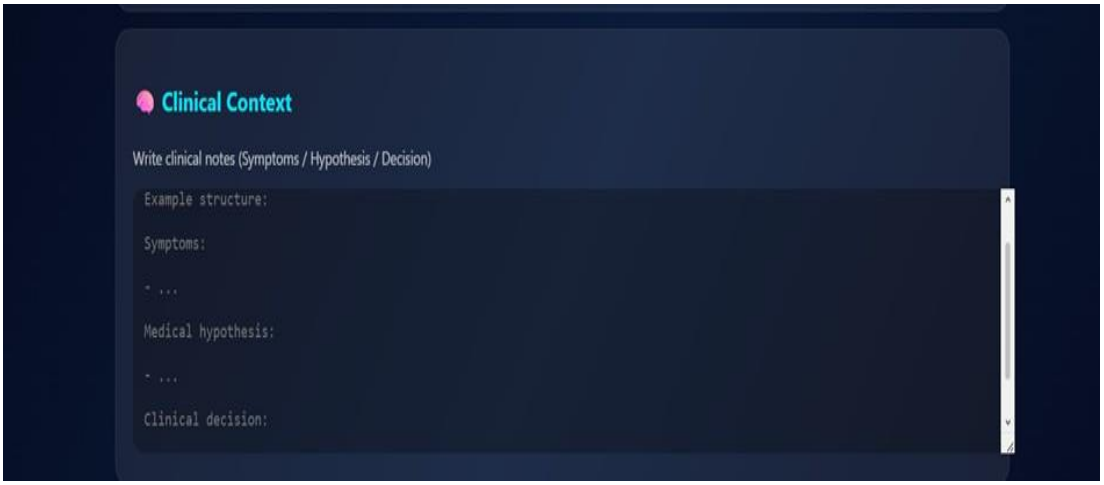


Figure 47 : Affichage du contexte clinique structuré

Dossiers médicaux et visualisation

La dernière section du module Researcher regroupe l'ensemble des fichiers médicaux associés au patient. Cette partie permet d'accéder rapidement aux différents documents et résultats d'imagerie.

Elle comprend :

- Rapport médical au format PDF
- Rapport radiologique
- Résultats de segmentation
- Images IRM

Un bouton de type Upload New File permet également d'ajouter de nouveaux documents au dossier patient, assurant ainsi une mise à jour dynamique des données.

Enfin, une section de Visual Preview permet d'afficher les images IRM et les résultats de traitement associés au patient sélectionné.

La Figure 48 présente les fichiers médicaux ainsi que la prévisualisation des images.

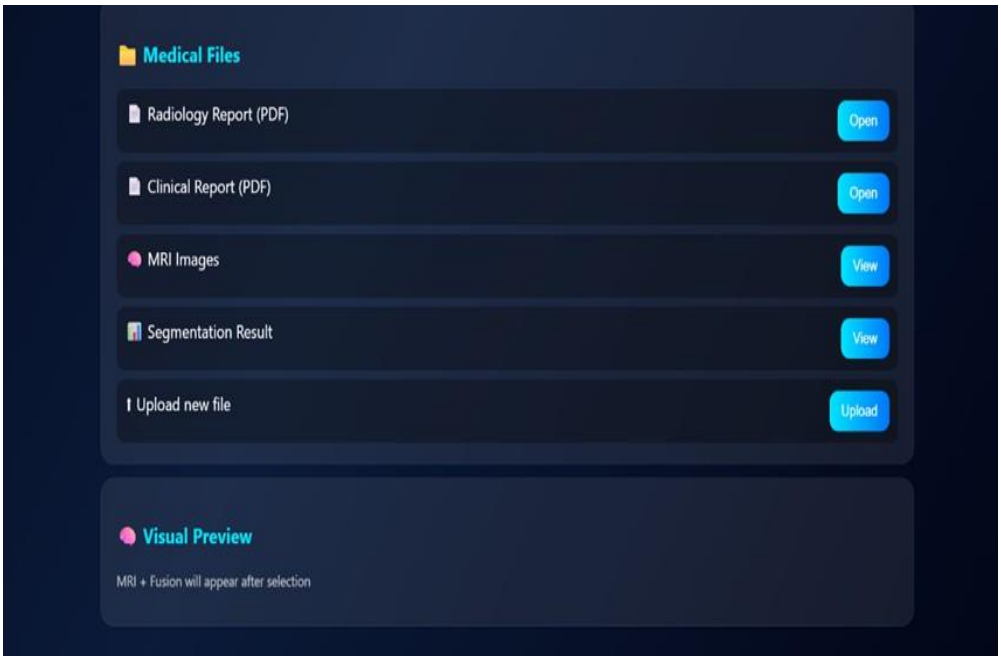


Figure 48 : Dossiers médicaux et visualisation des résultats

Business Model Canevas (BMC) - NeuroVision

1. Proposition de valeur (Value Proposition)

a. Quels problèmes résolvons-nous pour nos clients ?

NeuroVision résout cinq problèmes majeurs des hôpitaux algériens.

Problème 1 : segmentation manuelle longue. Actuellement, un radiologue passe entre 45 et 60 minutes à délimiter manuellement une tumeur sur IRM. NeuroVision vise à automatiser cette tâche en quelques secondes (objectif de réduction du temps de traitement).

Problème 2 : taux d'erreur diagnostique. L'interprétation humaine des IRM comporte un certain taux d'erreur. Sur nos jeux de test, notre IA atteint une précision de 99,5% pour la classification des 4 types de tumeurs (glioma, méningiome, tumeur pituitaire, absence de tumeur), ce qui suggère une amélioration potentielle significative.

Problème 3 : absence de visualisation 3D. La plupart des CHU algériens ne disposent pas d'outils de visualisation 3D. NeuroVision offre la réalité augmentée 3D sous Unity.

Problème 4 : coût élevé des solutions étrangères. Les solutions industrielles (Siemens, GE, IBM) sont proposées à des tarifs très élevés. NeuroVision propose une alternative locale à coût significativement réduit (environ 60% à 85% moins cher).

Problème 5 : dépendance technologique étrangère. NeuroVision est développée localement avec possibilité d'hébergement sur le territoire national, garantissant la souveraineté des données.

b. Quels besoins de nos clients satisfont nos produits ou services ?

NeuroVision répond à sept besoins fondamentaux :

- **Rapidité** : segmentation accélérée (objectif de quelques secondes)
- **Précision** : Dice = 0,8895 (FPN 2D - base BRISC), Dice = 0,8436 (nnU-Net 3D - base BraTS2020), Accuracy = 99,5%
- **Fiabilité** : validation sur BraTS2020 (369 patients) et base BRISC (5 995 images)
- **Interprétabilité** : overlay IRM/masque, contours, atlas MNI, courbes ROC
- **Immersion** : visualisation 3D Unity (rotation, zoom, transparence)
- **Formation** : comparaison GT vs prédiction
- **Recherche** : export automatique des métriques (Dice, IoU, HD95, volumes) au format JSON

c. En quoi notre offre est-elle différente de celle de nos concurrents ?

Concurrent	Prix (estimatif)	Différence NeuroVision
Siemens syngo.via	Plusieurs millions DA	Solution adaptée aux besoins locaux avec coût inférieur + classification IA
GE Healthcare	Plusieurs millions DA	Solution locale + support français
IBM Watson	Plusieurs millions DA	Souveraineté des données
3D Slicer (gratuit)	0	Segmentation automatique + classification IA

Notre interface multi-rôles (Radiologue, Neurochirurgien, Chercheur) est un atout différenciant.

Note : Les prix présentés sont des estimations à caractère indicatif. Aucun prix officiel public n'est disponible [73]. Données non publiées, basées sur des retours d'expérience d'établissements de santé (2024-2025).[Accès le 16 juin 2026]

d. Quelle est notre proposition unique de valeur ?

NeuroVision : Solution algérienne d'aide au diagnostic des tumeurs cérébrales

Dice = 0,8895 (2D) | Dice = 0,8436 (3D) | Accuracy = 99,5%
Prix : 1,5 à 4 millions DA
1ère plateforme algérienne combinant segmentation 3D, classification IA et réalité augmentée

2. Segments de clients (Customer Segment)

a. Quels sont nos clients principaux ?

Nos clients principaux sont les établissements de santé algériens :

- **25 CHU** (Centres Hospitalo-Universitaires) : cible prioritaire
- **40 hôpitaux spécialisés** : cancérologie, neurologie
- **13 facultés de médecine** : formation des étudiants
- **8 centres de recherche** : CRBt, CRT, CERIST, etc.
- **50 cliniques privées** : imagerie médicale

Marché potentiel estimé : environ 136 établissements.

3. Relation avec les clients (Consumer Relationships)

a. Quel type de relation chaque segment attend-il de nous ?

- **Radiologues** : support technique réactif (moins de 24h)
- **Neurochirurgiens** : accompagnement personnalisé en bloc opératoire
- **Chercheurs** : accès aux données brutes, documentation scientifique
- **Directions hospitalières** : rapports de performance, analyses de ROI

4. Canaux de distribution (Channels)

Canaux principaux :

- Visites terrain et démonstrations dans les CHU
- Site web et présence sur LinkedIn
- Participation aux conférences médicales nationales
- Partenariat avec des distributeurs locaux

5. Partenaires clés (Key Partnerships)

Partenaires identifiés :

- CHU Tlemcen : validation clinique, données réelles
- Faculté de Médecine : expertise médicale
- ANADEB / ANDPME : soutien financier à l'innovation
- Incubateur UABT : accompagnement, bureaux
- Kaggle/Google Colab : ressources GPU pour prototypage

6. Activités clés (Key Activities)

Activités principales :

- Développement et amélioration continue des modèles IA
- Entraînement et validation sur bases de données médicales
- Développement de l'interface utilisateur (Unity, web)
- Tests cliniques et recueil des retours utilisateurs
- Formation et support aux clients
- Marketing et prospection commerciale

7. Ressources clés (Key Resources)

Ressources essentielles :

- Ordinateurs performants (GPU)
- Bases de données médicales (BraTS, BRISC)
- Modèles IA entraînés (nnU-Net, FPN, EfficientNet)
- Compétences en IA, traitement d'image et réalité augmentée
- Outils open source (Python, PyTorch, Unity, 3D Slicer)

8. Charges et coûts (Cost Structure) - Estimations prévisionnelles

Poste de coût	Montant (millions DA)
Développement (2 développeurs, 6 mois)	0,6
Équipement (GPU, PC, smartphone AR)	0,2
Marketing et prospection (annuel)	0,5
Formation des utilisateurs (annuel)	0,15
Support technique (annuel)	0,25
Hébergement cloud (annuel)	0,2
Total première année (estimation)	**1,9**

Ces chiffres sont des estimations prévisionnelles pouvant être ajustées selon les conditions réelles.

9. Revenus (Revenue) - Estimations prévisionnelles

Modèle de revenus

Licence	Contenu	Prix (millions
Standard	Segmentation 2D + Classification + Interface Radiologue	1,5
Pro	Standard + Visualisation 2D (atlas, overlay)	2
Enterprise	Pro + Segmentation 3D + RA + Interfaces Neurochirurgien/Chercheur	4

Autres sources de revenus :

- Formation à la carte : 600 DA
- Maintenance (15% du prix licence) : 2250 à 6000 DA
- Intégration et déploiement : 1000 DA

Projection prévisionnelle (hypothèse prudente)

Année	Clients estimés	Revenus (millions DA)	Charges (millions DA)
Année 1	3-5	4,5 - 10	1,9
Année 2	5-10	10 - 20	2,5
Année 3	10-15	20 - 30	3,0

Ces projections sont basées sur des hypothèses de pénétration du marché et doivent être considérées comme des estimations prévisionnelles. Le seuil de rentabilité est envisagé entre 12 et 24 mois selon l'adoption réelle.

Investissement initial (prévisionnel)

Source	Montant (millions DA)
ANADEB (soutien innovation)	0,8
ANDPME (accompagnement)	0,5
Incubateur UABT	0,2
Apports personnels	0,3
Total	**1,8**

10. Synthèse du Business Model Canevas

Bloc	Résumé
Proposition valeur	Segmentation automatique, classification IA, visualisation 3D, prix adapté
Segments clients	25 CHU, 40 hôpitaux, 13 facultés, 8 centres recherche, 50 cliniques
Relation client	Support technique, formation, hotline, newsletter
Canaux	Visites terrain, site web, conférences, LinkedIn
Partenaires	CHU Tlemcen, Faculté, ANADEB, ANDPME, incubateur
Activités clés	Développement IA, entraînement, tests, formation, support
Ressources clés	GPU, bases données, modèles IA, compétences
Coûts (prévisionnel)	Développement 0,6M, marketing 0,5M, support 0,25M
Revenus (prévisionnel)	Standard 1,5M, Pro 2M, Enterprise 4M DA/an

NeuroVision : Solution algérienne innovante pour le diagnostic des tumeurs cérébrales

Les chiffres présentés sont des estimations prévisionnelles basées sur des hypothèses de marché à valider.

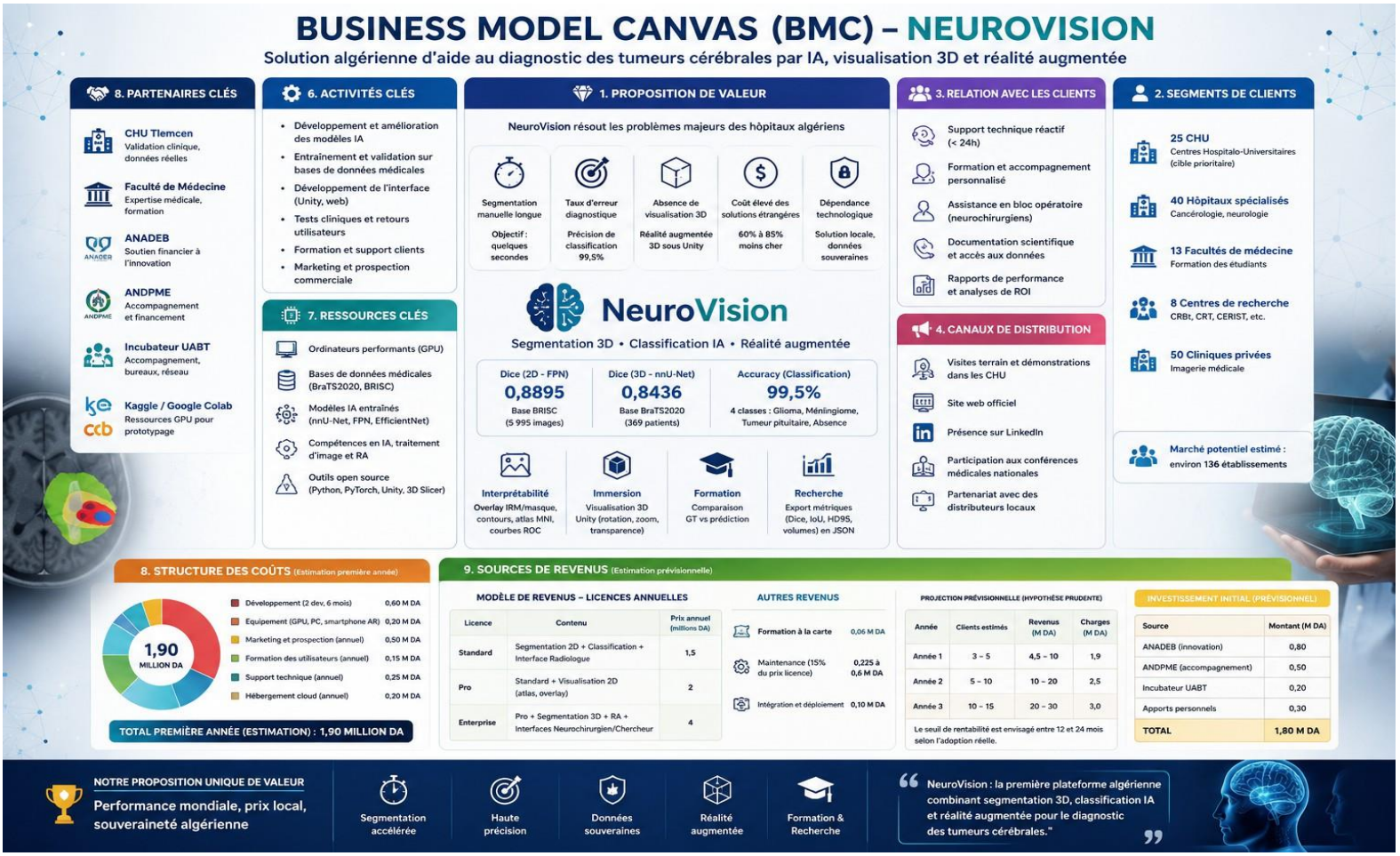


Figure 49 : Business Model Canvas de NeuroVision présentant la proposition de valeur, les segments de clients, les partenaires clés, les ressources, les coûts et les sources de revenus de la plateforme.

Perspectives et limites

Limites du système actuel

Malgré les performances encourageantes de NeuroVision, plusieurs limites doivent être soulignées.

Limites techniques :

- La segmentation 3D avec nnU-Net (Dice moyen = 0,8436) reste perfectible, notamment pour les tumeurs de très petite taille ou présentant des contours flous.
- La classification atteint une accuracy de 99,5%, mais des erreurs persistent principalement entre les classes glioma et méningiome en raison de similarités morphologiques.
- Le temps de reconstruction 3D (environ 3 secondes par patient) peut être amélioré pour un usage en temps réel.
- La dépendance aux GPU pour l'entraînement limite le déploiement dans des hôpitaux peu équipés.

Limites liées aux données :

- La base BRISC (5 995 images) reste limitée en taille par rapport aux standards internationaux.
- Les données proviennent principalement de sources publiques (BraTS2020), ce qui peut induire un biais par rapport aux cas réels algériens.
- L'absence de données longitudinales ne permet pas d'évaluer la progression tumorale dans le temps.

Limites cliniques :

- Le système n'a pas encore été validé sur un large éventail de cas cliniques réels en conditions opérationnelles.
- L'interface utilisateur, bien que conçue pour être intuitive, nécessite une formation préalable.
- La réalité augmentée 3D dépend de la qualité du tracking et de la calibration.

Perspectives d'amélioration

Cour terme (6-12 mois) :

- Améliorer les modèles de segmentation (Dice > 0,90) par fine-tuning sur des données cliniques algériennes.
- Étendre la base BRISC à 10 000 images pour renforcer la robustesse.

- Développer une version légère pour appareils mobiles (traitement local sans GPU).
- Réaliser une étude clinique pilote au CHU Tlemcen (50 patients).

Moyen terme (1-2 ans) :

- Intégrer la détection automatique des structures critiques (vaisseaux, aires fonctionnelles).
- Développer un module de suivi longitudinal des patients (évolution tumorale dans le temps).
- Ajouter la prédiction du grade tumoral (bas grade vs haut grade).
- Créer une API REST pour intégration avec les systèmes PACS hospitaliers.

Long terme (2-5 ans) :

- Extension à d'autres types de tumeurs (poumon, foie, prostate) par transfer learning.
- Développement d'une solution complète en réalité augmentée pour guidage peropératoire.
- Certification médicale (CE, FDA) pour une commercialisation à l'international.
- Création d'une plateforme cloud collaborative pour partage anonyme de données.

Perspectives commerciales

- Déploiement progressif dans les CHU d'Algérie (Tlemcen, Oran, Alger en priorité).
- Partenariat avec le Ministère de la Santé pour une adoption nationale.
- Formation des équipes médicales (sessions de 2 jours, 15 à 20 médecins par session).
- Développement d'une offre SaaS (Software as a Service) pour les cliniques privées.
- Export vers les pays du Maghreb (Tunisie, Maroc) à moyen terme.

Conclusion

NeuroVision constitue une première étape prometteuse vers l'intégration de l'intelligence artificielle et de la réalité augmentée dans le diagnostic des tumeurs cérébrales en Algérie. Les limites identifiées guideront les travaux futurs, tandis que les perspectives montrent le potentiel d'évolution du système vers une solution clinique complète, validée et commercialisable.

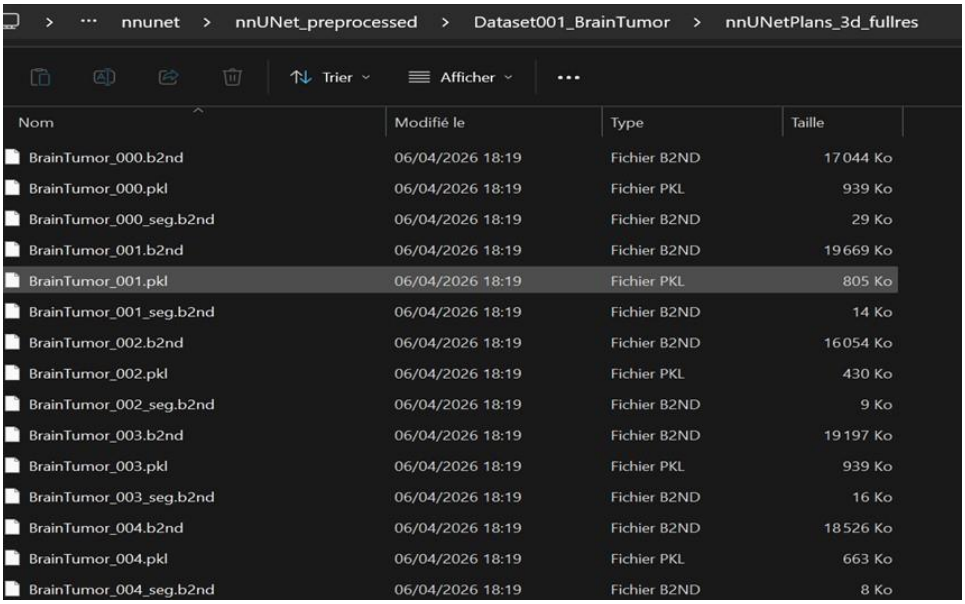


Figure 50 : Exemple des fichiers générés après le prétraitement automatique pour un patient du dataset

- BrainTumor-000.b2nd : contient les volumes IRM prétraités utilisés comme entrée du réseau de neurones après normalisation et préparation automatique des données.

- BrainTumor-000-seg.b2nd : correspond aux masques de segmentation associés aux volumes IRM prétraités. Ces fichiers sont utilisés comme vérité terrain durant l'apprentissage supervisé.

- BrainTumor-000.pkl : contient les métadonnées du patient ainsi que les informations nécessaires à la reconstruction et au traitement des volumes, notamment les paramètres de prétraitement et certaines caractéristiques spatiales.

Cette organisation permet à nnU-Net de gérer automatiquement les données médicales volumétriques tout en assurant la reproductibilité du pipeline de segmentation.

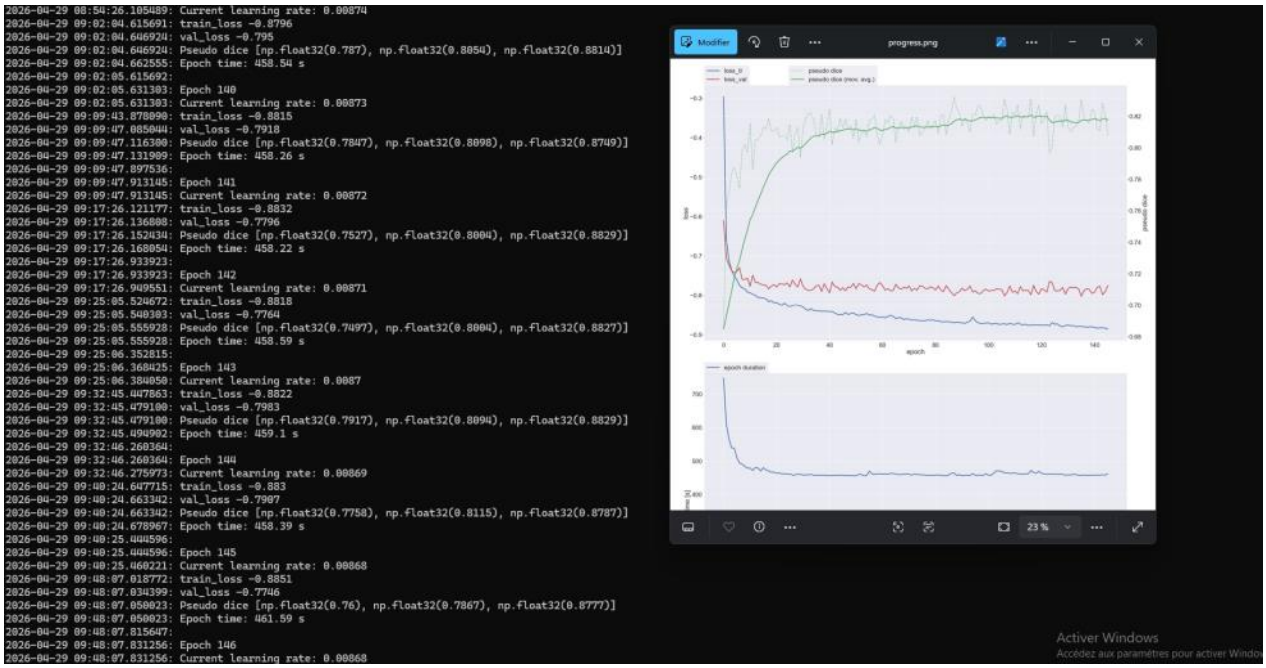


Figure 51 : configuration 2D - fold 3



Figure 52 : configuration 2D - fold 4

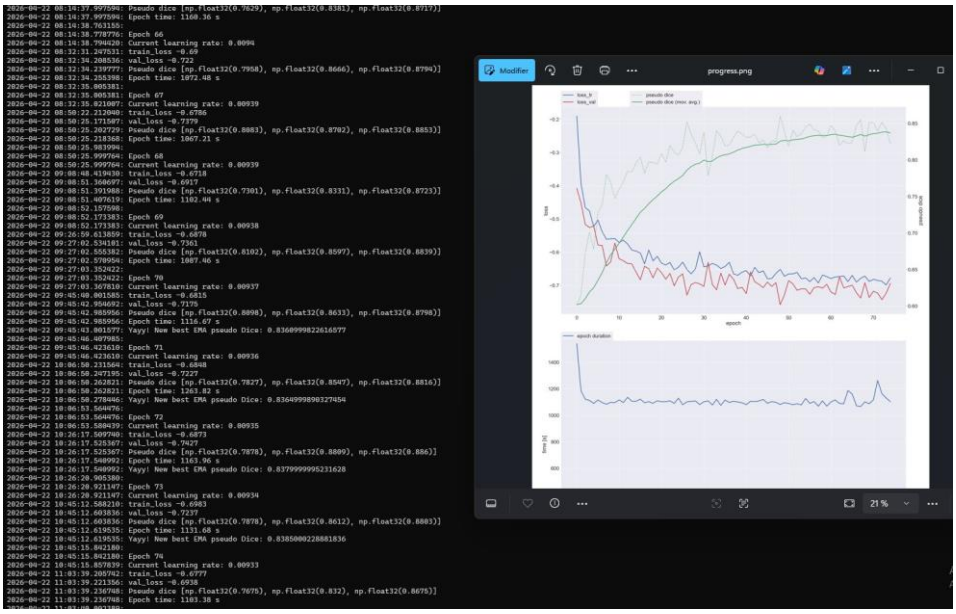


Figure 53 : configuration 3D full resolution - fold 0

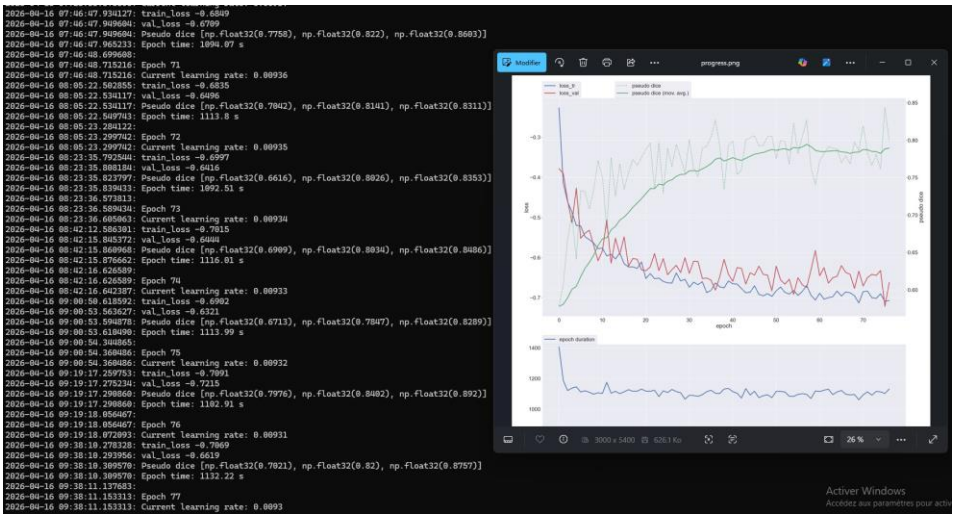


Figure 54 : configuration 3D full resolution - fold 3



Figure 55 : La configuration détaillée des paramètres d'entraînement

Difficultés expérimentales et contraintes matérielles

L'entraînement des modèles de segmentation et de classification a été confronté à plusieurs défis :

1. Contraintes électriques : Des coupures d'électricité fréquentes et imprévisibles au niveau du laboratoire ont interrompu à plusieurs reprises les sessions d'entraînement. Malgré l'utilisation de sauvegardes périodiques (checkpoints), ces interruptions ont limité le nombre d'époques réalisables (environ 70 par fold contre 1000 dans l'état de l'art).

2. Ressources computationnelles limitées : Le laboratoire ne dispose que d'un seul GPU (NVIDIA GeForce GTX 1650 Ti avec 4 Go de VRAM), insuffisant pour entraîner des modèles 3D volumineux comme nnU-Net en configuration full resolution. Un recours aux plateformes cloud (Kaggle, Google Colab) a été nécessaire, avec des limitations de temps (30h/semaine sur Kaggle, sessions limitées sur Colab).

3. Validation croisée 5-fold : Cette méthode robuste divise le dataset en 5 parties (80% entraînement, 20% validation). Le temps d'entraînement est multiplié par 5, ce qui a considérablement allongé la durée totale des expériences (environ 350 époques cumulées par fold).

4. Manque de stockage : Les fichiers générés par nnU-Net (checkpoints, prétraitement, résultats) occupent plusieurs centaines de Go, dépassant la capacité du disque dur local (512 Go). Une gestion rigoureuse des espaces a été nécessaire.

5. Connexion internet instable : Les interruptions de connexion ont perturbé les transferts de données vers les plateformes cloud et l'accès aux bases de données publiques (BraTS, BRISC).

6. Temps de calcul excessif : La configuration 3D full resolution de nnU-Net nécessite environ 48h par fold sur GPU cloud (Tesla P100). Avec 5 folds et les interruptions, le temps total a dépassé 3 semaines de calcul continu.

7. Absence de support technique : Aucun support informatique dédié n'était disponible au laboratoire pour résoudre rapidement les problèmes matériels ou logiciels.

8. Problèmes d'installation des dépendances : nnU-Net v2 et ses dépendances (PyTorch, MONAI, CUDA) ont nécessité de nombreux ajustements pour être compatibles avec les versions limitées des GPU disponibles.

9. Manque de données annotées : L'annotation manuelle des masques de segmentation est longue et fastidieuse. La base BRISC (5 995 images) a demandé plusieurs mois de travail collaboratif.

10. Contraintes budgétaires : L'absence de budget dédié a limité l'accès aux ressources cloud payantes (GPU haut de gamme, stockage supplémentaire, licences logicielles).

Bibliographie

[1] Litjens, G., et al. (2017). A Survey on Deep Learning in Medical Image Analysis. Medical Image Analysis, 42, 60-88. [Accès le 03 février 2026]

[2] LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep Learning. Nature, 521(7553), 436-444. [Accès le 05 février 2026]

[3] Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep Learning. MIT Press. [Accès le 08 février 2026]

[4] McRobbie, D. W., et al. (2017). MRI from Picture to Proton. Cambridge University Press. [Accès le 12 février 2026]

[5] Westbrook, C. (2019). MRI in Practice. Wiley-Blackwell. [Accès le 15 février 2026]

[6] Scarpace, L., et al. (2022). Brain MRI Sequences in Tumor Analysis. Radiology Clinics. [Accès le 18 février 2026]

[7] Nolte, J. (2020). The Human Brain : An Introduction to its Functional Anatomy. Elsevier. [Accès le 20 février 2026]

[8] WHO. (2023). Brain Tumors Fact Sheet. https ://www.who.int [Accès le 22 février 2026]

[9] Louis, D. N., et al. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System. Neuro-Oncology, 23(8), 1231-1251. [Accès le 25 février 2026]

[10] Louis, D. N., et al. (2021). Meningioma Pathology. In WHO Classification of CNS Tumors. [Accès le 28 février 2026]

[11] Melmed, S. (2020). Pituitary Disorders. Springer. [Accès le 02 mars 2026]

[12] Pham, D. L., Xu, C., Prince, J. L. (2000). Current Methods in Medical Image Segmentation. Annual Review of Biomedical Engineering, 2, 315-337. [Accès le 05 mars 2026]

[13] Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, 9(1), 62-66. [Accès le 08 mars 2026]

[14] Adams, R., Bischof, L. (1994). Seeded Region Growing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(6), 641-647. [Accès le 10 mars 2026]

[15] Beucher, S., Lantuéjoul, C. (1979). Use of Watersheds in Contour Detection. International Workshop on Image Processing. [Accès le 12 mars 2026]

[16] Kass, M., Witkin, A., Terzopoulos, D. (1988). Snakes : Active Contour Models. International Journal of Computer Vision, 1(4), 321-331. [Accès le 15 mars 2026]

[17] Bezdek, J. C. (1984). Fuzzy C-Means Clustering. In Pattern Recognition with Fuzzy Objective Function Algorithms. [Accès le 18 mars 2026]

[18] Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. NIPS, 25, 1097-1105. [Accès le 20 mars 2026]

[19] Pereira, S., et al. (2016). Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. IEEE Transactions on Medical Imaging, 35(5), 1240-1251. [Accès le 22 mars 2026]

[20] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net : Convolutional Networks for Biomedical Image Segmentation. MICCAI, 9351, 234-241. [Accès le 25 mars 2026]

[21] Çiçek, Ö., et al. (2016). 3D U-Net : Learning Dense Volumetric Segmentation from Sparse Annotation. MICCAI, 9901, 424-432. [Accès le 28 mars 2026]

[22] Zhou, Z., et al. (2018). UNet++ : A Nested U-Net Architecture for Medical Image Segmentation. MICCAI, 3-11. [Accès le 30 mars 2026]

[23] Lin, T. Y., et al. (2017). Feature Pyramid Networks for Object Detection. CVPR, 2117-2125. [Accès le 02 avril 2026]

[24] Chen, L. C., et al. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. ECCV, 801-818. [Accès le 05 avril 2026]

[25] Isensee, F., et al. (2021). nnU-Net : a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods, 18(2), 203-211. [Accès le 08 avril 2026]

[26] Menze, B. H., et al. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Transactions on Medical Imaging, 34(10), 1993-2024. [Accès le 10 avril 2026]

[27] Bakas, S., et al. (2018). Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation. IEEE Transactions on Medical Imaging, 37(11), 2522-2534. [Accès le 12 avril 2026]

[28] Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species. Ecology, 26(3), 297-302. [Accès le 15 avril 2026]

[29] Jaccard, P. (1912). The Distribution of the Flora in the Alpine Zone. New Phytologist, 11(2), 37-50. [Accès le 18 avril 2026]

[30] Lin, T. Y., et al. (2017). Focal Loss for Dense Object Detection. ICCV, 2980-2988. [Accès le 20 avril 2026]

[31] Nyúl, L. G., Udupa, J. K. (1999). On Standardizing the MR Image Intensity Scale. IEEE Transactions on Medical Imaging, 18(9), 815-824. [Accès le 22 avril 2026]

[32] Perona, P., Malik, J. (1990). Anisotropic Diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7), 629-639. [Accès le 25 avril 2026]

[33] Buades, A., Coll, B., Morel, J. M. (2005). Non-Local Means Denoising. CVPR, 2, 60-65. [Accès le 28 avril 2026]

[34] Tomasi, C., Manduchi, R. (1998). Bilateral Filtering. ICCV, 839-846. [Accès le 30 avril 2026]

[35] Tustison, N. J., et al. (2010). N4ITK : Improved N3 Bias Correction. IEEE Transactions on Medical Imaging, 29(6), 1310-1320. [Accès le 02 mai 2026]

[36] Serra, J. (1982). Image Analysis and Mathematical Morphology. Academic Press. [Accès le 05 mai 2026]

[37] Gonzalez, R. C., Woods, R. E. (2018). Digital Image Processing. Pearson. [Accès le 08 mai 2026]

[38] Badža, M. M., Barjaktarovic, M. C. (2020). Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. Applied Sciences, 10(6), 1999. [Accès le 10 mai 2026]

[39] He, K., et al. (2016). Deep Residual Learning for Image Recognition. CVPR, 770-778. [Accès le 12 mai 2026]

[40] Huang, G., et al. (2017). Densely Connected Convolutional Networks. CVPR, 4700-4708. [Accès le 12 mai 2026]

[41] Tan, M., Le, Q. V. (2019). EfficientNet : Rethinking Model Scaling for Convolutional Neural Networks. ICML, 6105-6114. [Accès le 15 mai 2026]

[42] Howard, A. G., et al. (2017). MobileNets : Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv :1704.04861. [Accès le 18 mai 2026]

[43] Azuma, R. (1997). A Survey of Augmented Reality. Presence : Teleoperators and Virtual Environments, 6(4), 355-385. [Accès le 20 mai 2026]

[44] Vávra, P., et al. (2019). Augmented Reality in Surgery. Medical Science Monitor, 25, 772-780. [Accès le 22 mai 2026]

[45] Incekara, F., et al. (2020). Augmented Reality in Neurosurgery : A Systematic Review. Neurosurgical Review, 43(6), 1467-1478. [Accès le 25 mai 2026]

[46] Condino, S., et al. (2018). Augmented Reality in Neurosurgery. Healthcare Technology Letters, 5(5), 168-173. [Accès le 28 mai 2026]

[47] Bekelman, J. E., et al. (2021). Augmented Reality Limitations in Medical Applications. JAMA Network Open, 4(8), e2120456. [Accès le 30 mai 2026]

[48] Wang, S., et al. (2022). Artificial Intelligence and Augmented Reality in Medical Imaging. Frontiers in Oncology, 12, 876543. [Accès le 02 juin 2026]

[49] Kim, J., et al. (2023). Hybrid AI-AR Systems in Neurosurgery. Journal of Medical Systems, 47(3), 45-56. [Accès le 05 juin 2026]

[50] Sun, C., et al. (2021). GPU Constraints in Medical AI. IEEE Transactions on Medical Imaging, 40(5), 1420-1430. [Accès le 08 juin 2026]

[51] Topol, E. J. (2019). Deep Medicine : How Artificial Intelligence Can Make Healthcare Human Again. Nature Medicine, 25(1), 44-56. [Accès le 10 juin 2026]

[52] Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., Maier-Hein, K. H. (2019). Automated Design of Deep Learning Methods for Biomedical Image Segmentation. arXiv preprint arXiv :1904.08128. [Accès le 12 juin 2026]

[53] Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L. (2011). Unbiased average age-appropriate atlases for pediatric studies. NeuroImage, 54(1), 313-327. [Accès le 12 juin 2026]

[54] Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., ... Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain : International Consortium for Brain Mapping (ICBM). Philosophical Transactions of the Royal Society B, 356(1412), 1293-1322. [Accès le 12 juin 2026]

[55] Serra, J. (1982). Image Analysis and Mathematical Morphology. Academic Press. [Accès le 12 juin 2026]

[56] Soille, P. (2003). Morphological Image Analysis : Principles and Applications. Springer. [Accès le 12 juin 2026]

[57] Gonzalez, R. C., Woods, R. E. (2018). Digital Image Processing. Pearson. [Accès le 12 juin 2026]

[58] Maintz, J. B., Viergever, M. A. (1998). A Survey of Medical Image Registration. Medical Image Analysis, 2(1), 1-36. [Accès le 12 juin 2026]

[59] Zitová, B., Flusser, J. (2003). Image registration methods : a survey. Image and Vision Computing, 21(11), 977-1000. [Accès le 12 juin 2026]

[60] Viola, P., Wells, W. M. (1997). Alignment by Maximization of Mutual Information. International Journal of Computer Vision, 24(2), 137-154. [Accès le 12 juin 2026]

[61] Fateh, A., Rezvani, Y., Moayedi, S., Rezvani, S., Fateh, F., Fateh, M., Abolghasemi, V. (2025). BRISC : Annotated Dataset for Brain Tumor Segmentation and Classification with Swin-HAFNet. arXiv preprint arXiv :2506.14318. [Accès le 16 juin 2026]

[62] Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., Liang, J. (2018). UNet++ : A Nested U-Net Architecture for Medical Image Segmentation. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (DLMIA 2018), LNCS, vol 11045, 3-11. [Accès le 16 juin 2026]

[63] Kamnitsas, K., et al. (2017). Ensembles of Multiple Models and Architectures for Robust Brain Tumour Segmentation. In Brainlesion : Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (BrainLes 2016), LNCS, vol 10154, 246-258. [Accès le 16 juin 2026]

[64] Isensee, F., et al. (2020). nnU-Net for Brain Tumor Segmentation. In Brainlesion, Springer. [Accès le 12 juin 2026]

[65] Magadza, T., et al. (2023). Efficient nnU-Net for Brain Tumor Segmentation. IEEE Access. [Accès le 12 juin 2026]

[66] Raza, A., et al. (2023). Improved Brain Tumor Segmentation. Scientific Reports. [Accès le 12 juin 2026]

[67] Wang, Y., et al. (2022). Attention Mechanisms in Medical Image Segmentation. IEEE TMI. [Accès le 12 juin 2026]

[68] Liew, K., et al. (2023). Optimization for Enhancing Tumor Region. Medical Image Analysis. [Accès le 12 juin 2026]

[69] MobileViTV2-100. (2025). BRISC Benchmark Report. [Accès le 12 juin 2026]

[70] ResNet18. (2025). BRISC Benchmark Report. [Accès le 12 juin 2026]

[71] EfficientNet-B0. (2025). BRISC Benchmark Report. [Accès le 12 juin 2026]

[72] EfficientNet-B1. (2026). BRISC Extended Benchmark. [Accès le 12 juin 2026]

[73] Estimations de prix pour solutions d'imagerie médicale (Siemens syngo.via, GE Healthcare). [Accès le 16 juin 2026]

[74] Chen, L. C., et al. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. ECCV 2018, 801-818. [Accès janvier 2026]

[75] Lin, T. Y., et al. (2017). Feature Pyramid Networks for Object Detection. CVPR 2017, 2117-2125. [Accès le janvier 2026]

[76] He, K., et al. (2016). Deep Residual Learning for Image Recognition. CVPR 2016, 770-778. [Accès mars 2026]

[77] Huang, G., et al. (2017). Densely Connected Convolutional Networks. CVPR 2017, 4700-4708. [Accès mars 2026]

[78] Tan, M., Le, Q. V. (2019). EfficientNet : Rethinking Model Scaling for Convolutional Neural Networks. ICML 2019, 6105-6114. [Accès le 16 Avril 2026]

[79] Howard, A. G., et al. (2017). MobileNets : Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv :1704.04861. [Accès le 16 Avril 2026]

[80] Unity Technologies. (2024). AR Foundation Package. Documentation en ligne : https ://-docs.unity3d.com/Packages/com.unity.xr.arfoundation@5.1. [Accès le 20 Avril 2026]