

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
Ministère de L'Enseignement Supérieur et de la Recherche Scientifique
ECOLE NATIONALE POLYTECHNIQUE
Département génie industriel



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique

LABORATOIRE
DES SCIENCES
DU NUMÉRIQUE
DE NANTES



MÉMOIRE DE PROJET DE FIN D'ÉTUDES
En vue de l'obtention du diplôme d'Ingénieur d'État en Génie Industriel
Option : Data Science et Intelligence Artificielle
Par : BAKHOUCHE Amel

Sujet

**Segmentation sémantique d'images médicales pour la
radiothérapie assistée par ordinateur**

Soutenue publiquement, le 04/07/2023, devant le jury composé de :

M.	Zouaghi Iskander	MCA	à ENP	Président du jury
M.	Fourar-Laidi Hakim	MCA	à ENP	Examineur
Mme	Diana Mateus	Professeure	à LS2N	Promotrice
M.	Arki Oussama	MCB	à ENP	Promoteur

ENP 2023

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
Ministère de L'Enseignement Supérieur et de la Recherche Scientifique
ECOLE NATIONALE POLYTECHNIQUE
Département génie industriel



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique

LABORATOIRE
DES SCIENCES
DU NUMÉRIQUE
DE NANTES



MÉMOIRE DE PROJET DE FIN D'ÉTUDES
En vue de l'obtention du diplôme d'Ingénieur d'État en Génie Industriel
Option : Data Science et Intelligence Artificielle
Par : BAKHOUCHE Amel

Sujet

**Segmentation sémantique d'images médicales pour la
radiothérapie assistée par ordinateur**

Soutenue publiquement, le 04/07/2023, devant le jury composé de :

M.	Zouaghi Iskander	MCA	à ENP	Président du jury
M.	Fourar-Laidi Hakim	MCA	à ENP	Examineur
Mme	Diana Mateus	Professeure	à LS2N	Promotrice
M.	Arki Oussama	MCB	à ENP	Promoteur

ENP 2023

Remerciements

Je tiens à exprimer ma profonde gratitude envers toutes les personnes qui ont contribué de manière significative au succès de mon stage et qui m'ont apporté une aide précieuse tout au long de la rédaction de ce rapport.

Je tiens tout d'abord à remercier chaleureusement ma référente, Madame Diana Mateus, pour les enseignements inestimables qu'elle m'a prodigués, ainsi que pour son soutien constant et ses précieux conseils qui m'ont guidé sur la voie de la réussite de la manière la plus bienveillante et pédagogique qui soit.

Mes remerciements vont également à mon encadrant de stage, Monsieur Arki, qui m'a accompagné tout au long de ce projet en me prodiguant des orientations éclairées et en me fournissant un encadrement attentif.

J'adresse également mes sincères remerciements à Louise, ma collègue doctorante, sans laquelle rien n'aurait été possible. Elle m'a accompagné et m'a prodigué un mentorat précieux tout au long de ces quatre mois.

Je tiens également à exprimer ma reconnaissance envers Vanessa Gonzallez Duques, doctorante, pour son inlassable disponibilité et son aide précieuse dans la résolution de mes interrogations, éclairant ainsi ma lanterne lors des moments d'incertitude.

Je souhaite également exprimer ma profonde gratitude envers Mira, qui a fait en sorte que je me sente véritablement chez moi au sein du laboratoire, et qui veille à l'intégration harmonieuse de chaque nouveau membre de l'équipe.

Je tiens à souligner le soutien technique précieux apporté par Richard Randriatoamanana et à l'en remercier chaleureusement.

Enfin, je souhaite adresser mes remerciements les plus sincères à l'ensemble de l'équipe du Laboratoire des Sciences du Numérique de Nantes, qui m'a accueilli avec une grande générosité. Je tiens à souligner tout particulièrement l'équipe du Signal, Images et Son, comprenant les doctorants, les maîtres de conférence et les stagiaires, pour m'avoir donné un aperçu concret de ce qu'est une véritable dynamique de laboratoire de recherche dans laquelle il est agréable de travailler.

Je suis profondément reconnaissante envers chacune de ces personnes pour leur soutien, leur expertise et leur bienveillance, qui ont grandement contribué à mon expérience de stage enrichissante.

Dédicaces

En préambule de ce mémoire, je remercie ALLAH de m'avoir aidé et de m'avoir donné la patience, la persévérance et le courage pour accomplir ce modeste travail.

À mes parents, je souhaite exprimer ma profonde reconnaissance pour avoir été ma source inépuisable de tendresse, de force et de vie. Je tiens à vous remercier du fond du cœur pour avoir toujours cru en moi et pour m'avoir constamment encouragé à poursuivre mes rêves. Votre foi inébranlable en moi est une source d'inspiration inestimable. Merci pour votre patience infinie et votre amour inconditionnel, qui ne connaissent aucune limite. À tous deux, vous êtes des parents extraordinaires. Votre soutien inlassable et vos sacrifices sont gravés à jamais dans mon cœur. Merci pour tout.

À mon frère, Mehdi, je souhaite exprimer ma profonde reconnaissance pour être bien plus qu'un simple frère. Tu es mon exemple, mon meilleur ami, mon alterego, mon confident, mon soutien, mon épaule sur laquelle je peux m'appuyer. Sans toi, ton amour, ton soutien et ta protection, je ne sais pas qui je suis.

Je te suis extrêmement reconnaissante d'avoir partagé avec moi les plus belles leçons de physiques et de vie. Tu as été un enseignant hors pair, m'insufflant ta passion pour la connaissance et m'aidant à grandir tant sur le plan intellectuel que personnel.

À Nora et Kamel, mes hôtes bienveillants pendant ces quatre mois, qui m'ont accueilli chaleureusement et ont veillé avec attention à chaque petit détail de mon quotidien, je vous remercie du plus profond de mon cœur. Votre générosité et votre prévenance resteront à jamais gravées en moi, et je n'oublierai jamais tout ce que vous avez fait.

À ma tante, qui est une source inépuisable de tendresse et d'amour et une seconde mère pour moi, je suis reconnaissante pour tout.

À ma chère grand-mère, toi, cette vaillante guerrière, tu as été la source de tant d'amour dans ma vie. Grâce à toi, j'ai appris les valeurs de générosité et de persévérance, et ta sagesse

m'a marqué lorsque tu m'as rappelé avec tendresse que "rien n'arrive sans peine". Ta présence bienveillante restera gravée dans mon cœur pour toujours.

À Nour, Google, ma binôme à l'école pendant un temps, mon acolyte dans la vie d'aujourd'hui et pour toujours, espérons-le. Ce que tu m'as apporté pendant ces 3 dernières années, ce que tu m'as appris, ce que tu as donné de ton temps et de ta patience pour moi, ont une valeur inestimable à mes yeux. Je te remercie pour tout.

À Amel, car les liens du cœur sont parfois plus puissants que les liens du sang, je tiens à exprimer ma profonde gratitude pour ton soutien et ton accompagnement au cours de ces trois dernières années. Ta présence à mes côtés était synonyme de famille, d'une sœur sur laquelle je pouvais compter. Je te remercie du fond du cœur pour tout ce que tu as fait pour moi.

À Lina, ma lavande, l'une des personnes que j'aime le plus au monde, ma sœur de cœur, je souhaite exprimer ma profonde gratitude. Merci pour tout ce que tu m'apportes, pour ton amitié si précieuse, pour ta présence apaisante qui m'entoure de réconfort. Je t'aime, bark, simplement.

À Papiche, ma soeur et mon amie, avec qui le liens ne cessent de se renforcer de jour en jour, merci pour ta présence, merci pour ta bienveillance, tu comptes énormément pour moi et tu compteras toujours.

À Khaled, Raouf, Jojo Aimen et Badis mes frères et ma soeur de cœur, qui ont toujours été présents comme une épaule solide sur laquelle je pouvais me reposer, je vous adresse un immense merci pour tout. Votre soutien indéfectible et votre amitié sincère ont illuminé mon chemin et ont été une source inestimable de réconfort.

À real home, Anissa, l'amour et l'attention incarnés. Maya, l'écoute bienveillante et la démonstration d'humanité. Nour, la sagesse et la sensibilité. Hadil, la combinaison parfaite de force et de fragilité. Votre amitié, à vous toutes, est inestimable et vaut bien plus que tout l'or du monde. Merci d'embellir ma vie.

À Faten, Meriem, Sofia, à notre amitié si naturelle, bienfaitante, et authentique, je vous adresse tout mon amour. Merci pour chaque instant de joie et de partage que vous m'avez généreusement offert. Votre présence dans ma vie est un cadeau inestimable.

À la Djorf Family, l'expérience vécue à la cité universitaire restera à jamais gravée dans mon cœur comme l'une des plus belles de ma vie, et cela grâce à vous. Merci à vous tous, de tout mon cœur.

À Nawel, l'amitié qui s'est épanouie au fil du temps est une bénédiction inestimable pour moi. Je tiens à exprimer ma profonde gratitude pour ta bienveillance, ta loyauté et ta générosité. Ta présence dans ma vie apporte une légèreté et une joie indescriptibles. Je t'adore.

À toutes les merveilleuses personnes de ma promotion à Daita, vous avez été la raison pour laquelle ces trois années sont les plus enrichissantes dans tous les aspects de ma vie. Avec tout mon amour et ma gratitude.

ملخص

تركز هذه الدراسة على تجزئة الأعضاء المعرضة للخطر المحيطة بأورام المخ باستخدام صور الأشعة المقطعية وصور الرنين المغناطيسي. كان هدفنا الرئيسي تسهيل تخطيط العلاج الإشعاعي المستهدف من خلال تحديد الأعضاء المعرضة للخطر بدقة. قمنا بتطوير منهجية قائمة على التعلم العميق باستخدام مجموعة بيانات تضم صوراً لـ ١٥ مشاركاً. على الرغم من التحديات التي يشكلها التصنيف المتعدد وتغيرات حجم الأعضاء وتشوهها الناتج عن الأورام، تمكنا من تحقيق تجزئة دقيقة من خلال تقييم أداء طريقتنا باستخدام مقاييس كمية وتحليلات بصرية عميقة. تُظهر النتائج المُستخلصة فعالية نهجنا، مما يفتح آفاقاً جديدة لتطبيقات سريرية متقدمة وتحسين تخصيص العلاجات في التصوير الطبي. الكلمات الرئيسية: الذكاء الاصطناعي، الرؤية الحاسوبية، التجزئة الدلالية، التصوير الطبي، أورام المخ، الأعضاء المعرضة للخطر، العلاج الإشعاعي.

الكلمات المفتاحية: الذكاء الاصطناعي، رؤية الحاسوب، تجزئة الصور، التصوير الطبي، أورام المخ، الأعضاء المعرضة للخطر، العلاج الإشعاعي المستهدفة.

Abstract

This study focuses on the segmentation of organs at risk surrounding brain tumors using Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) images. Our main objective was to facilitate the planning of targeted radiotherapy by accurately delineating the organs at risk. We developed a deep learning-based methodology, using a dataset of images from 51 subjects. Despite the challenges posed by multi-labeling and variations in organ size and deformation caused by tumors, we achieved precise segmentation by evaluating the performance of our method using quantitative measures and in-depth visual analyses. The obtained results demonstrate the effectiveness of our approach, thereby opening new prospects for advanced clinical applications and improved personalization of treatments in medical imaging. Key words : artificial intelligence, computer vision, semantic segmentation, medical imaging, brain tumors, organs at risk, radiotherapy

Keywords : artificial intelligence, computer vision, semantic segmentation, medical imaging, brain tumors, organs at risk, radiotherapy.

Résumé

Cette étude se concentre sur la segmentation d'organes à risque entourant les tumeurs cérébrales en utilisant des images tomographiques (CT) et des imageries par résonance magnétique (IRM). Notre objectif principal était de faciliter la planification d'une radiothérapie ciblée en délimitant précisément les organes à risque. Nous avons développé une méthodologie basée sur l'apprentissage profond, en utilisant un ensemble de données comprenant initialement des images de 51 sujets. Tout au long du développement de cette méthode, nous avons dû faire face à des défis liés à la multilabélisation et aux variations de taille et de déformation des organes provoquées par les tumeurs. Malgré cela, nous avons tout de même réussi à obtenir des segmentations de bonne qualité, comme en témoignent nos évaluations à la fois quantitatives et qualitatives visuelles approfondies. Mots Clés : intelligence artificielle, vision par ordinateur, segmentation d'images, imagerie médicale, tumeurs cérébrales, organes à risque, radiothérapie ciblée.

Mots-Clés : intelligence artificielle, vision par ordinateur, segmentation d'images, imagerie médicale, tumeurs cérébrales, organes à risque, radiothérapie ciblée.

Table des matières

Table des figures

Introduction Générale 15

1 État des lieux

1.1	Organisme d'accueil	18
1.1.1	L'École centrale de Nantes (ECN)	18
1.1.2	Laboratoire des Sciences du Numérique de Nantes (LS2N)	18
1.1.3	Equipe Signal, Image et Son (SIMS)	19
1.2	Enoncé de la problématique	20

2 État de l'art

2.1	Contexte médicale	23
2.1.1	Radiothérapie	23
2.1.2	Organes à risque	24
2.1.3	Radiothérapie assistée par ordinateur	25
2.1.4	Modalités de l'imagerie médicale	26
2.2	Format d'images médicales	29
2.2.1	DICOM	29
2.2.2	NIFTI	30
2.2.3	RTSTRUCT	30
2.2.4	Label Maps	30

2.2.5	Contexte technique	31
2.2.6	Architecture deep learning populaires	40
2.2.7	Défis de l'analyse d'images médicales	43
3	Méthodologie	
3.1	Définition du problème	46
3.1.1	Problème	46
3.1.2	Objectif	46
3.1.3	Entrées	46
3.1.4	Sorties	47
3.1.5	Contraintes	48
3.1.6	Approche proposée	48
3.2	Solution proposée	48
3.2.1	Pré-traitements	49
3.3	Architecture et solution proposée	58
3.3.1	Architecture Explorée	58
3.3.2	Principe générale de la méthode	59
3.3.3	Architecture utilisée	59
3.3.4	Fonction de coût	62
4	Validation expérimentale	
4.1	Details des bases de donnees et annotations	66
4.2	Application de la méthode à nos données	67
4.2.1	Données d'entrée et de sortie du modèle et prétraitement	67
4.2.2	Augmentation de données	68
4.2.3	Séparation de la base de données	69
4.2.4	Entraînement et validation	70
4.2.5	Entraînement sur les deux modalités	73
4.2.6	Test	74
4.2.7	Métriques utilisées	75

4.3	Résultats	78
4.3.1	Résultats quantitatifs	78
4.3.2	Résultats qualitatifs	82
4.3.3	Discussion et perspectives d'amélioration	84
4.4	Conclusion	86
	Conclusion Générale	87
	Bibliographie	89
	Annexe	94
A.1	Imfusion	94
A.2	3Dslicer	94
A.3	LS2N Infra Cloud Computing for AI (LICCIA)	95
A.4	Pytorch	95
A.5	MONAI	95

Table des figures

1.1	Organigramme du LS2N. [1]	19
2.1	Tomodensitométrie (TDM ou scanner CAT) du cerveau. [1]	27
2.2	IRM du cerveau. [2]	28
2.3	Une opération de convolution avec un filtre 3x3. [3]	38
2.4	Exemple d'opération de max pooling. [3]	39
2.5	Exemple de recalage rigide.	39
2.6	Architecture UNET [4]	40
2.7	Architecture UNETR. [5]	43
3.1	Exemple d'image tomosimétrique d'entrée.	46
3.2	Exemple d'image IRM d'entrée.	47
3.3	Exemple de vérité terrain d'entrée.	47
3.4	Dictionnaire de tous les organes à risques annotés chez au moins un des sujets.	50
3.5	Correction des erreurs d'orthographe des organes à risque.	51
3.6	Résultat de la liste ordonnée d'organes à risques.	51
3.7	Conversion de la RTstruct en Labelmaps séparées.	52
3.8	Copie de la liste obtenue lors de l'étape précédente.	52
3.9	Max compounding, pour attribuer à chaque pixel une seule et unique valeur.	52
3.10	LabelMap obtenue après la réalisation de ces étapes.	53
3.11	Résultat de la label map finale.	54

3.12	Ouverture des deux images dans la même scène.	55
3.13	Pré-rapprochement manuel approximatif de l'une deux images (IRM dans notre cas puisque la CT doit être fixe).	55
3.14	Recalage.	56
3.15	Aperçu des deux images recalées superposées.	56
3.16	Image IRM avant normalisation de l'espace.	57
3.17	Images IRM après normalisation de l'espace.	57
3.18	Architecture 3D UNET.	61
3.19	Formule de dice loss.	63
3.20	Dice Overlap.	63
4.2	Lieux et séquences IRM	67
4.1	Déséquilibre hommes/femmes de la base de données.	67
4.3	Sujet numéro 44 dans sa position originale	69
4.4	Sujet numéro 44 dans sa position retournée.	69
4.5	Principe de l'entraînement du modèle.	70
4.6	Courbes de métrique Dice lors de l'entraînement et de la vali- dation sur une base de données d'entraînement sur CT de 37 sujets.	71
4.7	Courbes de métrique Dice lors de l'entraînement et de la va- lidation sur une base de données d'entraînement sur IRM de 37 sujet.	72
4.8	Courbes de métrique Dice Cross Entropy lors de l'entraîne- ment et de la validation sur une base de données d'entraîne- ment sur CT de 37 sujet	73
4.9	Courbes de métrique Dice Cross Entropy lors de l'entraîne- ment et de la validation sur une base de données d'entraîne- ment sur IRM de 37 sujet	73

4.10	Courbes de métrique Dice Cross Entropy lors de l'entraînement et de la validation sur une base de données d'entraînement sur une mélange de 37 IRM ET CT	74
4.11	Composants du calcul de la distance de Hausdorff entre la courbe verte X et la courbe bleue Y.	76
4.12	Matrice de confusion.	77
4.13	Tableau des résultats des métriques obtenus lors de l'entraînement sur 37 sujets en utilisant la loss Dice sur des images IRM.	79
4.14	Tableau des résultats des métriques obtenus lors de l'entraînement sur 37 sujets en utilisant la loss Dice sur des images CT.	80
4.15	Tableau des résultats des métriques obtenus lors de l'entraînement sur 37 sujets en utilisant la loss Dice Cross Entropy sur des images IRM.	81
4.16	Tableau des résultats des métriques obtenus lors de l'entraînement sur 37 sujets en utilisant la loss Dice Cross Entropy sur des images CT.	81
4.17	Tableau des résultats des métriques obtenus lors de l'entraînement sur 37 sujets en utilisant la loss Dice Cross Entropy sur IRM et CT testé sur IRM.	81
4.18	Tableau des résultats des métriques obtenus lors de l'entraînement sur 37 sujets en utilisant la loss Dice Cross Entropy sur IRM et CT testé sur CT.	82
4.19	Présence d'îles sur les segmentations des CT	83
4.20	A gauche superposition de la vérité terrain et de l'image CT, à droite la superposition de notre prédiction (Dice Cross Entropy) et de l'image CT	84
4.21	Exemple de sous-segmentation (a) Véritée terrain (b) Prédiction	84

Introduction Générale

L'intelligence artificielle (IA) est devenue l'un des domaines les plus fascinants et dynamiques de la technologie ces dernières années, avec des progrès fulgurants depuis ses débuts. Aujourd'hui, elle trouve de nombreuses applications dans des domaines tels que la reconnaissance vocale, la recommandation, la reconnaissance faciale, les assistants virtuels, l'analyse des sentiments, la prédiction de résultats financiers, et bien d'autres.

Un sous-domaine qui a connu des avancées remarquables est la vision par ordinateur pour le traitement automatique des images et des vidéos. Les progrès récents ont été stimulés par l'évolution des matériels et des logiciels informatiques, capables de traiter un grand nombre de paramètres dans des modèles entraînaables. Parmi les tâches couramment abordées par la vision par ordinateur, on trouve la classification, qui consiste à étiqueter une image en identifiant l'objet représenté ; la détection, qui vise à localiser spatialement les objets dans l'image et à les classer ; et enfin, la segmentation, qui prédit une carte de classification au niveau des pixels.

La segmentation est la tâche la plus complète mais aussi la plus difficile, car elle nécessite la compréhension de l'image, la localisation des objets, le contourage et enfin la classification. L'un des domaines qui utilise le plus la segmentation d'images est l'imagerie médicale. En supervisant les images, il est possible de diagnostiquer des maladies telles que le cancer et de les traiter de manière optimale. Cependant, les données médicales sont souvent

incomplètes et limitées en quantité, ce qui rend l'application de techniques avancées comme le deep learning complexe, nécessitant une grande quantité de données.

Dans ce contexte, où de nombreuses recherches ont été menées sur la segmentation des tumeurs, notre étude se concentre sur la segmentation des organes à risque entourant une tumeur cérébrale. Pour cela, nous avons utilisé une base de données initiale de 51 sujets.

Cette tâche présente des défis inhérents liés à la multilabelisation et aux variations de taille et de déformation des organes causées par les tumeurs. Malgré ces défis, nous avons réussi à réaliser cette segmentation en utilisant deux modalités d'imagerie distinctes : les images tomодensitométriques (CT) et les imageries par résonance magnétique (IRM).

L'objectif principal de cette segmentation était de faciliter la planification d'une radiothérapie ciblée en permettant une délimitation précise des organes à risque.

Chapitre 1

État des lieux

1.1 Organisme d'accueil

1.1.1 L'École centrale de Nantes (ECN)

L'École centrale de Nantes (ECN), communément appelée Centrale Nantes, est l'une des 204 écoles d'ingénieurs françaises accréditées au 1er septembre 2020 à délivrer un diplôme d'ingénieur. Ayant une certaine importance sur la côte atlantique, elle fait partie des écoles d'ingénieurs d'excellence à la française d'après le Figaro (classement 2023). L'école est par ailleurs une référence dans les domaines de la mécanique, des matériaux et du génie industriel sur le plan national d'après le magazine l'Usine Nouvelle (classements 2022). Cela s'explique notamment de par son histoire et les besoins industriels de la région. Créée en 1919 sous le nom de l'Institut polytechnique de l'Ouest, elle est située dans le quartier Nantes Nord à proximité du campus de Nantes Université. Elle a pour mission principale de former des ingénieurs généralistes destinés en majeure partie aux entreprises. Néanmoins, l'école délivre également des diplômes d'ingénieur en apprentissage spécialisé en partenariat avec l'ITII Pays de la Loire, de mastère spécialisé, de master recherche et de doctorat, pour compter à ce jour près de 2 500 étudiants. Membre du réseau international Top Industrial Managers for Europe (TIME) au même titre que les autres écoles du Groupe des Écoles centrales dont elle fait partie, elle est depuis 2011 membre fondateur du réseau France AEROTECH (avec notamment l'École centrale de Lyon). Au 1er janvier 2022, l'école centrale de Nantes devient un établissement-composante de Nantes Université, dont elle est membre fondateur. [6]

1.1.2 Laboratoire des Sciences du Numérique de Nantes (LS2N)

Le Laboratoire des Sciences du Numérique de Nantes (LS2N) a été créé en janvier 2017 afin de répondre à l'ambition de faire progresser significativement la visibilité de la recherche en sciences du numérique à Nantes. Le

LS2N résulte essentiellement de la fusion des ex-UMR IRCCyN (Institut de Recherche en Communications et Cybernétique de Nantes) et LINA (Laboratoire d'Informatique de Nantes Atlantique). Cette jeune unité mixte de recherche met en commun les forces de recherche dans le domaine du numérique de trois établissements d'enseignement supérieur (Université de Nantes, Centrale Nantes, l'IMT Atlantique/campus de Nantes), du CNRS et de son partenaire Inria. Le LS2N est la plus grosse unité de recherche publique sur le site de Nantes et en région Pays de la Loire.

L'effectif du LS2N s'élève à plus de 480 personnes dont la moitié de permanents, répartis sur 5 sites : Faculté des Sciences et Techniques, Centrale Nantes, Polytech Nantes, IMT Atlantique et IUT de Nantes. [7]

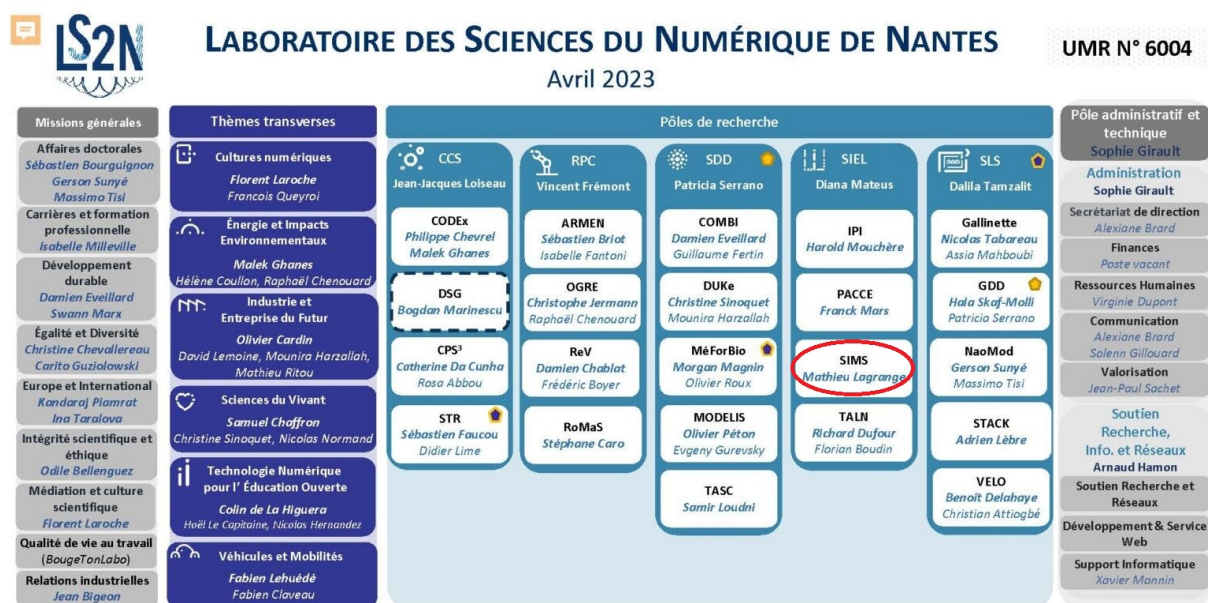


FIGURE 1.1 – Organigramme du LS2N. [1]

1.1.3 Equipe Signal, Image et Son (SIMS)

L'équipe Signal, Image et Son (SIMS) se consacre à la recherche dans le domaine du traitement statistique du signal et de l'image.

Les membres de l'équipe mènent des travaux de recherche qui abordent des questions méthodologiques basées sur des problématiques appliquées prove-

nant de divers domaines tels que la physique, l'ingénierie et l'instrumentation industrielle. Avec l'évolution des exigences en matière de précision des informations sur les grandeurs d'intérêt, les modalités d'acquisition des données se sont diversifiées. Cependant, cette multi-modalité entraîne une augmentation considérable du volume de données, ce qui pose souvent des problèmes aux méthodes classiques en termes de temps de calcul ou de consommation de mémoire.

Heureusement, la croissance des capacités de calcul des processeurs offre une opportunité pour le traitement de ces données, à condition d'avoir de nouvelles méthodes qui permettent une meilleure adaptation entre les algorithmes et l'architecture, tout en exploitant les structures intrinsèques des données.

Le défi principal de l'équipe SIMS est de concevoir de tels algorithmes en utilisant des outils mathématiques éprouvés. [7]

1.2 Enoncé de la problématique

J'ai effectué mon stage de fin d'études au sein de l'équipe SIMS, avec pour sujet une étude inscrite dans le thème de la segmentation sémantique des organes à risques dans des images médicales.

Malgré les avancées significatives de l'intelligence artificielle et, en particulier, de l'apprentissage profond (deep learning), l'application de ces techniques à la segmentation d'images médicales reste complexe en raison de la nécessité de disposer d'une grande quantité de données de haute qualité, ce qui est souvent difficile à obtenir dans le domaine médical.

Nous nous sommes focalisés sur la segmentation de ces organes, en utilisant une base de données initiale de 51 sujets. Cette tâche présente des difficultés inhérentes liées au caractère multi-organe et hétérogène des parties annotées, ainsi qu'aux variations de taille et de déformation provoquées par les tumeurs.

Dans notre approche, nous avons utilisé deux modalités d'imagerie dis-

tinctes, à savoir les images tomodensitométriques (CT) et les imageries par résonance magnétique (IRM). En nous appuyant sur les efforts conjugués des deux laboratoires, le LS2N, qui axe sa recherche sur la segmentation d'images, et le LTSI de Rennes, qui s'intéresse plus particulièrement à la génération d'images synthétiques d'une modalité à partir d'une autre.

Chapitre 2

État de l'art

2.1 Contexte médicale

Dans cette partie, nous allons explorer en détail les concepts médicaux fondamentaux liés à notre objectif de segmentation d'images médicales. Nous nous concentrerons sur les principaux aspects de la médecine qui sont pertinents pour notre étude et qui serviront de base solide pour notre approche. En examinant ces concepts, nous aborderons des sujets tels que les modalités d'imagerie médicale, les tumeurs et organes à risques et la radiologie assistée par ordinateur.

2.1.1 Radiothérapie

Les tumeurs cérébrales, également connues sous le nom de cancers du cerveau, sont des lésions tumorales qui se forment dans le crâne. Lorsqu'un cancer apparaît dans une partie du corps et se propage au cerveau, on parle de métastases cérébrales ou de tumeurs cérébrales métastatiques. Ces métastases cérébrales peuvent provenir de divers types de cancers tels que le cancer du sein, le cancer du poumon, le cancer colorectal, le mélanome, et d'autres. Elles sont détectées chez environ 30 000 nouveaux patients chaque année en France. [8] Lorsqu'un diagnostic de métastases cérébrales est établi, le traitement habituel comprend la chirurgie, la radiothérapie et/ou la chimiothérapie. La radiothérapie est souvent utilisée pour détruire les cellules cancéreuses présentes dans le cerveau. Elle consiste à utiliser des rayonnements ionisants, tels que les rayons X ou les rayons gamma, pour détruire les cellules cancéreuses ou empêcher leur croissance. L'objectif de la radiothérapie est de délivrer des doses précises de rayonnements à la tumeur tout en épargnant les tissus sains environnants autant que possible. Cependant, il est important de noter que la radiothérapie conventionnelle, qui n'est pas ciblée et précise, peut entraîner des inconvénients et des risques pour la santé du patient.

L'un des problèmes principaux liés à la radiothérapie non ciblée est son

impact sur les tissus sains environnants. Étant donné que le cerveau est un organe complexe et sensible, l'irradiation non ciblée peut endommager les tissus cérébraux sains et entraîner des effets secondaires indésirables. Les cellules saines du cerveau peuvent être affectées, entraînant des problèmes neurologiques tels que des troubles cognitifs, des problèmes de mémoire, des altérations de la vision, des maux de tête, et d'autres complications. [9] De plus, la radiothérapie non ciblée peut augmenter le risque de récurrence des métastases cérébrales. Si toutes les cellules cancéreuses ne sont pas détruites de manière précise, il est possible qu'elles continuent à se développer et à former de nouvelles tumeurs dans le cerveau.

C'est pourquoi l'utilisation de techniques de radiothérapie plus précises et ciblées est essentielle dans le traitement des métastases cérébrales. Les avancées telles que la radiothérapie assistée par ordinateur permettent de concentrer la dose de radiation sur les zones tumorales tout en préservant les tissus sains environnants. Cette approche réduit les risques de dommages aux tissus sains et peut améliorer les résultats cliniques tout en réduisant les effets secondaires associés à la radiothérapie conventionnelle. [10]

2.1.2 Organes à risque

Les organes à risque dans le crâne sont les structures sensibles qui peuvent être affectées par la radiothérapie lorsqu'elles sont exposées à des doses élevées de rayonnement. Certains des organes à risque couramment identifiés dans le cerveau comprennent :

Le chiasma optique est une structure située à la base du cerveau, où une partie des nerfs optiques se croise. Il est responsable de la transmission des informations visuelles entre les yeux et le cerveau. Lors de la radiothérapie dans la région cérébrale, il est important de protéger le chiasma optique pour prévenir les complications visuelles. [11]

Les yeux sont des organes sensibles qui doivent être préservés lors de la radiothérapie. Une exposition élevée aux rayonnements peut entraîner des problèmes de vision. Des précautions sont prises pour minimiser la dose de rayonnement reçue par les yeux afin de préserver leur santé.

Le tronc cérébral est une région du cerveau qui contrôle de nombreuses fonctions vitales du corps. Lors de la planification de la radiothérapie, il est essentiel de protéger le tronc cérébral pour éviter tout impact sur ces fonctions essentielles. [12]

Les nerfs oculomoteurs sont responsables du mouvement des yeux. Ils contrôlent les mouvements oculaires et la coordination entre les deux yeux. Lors de la radiothérapie, il est important de préserver ces nerfs pour maintenir une fonction oculaire normale. [13]

L'oreille interne est responsable de l'audition et de l'équilibre. Lorsque la radiothérapie est administrée dans la région cérébrale, il est crucial de préserver l'oreille interne pour éviter des problèmes auditifs et des troubles de l'équilibre. [14]

2.1.3 Radiothérapie assistée par ordinateur

La radiothérapie assistée par ordinateur (RAO) est une avancée majeure dans le domaine du traitement du cancer. [15] Elle permet d'améliorer la précision et l'efficacité de la radiothérapie en utilisant des techniques de planification et de délivrance assistées par ordinateur notamment dans :

Précision de la planification La RAO permet une planification précise du traitement en utilisant des images CT et IRM pour délimiter les tumeurs et les organes à risque. Grâce à la segmentation semi-automatique, les contours

des structures anatomiques sont définis de manière plus précise, ce qui facilite la conception d'un plan de traitement optimal.

Adaptation individuelle La RAO permet de prendre en compte les caractéristiques spécifiques de chaque patient. En utilisant les données anatomiques et radiologiques, il est possible d'adapter le traitement en fonction de la taille, de la forme et de la localisation de la tumeur, ainsi que de la proximité des organes à risque. Cela permet d'optimiser la délivrance de la dose de rayonnement et de réduire les effets indésirables sur les tissus sains environnants.

Contrôle en temps réel La RAO peut également inclure des systèmes de contrôle en temps réel qui surveillent la position du patient pendant le traitement. Cela permet d'ajuster la délivrance de la dose en cas de mouvements ou de déviations du patient, garantissant ainsi une précision accrue.

Avantages pour les patients La RAO offre plusieurs avantages pour les patients, tels qu'une réduction des effets indésirables sur les tissus sains, une amélioration de la qualité de vie pendant et après le traitement, ainsi qu'une augmentation des taux de contrôle local de la tumeur. De plus, la RAO permet souvent de réduire la durée totale du traitement, ce qui peut être bénéfique pour les patients.

2.1.4 Modalités de l'imagerie médicale

L'imagerie médicale est le processus permettant d'obtenir une vue de l'intérieur du corps humain à l'aide d'outils et de techniques spécifiques. Elle facilite le diagnostic et peut être utilisée pour suivre l'évolution des traitements. Les techniques d'acquisition, ou modalités, comprennent l'imagerie par résonance magnétique (IRM), l'échographie, la radiographie, etc. Certaines modalités sont tomographiques, ce qui signifie que les images sont acquises en assemblant des coupes fines en 2D pour former un volume en 3D.

Par exemple, l'IRM et la tomodensitométrie (scanner) produisent toutes deux des images tomographiques. Lors de l'analyse des organes internes, notamment dans le crâne, les modalités les plus couramment utilisées sont l'IRM et la tomodensitométrie, car elles permettent de visualiser tous les organes en trois dimensions, offrant ainsi un aperçu du volume et de la disposition des structures.

Tomodensitométrie (scanner) Le scanner est une technique d'imagerie médicale utilisant des rayons X qui permet d'obtenir différentes images du corps en mesurant l'absorption d'un faisceau de rayons X balayant le patient dans différentes directions. Ces images sont ensuite analysées par un ordinateur qui reconstruit des coupes de différentes parties du corps. Le scanner réalise des coupes virtuelles du corps humain en tranches fines, ce qui permet une reconstitution précise des images de la partie visée. Il offre ainsi une vision détaillée des organes, détecte les tumeurs et permet de définir leur volume, leur consistance, leur étendue aux organes voisins ainsi que leurs contours. Cet examen fournit des images (2D ou 3D) très précises des structures anatomiques de l'organisme. [16] [17]



FIGURE 2.1 – Tomodensitométrie (TDM ou scanner CAT) du cerveau. [1]

Imagerie par Résonance Magnétique (IRM) L'IRM est une technique d'imagerie médicale utilisant un champ magnétique et des ondes radio. Elle permet de visualiser les zones anatomiques du cerveau avec une précision millimétrique. L'IRM est utilisée en neurosciences pour cartographier les différentes zones du cerveau chez des individus en bonne santé ainsi que chez des per-

sonnes atteintes de troubles neurologiques. En pratique médicale, elle est utilisée pour distinguer les tissus pathologiques des tissus sains, par exemple pour détecter les tumeurs cérébrales. L'IRM est également utilisée pour étudier d'autres organes tels que la colonne vertébrale, les articulations et les tissus mous. Cette technique permet d'obtenir des images en 2D ou 3D et est particulièrement utile pour détecter des lésions qui ne sont pas visibles sur les radiographies, les échographies ou les scanners. Elle est principalement utilisée pour étudier les lésions infectieuses ou inflammatoires, les anomalies vasculaires et les tumeurs du cerveau, ainsi que les hernies discales de la colonne vertébrale et les lésions ligamentaires ou méniscales au niveau des articulations. [18]



FIGURE 2.2 – IRM du cerveau. [2]

Comparaison entre les deux L'IRM permet d'obtenir des images du cerveau et de la moelle épinière avec des contrastes plus élevés que le scanner, ce qui la rend efficace pour l'observation des tissus mous. Cependant, l'IRM ne permet pas l'étude des tissus durs tels que les os et les dents qui sont pauvres en hydrogène. Il est important de noter que chaque technique de diagnostic a ses propres avantages et qu'elles sont complémentaires. Ainsi, il n'est pas question de remplacer un examen par un autre, mais plutôt d'utiliser les différentes techniques en fonction des besoins spécifiques du patient. [16]

2.2 Format d'images médicales

L'efficacité et la précision de la segmentation d'images médicales reposent en grande partie sur la qualité et le format des données d'entrée. Les formats d'images médicales sont spécialement conçus pour stocker et représenter les données issues des différentes modalités d'imagerie, telles que l'IRM, la TDM, l'échographie..etc Dans cette partie de l'état de l'art, nous explorerons les principaux formats d'images médicales utilisés dans le contexte de la segmentation avec le deep learning. Nous aborderons des formats couramment utilisés tels que le DICOM (Digital Imaging and Communications in Medicine), qui est largement adopté dans les environnements cliniques pour le stockage et l'échange des données d'imagerie médicale. Nous discuterons également d'autres formats tels que le NIfTI (Neuroimaging Informatics Technology Initiative) et le RTSTRUCT. [19] [16]

2.2.1 DICOM

Le format DICOM (Digital Imaging and Communications in Medicine) est un format de fichiers couramment utilisé dans le domaine de l'imagerie médicale.

Le format DICOM est une norme utilisée pour échanger des images médicales et des informations associées entre différents systèmes. Il permet de stocker des données d'imagerie telles que des radiographies, des scanners, des IRM, etc. Ce format est largement adopté dans le domaine médical en raison de sa compatibilité interopérable entre les équipements et les logiciels. Il facilite la centralisation et le partage des images et des données relatives aux patients. De plus, il offre une large gamme dynamique pour les images enregistrées, pouvant atteindre jusqu'à 16 bits. [20] [21] [16]

2.2.2 NIFTI

Le format NIfTI est spécifiquement utilisé dans le domaine de la neuroimagerie. Il s'agit d'un format de fichier pour le stockage et la manipulation des données d'imagerie cérébrale obtenues par l'IRM fonctionnelle (IRMf) et d'autres techniques d'imagerie du cerveau. Le format NIfTI est basé sur le format de fichier Analyze, mais il inclut des améliorations pour faciliter l'utilisation et l'analyse des données d'imagerie cérébrale. Il permet de stocker à la fois les images anatomiques et les données fonctionnelles liées à l'activité cérébrale. [19] [20]

2.2.3 RTSTRUCT

Les fichiers au format RTSTRUCT (Radiotherapy Structure Set) sont utilisés dans le domaine de la radiothérapie pour le transfert des structures anatomiques et des données associées entre les appareils au sein et à l'extérieur du département de radiothérapie. Ils sont conformes à la norme DICOM (Digital Imaging and Communications in Medicine) et contiennent principalement des informations sur les régions d'intérêt (ROI) et les points d'intérêt, tels que les points de référence de dose. Les fichiers RTSTRUCT peuvent être utilisés pour stocker des contours de segmentation, qui représentent les limites des structures anatomiques sur les images médicales. [22]

2.2.4 Label Maps

Une carte de labels, connue également sous le nom de cartographie des étiquettes ou de cartographie d'étiquetage, se réfère à une représentation visuelle d'une image où chaque pixel ou voxel est associé à une étiquette ou à une classe spécifique. Son utilisation prédominante se situe dans le domaine de la segmentation des images médicales, permettant ainsi l'identification et la délimitation des différentes structures anatomiques présentes dans l'image. [23]

La carte de labels attribue un identifiant unique à chaque structure ou région d'intérêt présente dans l'image. En considérant une image IRM du cerveau comme exemple, la carte de labels pourrait asseoir une identité spécifique aux régions telles que le cortex, l'hippocampe ou le ventricule latéral.

Chaque pixel ou voxel, présent au sein de la carte de labels, renferme une valeur numérique associée à l'étiquette attribuée à la structure anatomique à laquelle il appartient. Par conséquent, les pixels ou voxels appartenant à une même structure affichent une valeur identique au sein de la carte de labels.

La carte de labels est fréquemment employée conjointement à l'image d'origine, dans des méthodes d'apprentissage supervisé automatique, dans le but d'orienter la segmentation automatique des structures anatomiques. Elle apporte ainsi des informations précises quant à la localisation spatiale des différentes régions d'intérêt, offrant ainsi aux algorithmes de segmentation une meilleure compréhension de la forme et de la position des structures à segmenter.

2.2.5 Contexte technique

Dans cette partie, nous approfondirons notre exploration du contexte technique en mettant l'accent sur les domaines du machine learning, du deep learning et des réseaux de neurones convolutionnels (CNN).

Apprentissage automatique L'apprentissage automatique constitue une méthode qui consiste à entraîner un système à l'aide d'exemples pertinents plutôt que de le programmer explicitement, à travers l'utilisation d'exemples pertinents liés à la tâche donnée. Le système d'apprentissage automatique détermine ensuite une structure statistique permettant d'acquérir des représentations significatives des données d'entrée. Ces représentations apprises permettent alors au système d'élaborer ses propres règles en vue de l'automatisation de la tâche. On dénombre quatre types d'apprentissage automa-

tique : supervisé, non supervisé, auto-supervisé et par renforcement. Mon projet repose sur de l'apprentissage supervisé qui requiert trois composantes essentielles [24] :

- Des points de données d'entrée
- Des exemples de sortie attendue
- Des métriques de performance

Les mesures de performance utilisées dans un algorithme d'apprentissage automatique établissent le lien entre l'entrée et la sortie en fournissant un signal de rétroaction pour corriger les sorties de l'algorithme afin de les rapprocher davantage de la sortie escomptée. Cela permet au système d'apprendre à partir des données fournies et de transformer l'entrée en représentations optimales selon une métrique de performance donnée.

Apprentissage profond Le deep learning est une sous-discipline de l'apprentissage automatique qui pousse encore plus loin l'approche de l'apprentissage automatique en apprenant de nombreuses couches successives de représentations des données. Ces couches forment ce que l'on appelle la profondeur du modèle. Le deep learning nécessite l'utilisation de réseaux neuronaux pour apprendre les multiples caractéristiques et représentations des données. [24] [25]

2.2.5.1 Vision par ordinateur

La vision par ordinateur représente un domaine de l'intelligence artificielle (IA) qui confère aux ordinateurs et aux systèmes la capacité d'extraire des informations pertinentes à partir de diverses sources visuelles telles que des images numériques et des vidéos. Cette capacité leur permet ensuite d'agir ou de formuler des recommandations en se basant sur ces informations. Alors que l'IA permet aux ordinateurs de raisonner, la vision par ordinateur leur offre la possibilité de voir, d'observer et de comprendre.

Le fonctionnement de la vision par ordinateur présente des similitudes avec celui de la vision humaine, bien que les êtres humains aient une avance

considérable dans ce domaine. La vision humaine bénéficie en effet de décennies d'expérience permettant d'apprendre à distinguer les objets, à évaluer leur distance, à détecter leur mouvement et à identifier d'éventuels problèmes dans une image.

La vision par ordinateur entraîne les machines à accomplir ces mêmes tâches, mais avec un temps d'apprentissage considérablement réduit, en utilisant des caméras, des données et des algorithmes plutôt que des rétines, des nerfs optiques et un cortex visuel. Grâce à cette méthode, un système formé pour inspecter des produits ou surveiller des actifs de production peut analyser des milliers d'éléments ou de processus par minute, permettant ainsi de détecter des défauts ou des problèmes imperceptibles et de dépasser rapidement les capacités humaines. [25]

2.2.5.2 Segmentation sémantique

La segmentation sémantique vise à étiqueter une image au niveau des pixels, ce qui nécessite de localiser, délimiter et classifier chaque objet. Cette tâche est importante pour la planification d'interventions ou la comparaison des changements physiques en réponse à un traitement. [26] Dans le domaine de la segmentation d'images, il existe deux types principaux de segmentation : la segmentation manuelle et la segmentation automatique. Chacun de ces types présente des caractéristiques distinctes et joue un rôle important dans l'analyse et le traitement des images.

La segmentation manuelle réalisée par des opérateurs ou des experts qui effectuent le processus de segmentation de manière manuelle. Ils identifient et délimitent les régions ou les objets d'intérêt dans une image en utilisant leur expertise et leurs connaissances. Bien que ce processus puisse être laborieux et soumis à des variations inter-opérateurs, il offre un contrôle précis sur les résultats de la segmentation .

la segmentation automatique également appelée segmentation assistée par ordinateur, utilise des algorithmes et des techniques informatiques pour effectuer la segmentation d'images de manière automatisée, sans intervention humaine directe. Les méthodes de segmentation automatique peuvent être basées sur des seuils, des contours, des régions, des réseaux de neurones ou d'autres approches. Ces méthodes automatisées permettent de gagner du temps et peuvent être appliquées à un grand nombre d'images. [27]

Dans le domaine de la segmentation automatique, il existe plusieurs techniques et algorithmes spécifiques. Voici quelques exemples courants :

A. Segmentation basée sur les seuils : Cette technique consiste à définir des seuils ou des valeurs de seuil pour séparer les régions d'intérêt de l'image en fonction des niveaux d'intensité des pixels. Elle est largement utilisée dans diverses applications de traitement d'images.

B. Segmentation par contour actif Également connue sous le nom de "snake" ou "active contour", cette méthode utilise des modèles mathématiques pour détecter et suivre les contours des objets dans une image. Elle est souvent utilisée dans des tâches telles que la détection des contours et la segmentation des contours.

C. Segmentation par régions Cette approche divise l'image en régions homogènes en se basant sur des critères tels que la couleur, la texture, la forme, etc. Elle permet de regrouper les pixels similaires et de créer des segments cohérents.

D. Segmentation par réseaux de neurones Les réseaux de neurones, en particulier les réseaux de neurones convolutifs (CNN), sont utilisés pour effectuer la segmentation automatique en apprenant à reconnaître les caractéristiques des différentes classes d'objets dans les images. Les CNN ont connu

des avancées significatives ces dernières années et ont été largement utilisés dans des applications de segmentation d'images.

2.2.5.3 Les réseaux neuronaux

Les réseaux neuronaux sont des modèles structurés avec des couches empilées les unes sur les autres, chaque étape filtrant successivement les données d'entrée pour apprendre une représentation finale utile pour une tâche donnée [24]. La structure d'un réseau neuronal à quatre couches de base utilisé pour identifier des chiffres manuscrits à partir de l'ensemble de données MNIST est illustrée dans la Figure 1. L'entraînement d'un réseau neuronal implique quatre composantes :

- Couches
- Données d'entrée
- Fonction de perte
- Optimiseur

La difficulté dans l'optimisation de ces réseaux pour différentes tâches réside dans la configuration de ces poids, car il peut exister des dizaines de millions de paramètres qui dépendent les uns des autres. Pour déterminer une configuration de poids, une fonction de perte est utilisée comme mesure de performance pour évaluer la façon dont les représentations de couches en sortie du réseau se comportent par rapport à la sortie attendue. Ce score est ensuite utilisé comme signal de rétroaction pour que l'optimiseur ajuste les poids des couches du réseau à l'aide d'un algorithme de rétropropagation.

L'algorithme de rétropropagation commence par la valeur finale de la perte et travaille des couches supérieures aux couches inférieures pour calculer la contribution de chaque paramètre dans cette valeur. La descente de gradient stochastique consiste à prendre la dérivée de la fonction de perte en un point local lors de l'entraînement afin de déterminer les ajustements nécessaires pour minimiser la perte et optimiser le modèle selon une métrique d'évalua-

tion définie.

Cette boucle d'entraînement commence par une initialisation aléatoire des poids, ce qui entraîne un score de perte élevé puisque ces poids initiaux sont probablement incorrects pour représenter les données de manière utile. À chaque lot d'entraînement, cependant, le réseau ajuste ses poids et commence lentement à s'améliorer jusqu'à ce qu'il atteigne un score de perte acceptable.

Couches denses Traditionnellement, pour les ensembles de données autres que les images, des couches connectées de manière dense sont utilisées, également appelées couches entièrement connectées, dans lesquelles chaque neurone d'une couche est connecté à tous les neurones de la couche précédente. Cependant, ces couches denses ne sont capables d'apprendre que des motifs globaux dans leur espace de caractéristiques d'entrée, tandis que les couches de convolution sont capables d'apprendre des motifs locaux. Cela signifie que les images d'entrée peuvent être découpées en caractéristiques telles que les contours et les textures, qui peuvent être apprises et sont plus utiles que les motifs globaux à des fins de classification

2.2.5.4 Réseau neuronal convolutif (CNN ou ConvNet)

Les CNN sont des algorithmes puissants du domaine de l'apprentissage profond, plus précisément du Deep Learning. Les CNN sont principalement utilisés dans le traitement d'images et ont été à l'origine de progrès significatifs dans ce domaine. [28] [29] [24]

Un CNN est composé de différentes couches, chacune ayant un rôle spécifique dans le processus de traitement des données visuelles. La première couche est généralement une couche de convolution suivie par une couche de pooling.

Les couches convolutives et de pooling sont alternées plusieurs fois pour permettre l'apprentissage de caractéristiques de plus en plus complexes et abstraites. Enfin, les caractéristiques extraites sont acheminées vers des couches

entièrement connectées, qui réalisent la segmentation ou d'autres tâches spécifiques.

Les CNN bénéficient de plusieurs avantages. Ils sont capables de détecter des motifs et des structures complexes dans les images, tels que des contours, des textures ou des objets, grâce à leur capacité à apprendre des représentations hiérarchiques à partir des données. De plus, les CNN sont résilients aux variations de position, de taille et d'orientation des objets dans les images.

Convolution L'opération de convolution est utilisée dans les réseaux neuronaux pour traiter des données d'entrée telles que des images ou des vidéos. L'idée principale derrière la convolution est de réduire le nombre de connexions nécessaires entre les couches de neurones, tout en conservant la capacité de détecter et de reconnaître des caractéristiques importantes dans les données. [24] Lorsque l'entrée est une image, par exemple, une connexion pondérée serait nécessaire entre chaque pixel de l'image et chaque neurone de la couche suivante. Cela entraînerait un grand nombre de connexions pondérées et un grand nombre de paramètres à optimiser. Pour rendre ce processus plus efficace, la convolution propose une approche différente.

Au lieu de connecter chaque pixel à chaque neurone, la convolution utilise des régions locales de l'image pour créer des connexions. Ces régions locales sont également appelées "filtres" ou "masques". Par exemple, un filtre 5x5 serait appliqué à l'image en se déplaçant pixel par pixel. Chaque position du filtre est multipliée par la valeur du pixel correspondant, puis les produits sont sommés pour obtenir une valeur unique. Cette valeur unique est ensuite utilisée comme activation pour le neurone correspondant dans la couche suivante.

Le grand avantage de cette approche est que le même filtre est utilisé pour tous les pixels de l'image. En d'autres termes, le filtre est "glissé" sur toute l'image. Cela permet de détecter des caractéristiques indépendamment de leur position exacte dans l'image. Par exemple, un filtre peut être conçu

pour détecter les bords, et il sera capable de les détecter peu importe où ils se trouvent dans l'image.

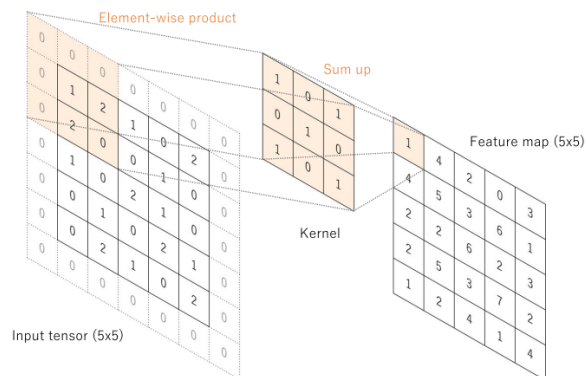


FIGURE 2.3 – Une opération de convolution avec un filtre 3x3. [3]

Strides Un facteur pouvant contribuer à une taille de sortie différente dans les réseaux de neurones convolutifs (CNN) est appelé le stride de la convolution. Le stride est un paramètre de l'opération de convolution qui définit la distance entre les patchs extraits de la carte des caractéristiques d'entrée. Avec un stride de 2, la largeur et la hauteur de la carte des caractéristiques de sortie sont réduites de moitié.

Pooling Le pooling est une opération couramment utilisée dans les réseaux neuronaux pour réduire la dimensionnalité des données et extraire les caractéristiques les plus importantes. Le pooling maximum (max pooling), précisément consiste à sélectionner la valeur la plus élevée dans un groupe de sorties de la couche précédente. Cela permet de mettre en évidence les caractéristiques les plus importantes et de réduire le bruit ou les variations mineures dans les données.

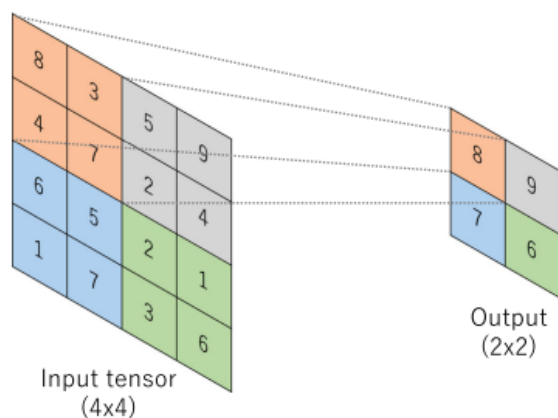


FIGURE 2.4 – Exemple d'opération de max pooling. [3]

2.2.5.5 Le recalage

En traitement d'images, le recalage est une technique consistant à faire correspondre des images afin de comparer ou de combiner leurs informations respectives. La méthode repose sur les mêmes principes physiques et le même type de modélisation mathématique que la corrélation d'images.

Cette mise en correspondance se fait en trouvant des transformations géométriques qui permettent le transfert d'une image à une autre. La technologie a de nombreuses applications, de l'imagerie médicale (par exemple pour combiner plusieurs modalités d'imagerie).

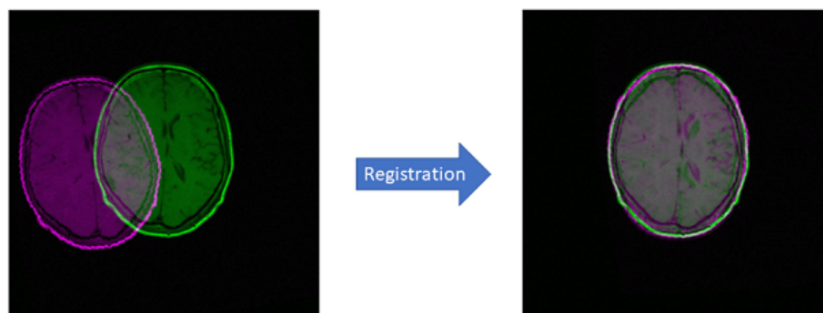


FIGURE 2.5 – Exemple de recalage rigide.

2.2.6 Architecture deep learning populaires

2.2.6.1 UNET

L'architecture U-Net est un modèle de réseau de neurones convolutifs largement utilisé dans le domaine de la segmentation sémantique des images, en particulier dans le contexte de l'imagerie biomédicale. Son nom, U-Net, fait référence à sa structure en forme de U. [4]

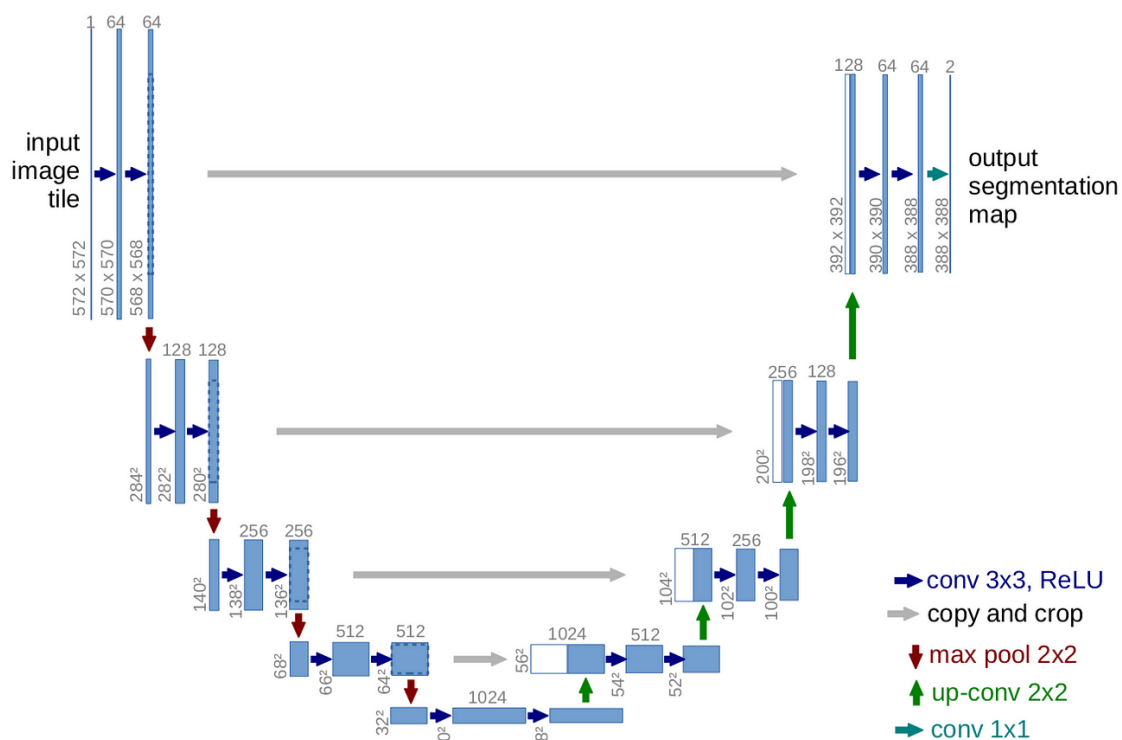


FIGURE 2.6 – Architecture UNET [4]

Le principe de base de l'architecture U-Net est de combiner une "voie contractante" (contracting path) et une "voie expansive" (expansive path). La voie contractante est responsable de l'extraction des caractéristiques à partir de l'image d'entrée, tandis que la voie expansive est responsable de la génération de la carte de segmentation précise.

La voie contractante est composée de plusieurs couches de convolution qui réduisent progressivement la dimension spatiale de l'image tout en augmen-

tant le nombre de canaux de caractéristiques. Cette réduction de dimension est réalisée en utilisant des opérations de convolution, de mise en commun (pooling) ou de sous-échantillonnage (downsampling) successives. L'objectif de cette voie est d'obtenir une représentation compacte et abstraite de l'image, tout en conservant les informations essentielles pour la segmentation.

La voie expansive utilise des opérations de convolution transposée (ou déconvolution) pour effectuer une upsampling (sur-échantillonnage) de l'image réduite à partir de la voie contractante. Cette étape permet de restaurer progressivement la dimension spatiale de l'image et d'obtenir une carte de segmentation finale. Des opérations de concaténation sont réalisées entre les couches de la voie expansive et les couches correspondantes de la voie contractante afin de fusionner les informations de localisation et de contexte.

L'architecture U-Net est caractérisée par sa capacité à effectuer une segmentation précise des objets dans les images. Elle est particulièrement adaptée aux tâches où la localisation précise des contours et des régions est cruciale, comme dans l'imagerie médicale pour la détection de structures anatomiques.

2.2.6.2 UNETR

Pour comprendre Unetr il faut d'abord commencer par expliquer les ViTs :

Vision Transformers (ViTs) Les transformeurs pour la Vision sont une classe d'architectures de réseaux neuronaux qui appliquent le modèle du transformeur, développé initialement pour le traitement du langage naturel (NLP), au domaine de la vision par ordinateur. [30]

Traditionnellement, les réseaux neuronaux convolutifs (CNN) étaient l'architecture dominante pour les tâches liées à l'image. Cependant, les ViTs offrent une approche alternative en exploitant le mécanisme d'auto-attention des transformateurs. Au lieu d'utiliser des couches convolutionnelles, les ViTs

traitent les images comme des séquences de patches et utilisent l'auto-attention pour capturer les relations entre ces patches.

En traitant les images comme des séquences, les ViTs permettent de capturer des informations contextuelles globales, ce qui peut être bénéfique pour diverses tâches de vision. Les ViTs introduisent également des encodages positionnels pour fournir des informations spatiales au modèle du transformeur. La combinaison de l'auto-attention et des encodages positionnels permet aux ViTs de modéliser les dépendances à longue distance et de capturer les détails fins des images.

Les ViTs ont démontré des performances remarquables sur diverses références en vision par ordinateur, notamment la segmentation sémantique. Cependant, ils nécessitent souvent de grandes quantités de données et des ressources informatiques importantes pour l'entraînement.

Transformateurs U-Net (UNeTR) : S'appuyant sur le succès des ViTs, les chercheurs ont exploré l'intégration des transformateurs dans l'architecture U-Net, ce qui a donné naissance aux Transformateurs U-Net (UNeTR). L'architecture U-Net, initialement proposée pour la segmentation d'images biomédicales, est largement utilisée pour son efficacité à capturer les détails fins et le contexte spatial. [5]

Les UNeTR combinent la structure encodeur-décodeur de l'U-Net avec le mécanisme d'auto-attention des transformateurs. Dans les UNeTR, l'encodeur traite l'image d'entrée à l'aide d'une série de couches d'auto-attention, capturant ainsi des informations contextuelles globales. Le décodeur utilise ensuite une combinaison de skip connections et d'upsampling basé sur les transformers pour récupérer les détails spatiaux et générer la carte finale de segmentation.

Les UNeTR ont montré des résultats prometteurs dans diverses tâches de segmentation, notamment la segmentation d'images médicales, où ils ont atteint des performances de pointe. En combinant les points forts de l'U-Net et des transformateurs, les UNeTR offrent un cadre puissant pour capturer

les caractéristiques détaillées des images tout en exploitant les informations contextuelles globales mais sont très coûteuses en ressources et en temps de calcul.

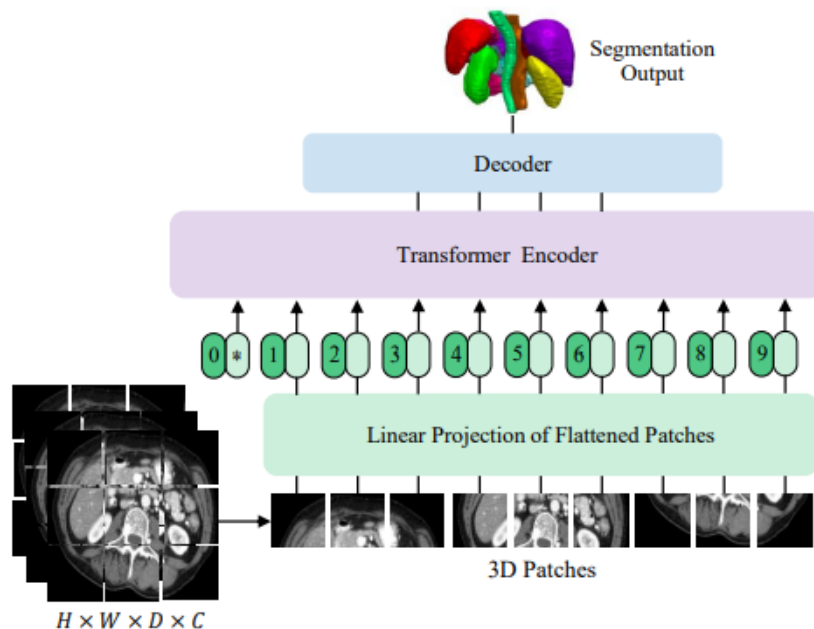


FIGURE 2.7 – Architecture UNETR. [5]

2.2.7 Défis de l'analyse d'images médicales

L'annotation manuelle des données médicales est une tâche coûteuse et chronophage. Les experts médicaux doivent identifier et étiqueter manuellement les structures anatomiques d'intérêt ou les anomalies dans les images, ce qui peut nécessiter une expertise et un investissement en temps important. Par conséquent, les ensembles de données annotées disponibles peuvent être limités en taille et en diversité. [27]

Le manque de données annotées pose plusieurs défis à l'analyse d'images médicales. Voici quelques-uns de ces défis :

Limitations de la généralisation les modèles d'apprentissage automatique ont besoin de données représentatives et diversifiées pour se généraliser efficacement à de nouvelles données. Le manque de données annotées conduit

à une généralisation insuffisante du modèle, ce qui affecte sa précision et sa fiabilité lorsqu'il est utilisé pour analyser de nouvelles images médicales.

Risque de surapprentissage : lorsque l'ensemble de données est limité, il existe un risque accru de surajustement, également appelé surapprentissage. Le surapprentissage se produit lorsqu'un modèle apprend à mémoriser des données d'entraînement spécifiques au lieu de généraliser des modèles et des fonctionnalités plus larges. Peut entraîner de mauvaises performances lorsque le modèle est appliqué à de nouvelles données.

Biais et variabilité dans les annotations les annotations faites par des experts médicaux peuvent souffrir de biais et de parti pris. Différents annotateurs peuvent interpréter les images différemment ou avoir des critères d'annotation différents, ce qui peut introduire des incohérences dans l'ensemble de données annoté. De plus, les annotations peuvent être sujettes à des erreurs humaines, ce qui affecte la qualité et la fiabilité des données annotées.

Dans le cadre spécifique de notre projet de fin d'études, le manque de données ainsi que la variabilité inhérente aux annotations constituent les défis les plus significatifs auxquels nous sommes confrontés. Il est donc primordial de prendre en compte ces facteurs lors de la conception et de l'évaluation de nos méthodes de segmentation.

Chapitre 3

Méthodologie

3.1 Définition du problème

3.1.1 Problème

Segmentation automatique des organes à risques adjacents à une/des tumeur sur des images CT/IRM pour la planification précise de la radiothérapie.

3.1.2 Objectif

Développer un algorithme de segmentation basé sur un réseau de neurones capable de localiser et de prédire avec précision l'appartenance de chaque pixel d'une image à l'un des organes à risques adjacents à une tumeur à partir d'une base de données d'images CT/IRM où uniquement les CT sont annotées. L'objectif principal est de faciliter la planification de la radiothérapie en évitant les dommages potentiels aux tissus sains.

3.1.3 Entrées

- Un ensemble d'images volumétriques CT/IRM du crâne des patients.
- Des annotations manuelles sur CT où les experts ont marqué les contours des organes à risques et de la/les tumeur(s).

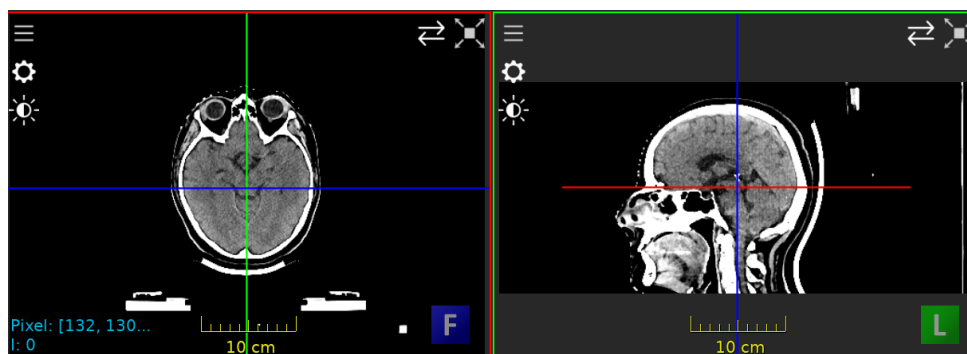


FIGURE 3.1 – Exemple d'image todosimétrique d'entrée.

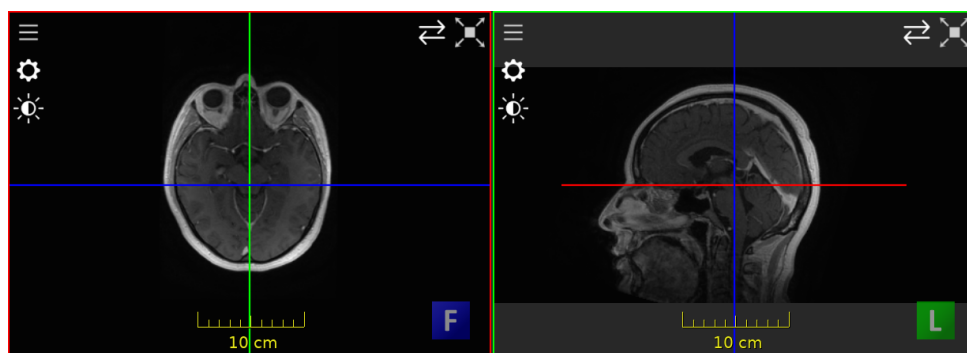


FIGURE 3.2 – Exemple d’image IRM d’entrée.

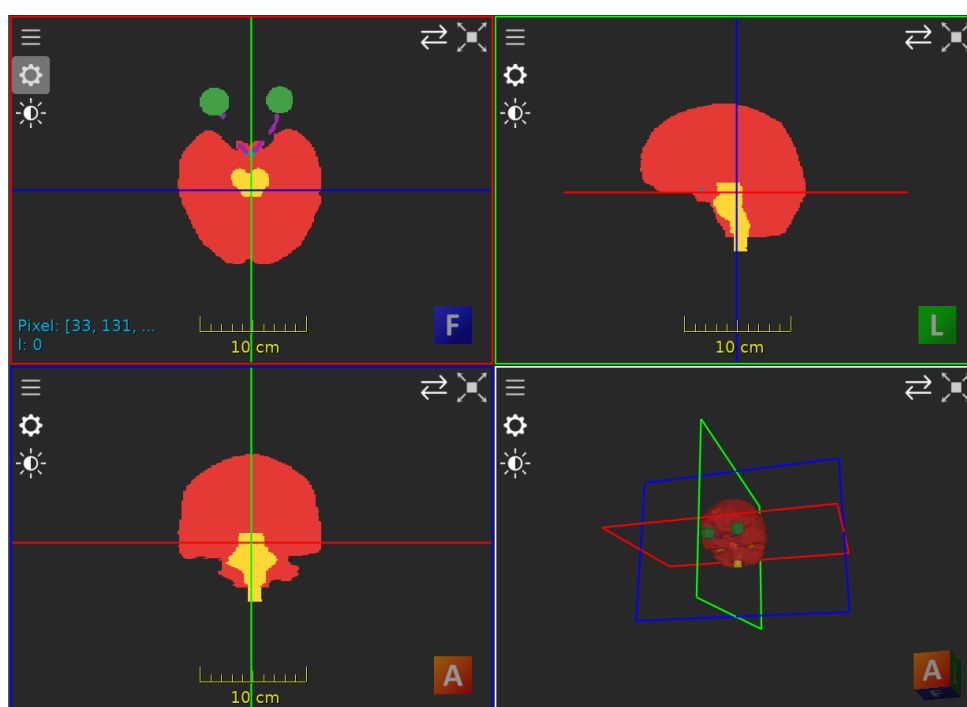


FIGURE 3.3 – Exemple de vérité terrain d’entrée.

3.1.4 Sorties

Une cartographie labellisée des organes à risques adjacents à la tumeur dans chaque image CT/IRM, fournissant des contours précis des organes à risques identifiés. Ces contours permettent de délimiter les régions des organes à risques et de les différencier des tissus environnants.

3.1.5 Contraintes

- Le nombre, la texture, et l'emplacement des tumeurs peuvent varier d'un patient à un autre.
- Les organes à risques à segmenter peuvent varier en fonction de la/les tumeurs. L'algorithme de segmentation doit être capable de traiter des images CT/IRM avec différentes résolutions et niveaux de bruit, afin d'être applicable à diverses situations cliniques.
- La précision de la segmentation doit être bonne, en minimisant l'erreur de contour, pour permettre une planification précise de la radiothérapie et une protection maximale des tissus sains.

3.1.6 Approche proposée

- Prétraitement des images CT/IRM pour améliorer la qualité de l'image et normaliser les intensités.
- Construction d'un modèle de segmentation basé sur l'apprentissage automatique, tel que les réseaux de neurones convolutifs (CNN) ou les réseaux de neurones profonds (DNN), en utilisant les images annotées comme données d'entraînement.
- Entraînement du modèle en optimisant les hyperparamètres et en utilisant des techniques d'augmentation de données pour améliorer la généralisation.
- Validation du modèle sur des données CT/IRM supplémentaires non utilisées lors de l'entraînement, afin de tester sa capacité à généraliser à de nouveaux cas.

3.2 Solution proposée

Les solutions proposées pour la segmentation d'images médicales offrent une approche complète et méthodique pour préparer les données en vue d'un

traitement efficace. En combinant des méthodes de préparation de données, de recalage et de normalisation, ces solutions permettent de créer des cartes de labels cohérentes, optimisant ainsi la segmentation précise des organes d'intérêt tout en tenant compte de la symétrie anatomique.

3.2.1 Pré-traitements

Préparation des données

- Conversion des étiquettes du format RTSTRUCT au format NIfTI pour une meilleure compatibilité avec les langages et les frameworks utilisés.
- Uniformisation des annotations pour garantir la cohérence et la comparabilité des résultats entre les patients.

Pour assurer une représentation équilibrée des organes d'intérêt tout en évitant que les organes de l'arrière-plan ne prédominent indûment lors de l'étape du "max compounding", nous avons élaboré une stratégie basée sur un dictionnaire de valeurs de labels. Ce dictionnaire était conçu de manière à attribuer les valeurs les plus basses aux organes d'arrière-plan et aux petits organes, et des valeurs plus élevées aux organes que nous souhaitions mettre en évidence. Ainsi, lors de l'étape ultérieure de fusion des labels, les organes d'intérêt bénéficieraient d'une pondération supérieure, minimisant ainsi le risque d'écrasement par les organes de moindre importance.

La première étape de mise en œuvre de cette stratégie consistait à sélectionner le chemin d'accès vers le fichier "rt struct". Ce fichier contient des informations structurées sur les structures anatomiques présentes dans l'image médicale. En identifiant et en ouvrant ce fichier, nous avons pu déterminer l'ordre dans lequel les informations seraient affichées. Cette étape était cruciale pour garantir la cohérence dans le traitement des labels, car elle permettait d'organiser les organes de manière prédictible et reproductible.

```
{'Background': 0, 'C_ext': 1, 'Whole_Brain': 2, 'Cerveau_hors_PTV': 3, 'Cerveau_et_PTV': 4, 'Brain_Stem': 5, 'Cervelet': 6, 'Moelle': 7, 'Left_Eye': 8, 'Right_Eye': 9, 'Cristallin_D': 10, 'Cristallin_G': 11, 'Retine_D': 12, 'Retine_G': 13, 'Optic_Chiasm': 14, 'NOD': 15, 'NOG': 16, 'Right_Hippocampus': 17, 'Left_Hippocampus': 18, 'Nez': 19, 'OID': 20, 'OIG': 21, 'Oreille_Moyenne_G': 22, 'Right_Lens': 23, 'Left_Lens': 24, 'Pituitary': 25, 'Hypophyse': 26, 'Tumor': 27, 'PTV': 28, 'CTV': 29, 'GTV': 30}
```

FIGURE 3.4 – Dictionnaire de tous les organes à risques annotés chez au moins un des sujets.

À l'aide du dictionnaire préalablement défini, nous avons réorganisé les labels en fonction de l'ordre établi dans le fichier "rt struct". Cela signifie que chaque organe a été associé à une valeur spécifique du dictionnaire, permettant ainsi de créer une correspondance précise entre les organes et les valeurs qui leur étaient attribuées. Par conséquent, les organes d'intérêt recevaient des valeurs plus élevées, tandis que les organes d'arrière-plan et les petits organes se voyaient attribuer des valeurs plus basses.

Avant de procéder à cette réorganisation, il était important de corriger d'éventuelles erreurs de typographie dans les labels. Ces erreurs peuvent survenir lors de la saisie manuelle des informations ou de la conversion des données. En corrigeant ces erreurs, nous avons obtenu une séquence ordonnée de chiffres représentant de manière précise et cohérente les différents organes présents dans l'image médicale.

```

label_names = [word.replace("brain_stem", "Brain_Stem") for word in label_names]
label_names = [word.replace("Brain-Stem", "Brain_Stem") for word in label_names]
label_names = [word.replace("Optic_Chiasm", "Optic_Chiasm") for word in label_names]
label_names = [word.replace("Whole_Brain", "Whole_Brain") for word in label_names]
label_names = [word.replace("Right-Lens", "Right_Lens") for word in label_names]

```

FIGURE 3.5 – Correction des erreurs d'orthographe des organes à risque.

```

Updated label numbers for key 1: 11 12 5 6 3 4 13 7 8 9 10 14 15 2 1
Updated label numbers for key 2: 11 12 5 6 4 16 17 18 19 7 8 3 9 10 2 13 1
Updated label numbers for key 4: 11 20 12 3 6 7 5 8 15 4 2 9 10 13 1
Updated label numbers for key 5: 11 12 5 6 4 2 21 9 10 7 8 3 16 17 20 13 1
Updated label numbers for key 6: 11 12 3 6 7 14 5 8 15 4 22 2 13 9 10 1
Updated label numbers for key 7: 20 11 12 5 6 3 4 8 7 9 10 21 13 14 15 2 23 24 1
Updated label numbers for key 8: 11 12 5 6 3 4 13 12 12 11 11 11 7 8 9 10 11 14 15 2 1
Updated label numbers for key 10: 11 12 5 6 22 3 4 13 14 15 2 10 9 7 8 23 24 1
Updated label numbers for key 11: 11 12 5 6 22 3 4 9 10 7 8 13 2 1
Updated label numbers for key 15: 11 12 5 6 22 3 4 25 7 8 21 9 10 2 1
Updated label numbers for key 19: 11 11 12 3 6 7 14 5 8 15 26 4 22 2 13 27 23 24 12 9 10 18 19 11 1
Updated label numbers for key 20: 11 12 5 6 3 4 7 8 14 15 2 13 1
Updated label numbers for key 21: 11 12 3 6 7 14 5 8 15 4 22 2 13 1
Updated label numbers for key 22: 11 12 3 6 7 9 5 8 10 4 2 13 1
Updated label numbers for key 23: 11 12 5 6 22 3 4 2 21 9 10 7 8 13 1
Updated label numbers for key 26: 11 12 5 6 22 3 4 7 8 21 9 10 13 2 1
Updated label numbers for key 27: 20 12 12 3 6 7 5 8 4 2 11 11 13 11 1
Updated label numbers for key 33: 12 12 12 5 6 22 3 4 21 10 9 7 8 2 11 11 11 13 1
Updated label numbers for key 36: 12 12 12 3 11 11 6 7 11 5 8 11 13 4 2 9 10 1

```

FIGURE 3.6 – Résultat de la liste ordonnée d'organes à risques.

Une fois ces étapes préliminaires réalisées, nous avons utilisé un logiciel spécialisé appelé "imfusion" pour créer des cartes de labels séparées, assignant à chaque organe le label normalisé approprié. L'utilisation de ce logiciel nous a permis d'automatiser le processus de création des cartes de labels et de garantir leur cohérence avec les informations structurées du fichier "rtstruct".

Enfin, nous avons appliqué une opération de fusion des labels appelée "max compounding". Cette opération consiste à attribuer à chaque pixel le label correspondant à la valeur la plus élevée parmi tous les labels présents à cet emplacement. Ainsi, chaque pixel de l'image finale se voit attribuer le label de l'organe le plus dominant à cet endroit. Cela nous a permis d'obtenir une carte de labels ordonnée, où chaque pixel est étiqueté avec le label correspondant à l'organe prédominant.

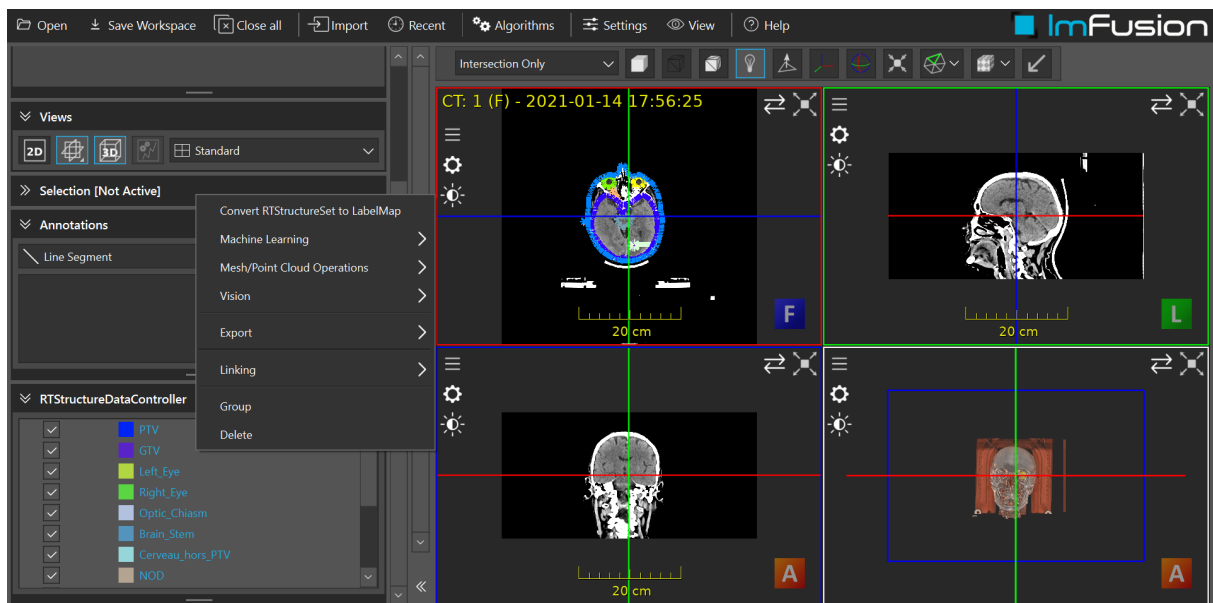


FIGURE 3.7 – Conversion de la RTstruct en Labelmaps séparées.

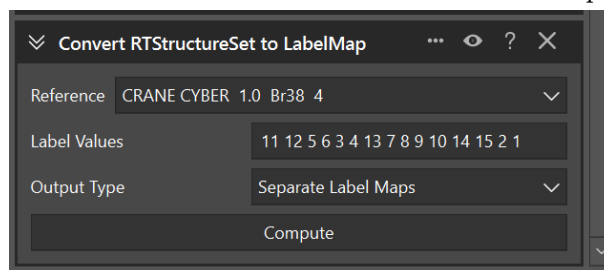


FIGURE 3.8 – Copie de la liste obtenue lors de l'étape précédente.

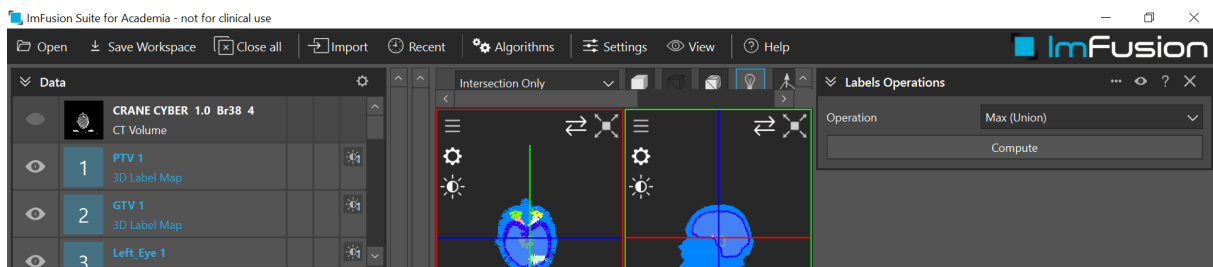


FIGURE 3.9 – Max compounding, pour attribuer à chaque pixel une seule et unique valeur.

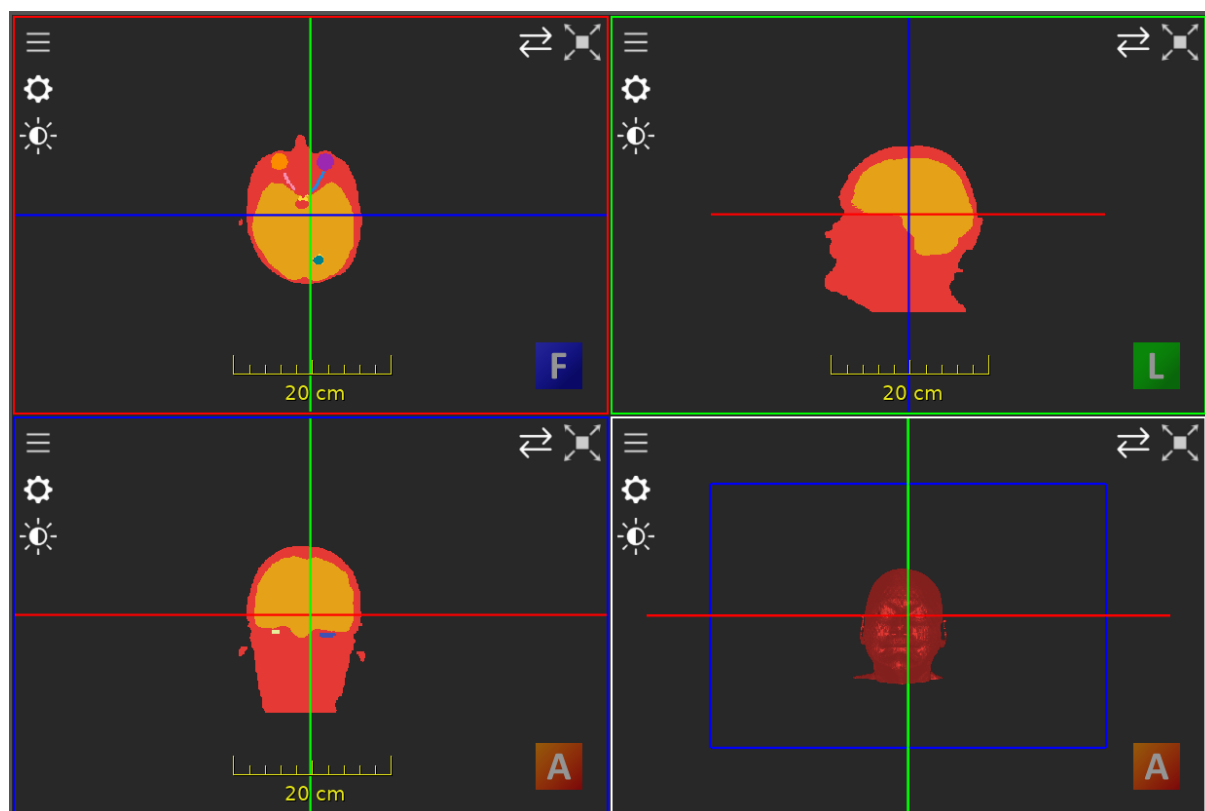


FIGURE 3.10 – LabelMap obtenue après la réalisation de ces étapes.

Ensuite, nous avons découvert un aspect supplémentaire qui a contribué à améliorer l'apprentissage du modèle : l'attribution de labels identiques aux organes symétriques. En observant attentivement les images médicales, nous avons constaté que de nombreux organes présentent une symétrie bilatérale, tels que les yeux, les nerfs oculaires ou les oreilles internes. En attribuant des labels identiques à ces organes symétriques, nous avons pu exploiter cette régularité anatomique pour renforcer la capacité d'apprentissage du modèle. En considérant les organes symétriques comme des entités unifiées avec des labels identiques, le modèle a été en mesure de généraliser plus efficacement et de mieux comprendre les relations spatiales et structurelles entre les organes, améliorant ainsi la précision de la segmentation. Cette approche s'est avérée particulièrement pertinente dans le contexte de la segmentation d'images médicales, où la symétrie anatomique est souvent présente et peut fournir des informations supplémentaires pour guider le processus de segmentation.

Voici enfin à quoi ressemble la carte label finale

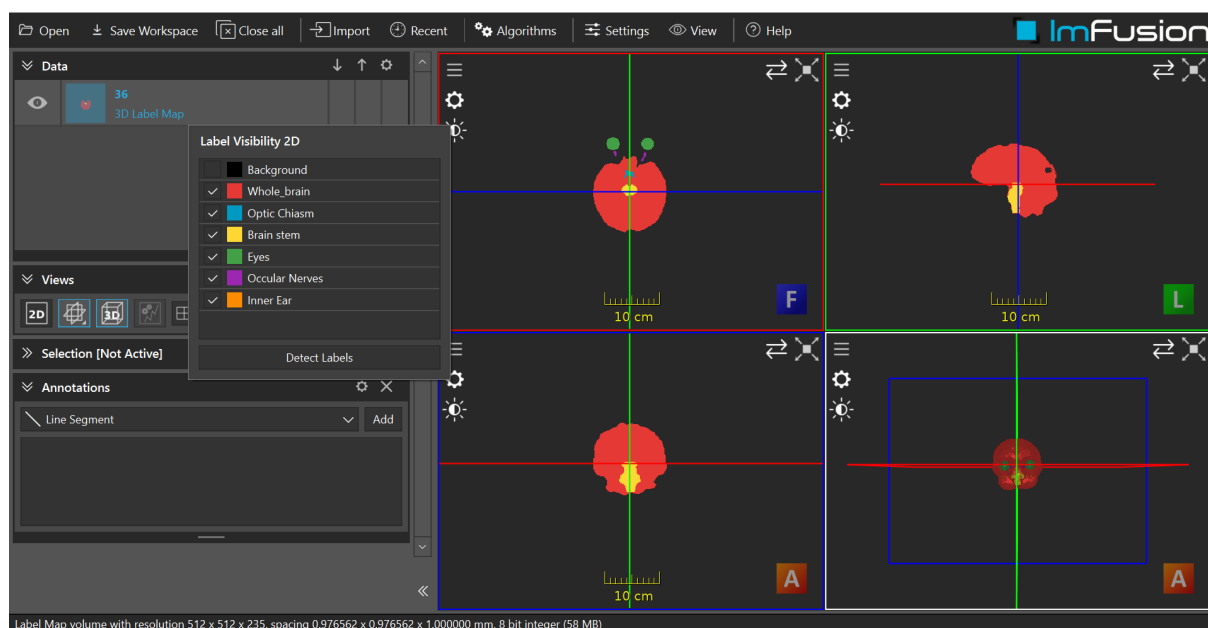


FIGURE 3.11 – Résultat de la label map finale.

Recalage de données

Il s'agit d'un processus de transformation géométrique qui vise à minimiser les différences spatiales entre les images en effectuant des translations, des rotations et des mises à l'échelle. Le recalage rigide suppose que les déformations entre les images sont principalement dues à des mouvements rigides, c'est-à-dire des déplacements globaux sans déformations locales importantes. Par conséquent, il ne permet pas de modéliser des déformations complexes telles que les déformations non linéaires. [31] [32] L'objectif principal du recalage rigide est d'aligner les images de sorte que les structures anatomiques d'intérêt se superposent le mieux possible. Le recalage rigide est réalisé en recherchant les paramètres de transformation (translation, rotation, mise à l'échelle) qui minimisent une mesure de similarité entre les images à recalibrer. Une fois que les paramètres de transformation optimaux sont trouvés, ils sont appliqués à l'image à recalibrer afin de l'aligner spatialement avec l'image de référence. Ainsi, le recalage rigide permet de créer un

alignement précis entre les images, ce qui facilite la comparaison et l'analyse ultérieure des données médicales. J'ai donc effectué un recalage rigide automatique à l'aide du logiciel Imfusion en suivant les étapes suivantes :

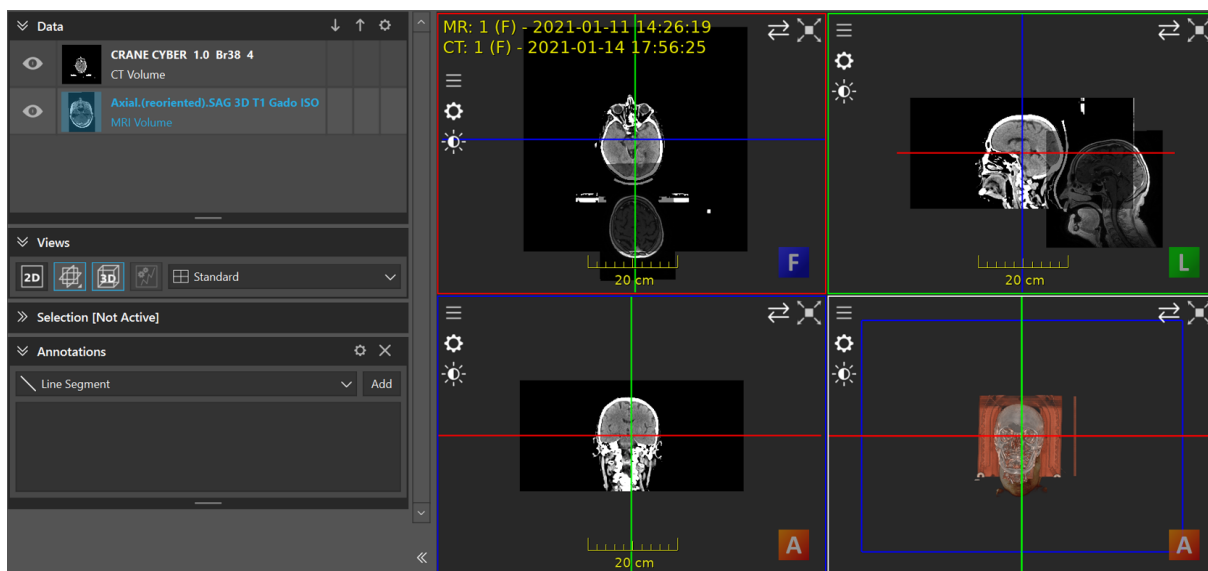


FIGURE 3.12 – Ouverture des deux images dans la même scène.

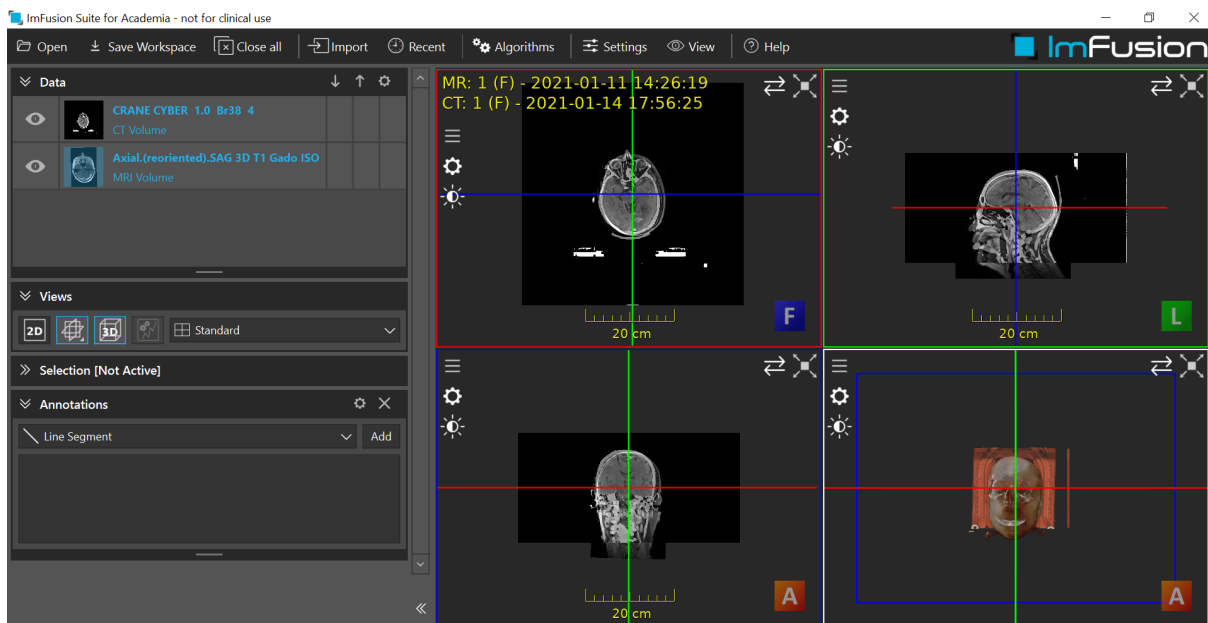


FIGURE 3.13 – Pré-rapprochement manuel approximatif de l'une deux images (IRM dans notre cas puisque la CT doit être fixe).

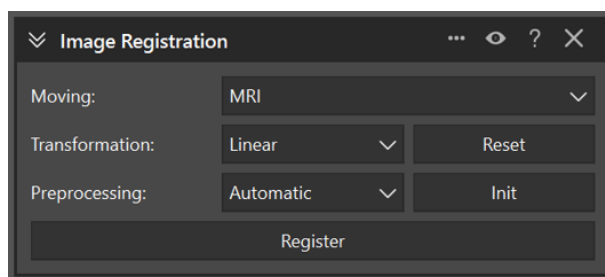


FIGURE 3.14 – Recalage.

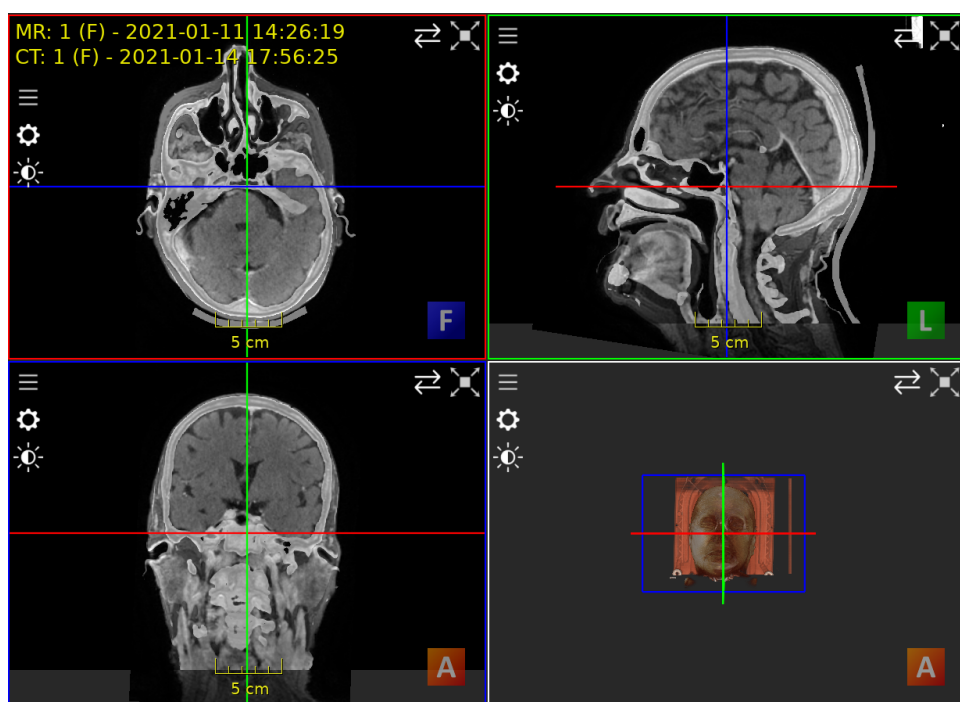


FIGURE 3.15 – Aperçu des deux images recalées superposées.

Normalisation de l'espace des images

- Une fois le processus de recalage terminé, il est apparu que les dimensions et l'espace des images différaient d'une modalité à une autre. Par conséquent, il était impératif de rétablir un espace cohérent afin de pouvoir entraîner un modèle sur les deux modalités en utilisant la même carte d'étiquetage. Pour réaliser cette opération, Nous avons utilisé l'outil ImFusion, Nous avons procédé à la réinitialisation de l'espace des images de manière à obtenir une cohérence entre les modalités. Cette étape était cruciale pour assurer la précision et la

fiabilité de l'entraînement du modèle, en garantissant une concordance précise entre les données et les étiquettes correspondantes. Les étapes suivantes :

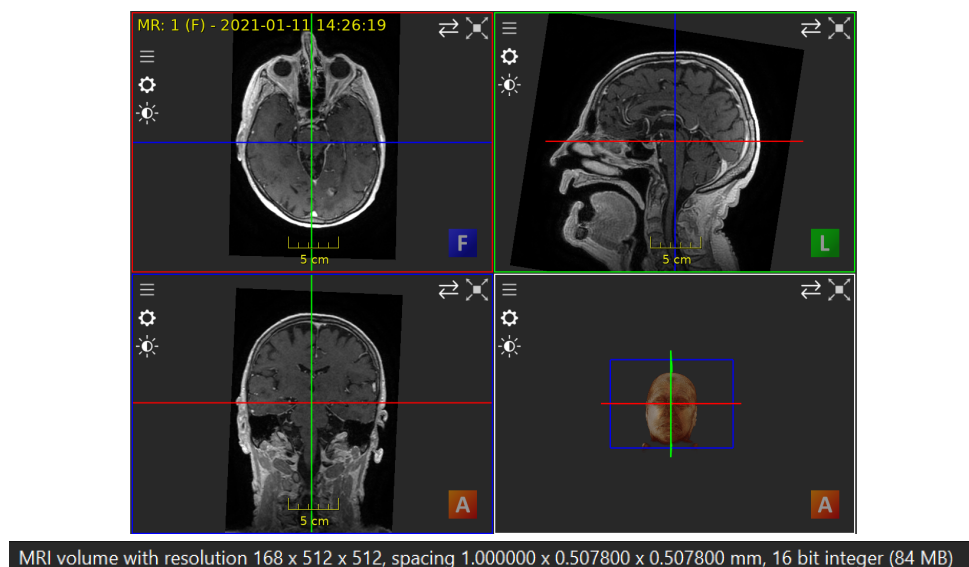


FIGURE 3.16 – Image IRM avant normalisation de l'espace.

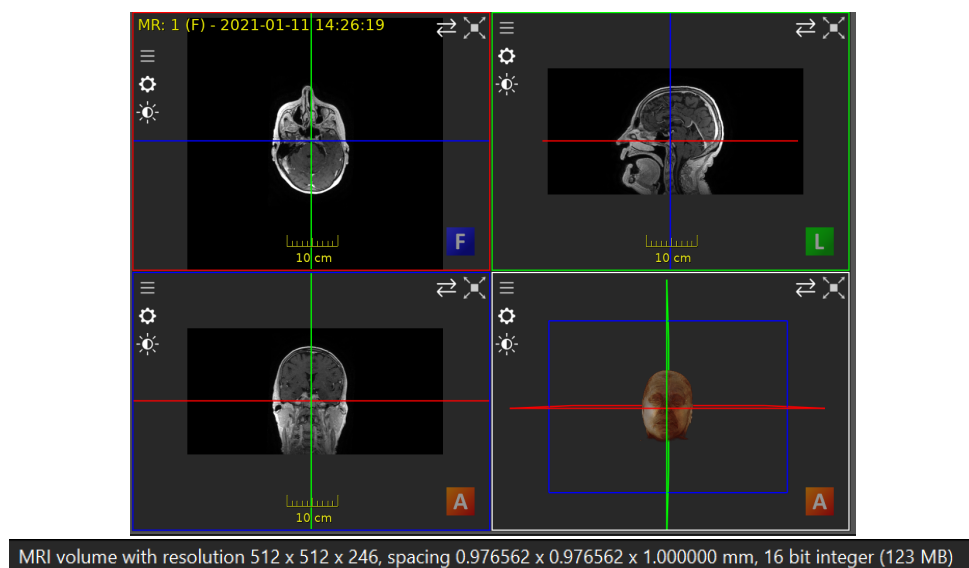


FIGURE 3.17 – Images IRM après normalisation de l'espace.

Normalisation des intensités des images

- Les images de modalités différentes peuvent avoir des plages d'intensité différentes, ce qui peut affecter les performances des modèles d'ap-

prentissage automatique. Pour remédier à cela, une normalisation des intensités des images entre 0 et 1 est effectuée pour rendre les plages d'intensité comparables entre les modalités.

Sauvegarde et enregistrement

- Une fois tous les pré-traitements faits, Nous avons procédé à l'exportation des images IRM et CT, qui étaient initialement stockées sous forme de coupes DICOM, vers le format de volume NIfTI. qui comme mentionné ci dessus est plus couramment utilisé dans le domaine de l'imagerie médicale en raison de ses avantages et de sa compatibilité avec de nombreux outils et bibliothèques de traitement d'images.

3.3 Architecture et solution proposée

3.3.1 Architecture Explorée

Une fois les données traitées et l'état de l'art réalisé, nous avons pu choisir le modèle le plus approprié pour la tâche de segmentation, applicable dans les délais du stage. Nous avons fait le choix d'utiliser l'architecture UNET, une architecture de réseau de neurones convolutionnels largement utilisée dans le domaine de la segmentation d'images médicales. Cette architecture, notée UNET, a été introduite par Ronneberger et al. [4] et s'est avérée très performante pour la segmentation sémantique.

Dans un premier temps, nous allons présenter le principe général de la méthode UNET et expliquer en détail la composition de l'architecture choisie et de la fonction de coût. Ensuite, nous aborderons l'application de cette méthode à nos données spécifiques, en décrivant notamment les données d'entrée et de sortie du modèle, ainsi que notre choix de séparation de la base de données en ensembles d'entraînement, de validation et de test. Nous discuterons également des hyperparamètres utilisés, du processus d'entraînement,

de la validation et des tests de notre modèle.

3.3.2 Principe générale de la méthode

3.3.3 Architecture utilisée

UNET est une architecture de réseau encodeur-décodeur en forme de U, qui se compose de quatre blocs encodeurs et de quatre blocs décodeurs connectés via un pont. Le réseau encodeur (voie contractante) réduit de moitié les dimensions spatiales et double le nombre de filtres (canaux de caractéristiques) à chaque bloc encodeur. De même, le réseau décodeur double les dimensions spatiales et divise par deux le nombre de canaux de caractéristiques. [4] [29]

3.3.3.1 Réseau Encodeur

Le réseau encodeur agit comme un extracteur de caractéristiques et apprend une représentation abstraite de l'image d'entrée à travers une séquence de blocs encodeurs. Chaque bloc encodeur se compose de deux convolutions 3x3, où chaque convolution est suivie d'une fonction d'activation ReLU (Rectified Linear Unit). La fonction d'activation ReLU introduit une non-linéarité dans le réseau, ce qui aide à une meilleure généralisation des données d'entraînement. La sortie de la ReLU agit comme une skip connection pour le bloc décodeur correspondant.

Ensuite, suit un max-pooling 2x2, où les dimensions spatiales (hauteur et largeur) des cartes de caractéristiques sont réduites de moitié. Cela réduit le coût computationnel en diminuant le nombre de paramètres entraîna-

3.3.3.2 Skip connections (Connexion de saut)

Ces skip connections fournissent des informations supplémentaires qui aident le décodeur à générer de meilleures caractéristiques sémantiques. Elles

agissent également comme une connexion de raccourci qui facilite le flux indirect des gradients vers les couches antérieures sans aucune dégradation. En termes simples, nous pouvons dire que la connexion de saut aide à un meilleur flux de gradient lors de la rétropropagation, ce qui permet au réseau d'apprendre une meilleure représentation.

3.3.3.3 Bridge (Pont)

Le pont connecte le réseau encodeur et le réseau décodeur et complète le flux d'informations. Il se compose de deux convolutions 3x3, où chaque convolution est suivie d'une fonction d'activation ReLU.

3.3.3.4 Réseau Décodeur

Le réseau décodeur est utilisé pour prendre la représentation abstraite et générer un masque de segmentation sémantique. Le bloc décodeur commence par une convolution transpose 2x2. Ensuite, il est concaténé avec la carte de caractéristiques de connexion de saut correspondante provenant du bloc encodeur. Ces connexions de saut fournissent des caractéristiques provenant de couches antérieures qui sont parfois perdues en raison de la profondeur du réseau. Ensuite, deux convolutions 3x3 sont utilisées, où chaque convolution est suivie d'une fonction d'activation ReLU.

La sortie du dernier décodeur passe par une convolution 1x1 avec une activation sigmoïdale. La fonction d'activation sigmoïdale donne le masque de segmentation représentant la classification pixel par pixel.

3.3.3.5 3D-UNET

Le 3D U-Net est une extension de l'architecture U-Net spécifiquement conçue pour la segmentation d'images médicales en 3D. Elle a été largement adoptée et a obtenu d'excellentes performances dans diverses applications d'imagerie médicale. [9]

L'architecture du 3D U-Net étend l'U-Net original en intégrant une dimension supplémentaire pour traiter les données volumétriques. Elle est couramment utilisée pour segmenter des organes, des tumeurs ou des structures dans des images médicales en 3D telles que des scanners CT ou des volumes IRM.

Pour améliorer la précision de segmentation du réseau, le 3D U-Net utilise souvent différentes techniques telles que la normalisation par lots (batch normalization), le dropout et les connexions résiduelles. De plus, il utilise généralement une fonction d'activation softmax à la couche de sortie pour produire une distribution de probabilité pour chaque classe.

L'entraînement d'un 3D U-Net nécessite un ensemble de données étiquetées d'images en 3D, où chaque image est associée à un masque de segmentation correspondant. Le réseau est entraîné à l'aide d'une fonction de perte telle que la perte de Dice ou la perte de l'entropie croisée, qui mesure la dissimilarité entre la segmentation prédite et la vérité terrain.

L'architecture 3D U-Net a montré des résultats prometteurs dans diverses tâches de segmentation d'images médicales. Sa capacité à capturer à la fois des informations contextuelles locales et globales en fait une méthode particulièrement adaptée pour traiter des données volumétriques.

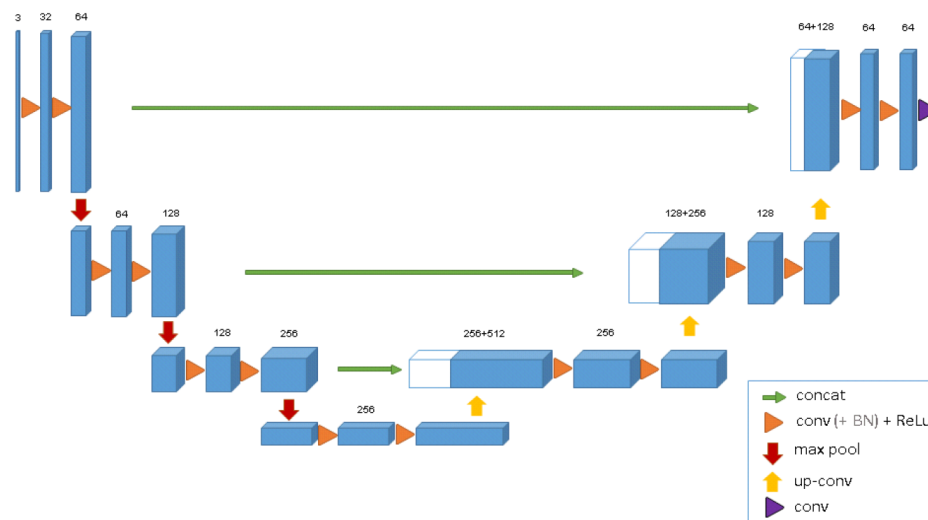


FIGURE 3.18 – Architecture 3D UNET.

3.3.3.6 Avantages

Les avantages les plus importants de l'architecture 3D U-Net par rapport à l'U-Net classique sont :

Capacité à gérer des données de volume 3D L'architecture U-Net 3D est conçue pour la segmentation d'images médicales 3D, telles que les tomodensitogrammes ou les volumes IRM. Il peut capturer des informations spatiales dans toutes les dimensions, aidant à mieux représenter l'anatomie complexe présente dans les données volumétriques.

Préservation de la résolution spatiale L'architecture U-Net 3D intègre des opérations de suréchantillonnage dans le chemin du décodeur, permettant de récupérer la résolution spatiale à partir des données volumétriques. Cette capacité est essentielle pour obtenir des résultats de segmentation détaillés et précis, en préservant les caractéristiques fines de l'anatomie dans les données 3D.

3.3.4 Fonction de coût

La Dice Loss est dérivée du coefficient de Dice de Sørensen-Dice, une mesure statistique développée dans les années 1940 pour évaluer la similarité entre deux échantillons. Elle a été introduite dans le domaine de la vision par ordinateur par Milletari et al. [28] en 2016 dans le contexte de la segmentation d'images médicales en 3D.

Le coefficient de Dice est utilisé pour quantifier le degré de chevauchement entre deux ensembles. Dans le cas de la détection de contours, les pixels prédits et les pixels de vérité terrain sont binaires, avec une valeur de 1 pour les pixels de contour et de 0 pour les pixels non contour. Le coefficient de Dice est calculé en comparant les pixels correspondants dans les deux ensembles. Le dénominateur du coefficient de Dice représente la somme totale des pixels

de contour dans les deux ensembles, tandis que le numérateur représente la somme des pixels de contour correctement prédits. [33]

$$L_{Dice} = 1 - \frac{2 \sum_{c=1}^C \sum_{i=1}^N g_i^c s_i^c}{\sum_{c=1}^C \sum_{i=1}^N g_i^c + \sum_{c=1}^C \sum_{i=1}^N s_i^c}$$

FIGURE 3.19 – Formule de dice loss.

Le coefficient de Dice peut également être interprété dans le cadre de la théorie des ensembles. Lorsque deux ensembles A et B se chevauchent parfaitement, le coefficient de Dice atteint sa valeur maximale de 1. À mesure que le chevauchement diminue, le coefficient de Dice diminue également, atteignant sa valeur minimale de 0 lorsque les ensembles ne se chevauchent pas du tout. Par conséquent, le coefficient de Dice est compris entre 0 et 1, et on peut utiliser 1 - Dice comme Dice Loss pour maximiser le chevauchement entre les ensembles.

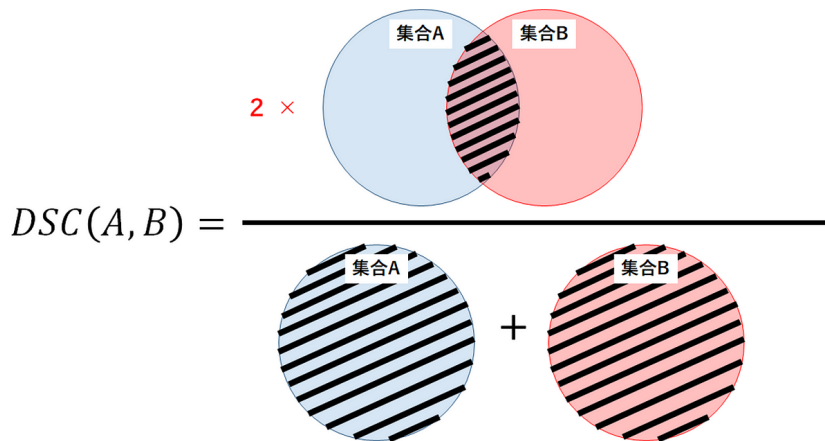


FIGURE 3.20 – Dice Overlap.

Dans les tâches de détection de contours, les pixels de contour prédits et les pixels de contour de vérité terrain peuvent être considérés comme deux ensembles distincts. En utilisant la Dice Loss, on cherche à entraîner ces ensembles à se chevaucher de plus en plus. La Dice Loss prend en compte à la fois l'information de chevauchement local et global en considérant le nombre

total de pixels de contour ainsi que le chevauchement entre les ensembles. Ceci est essentiel pour obtenir une segmentation précise des contours.

En résumé, la Dice Loss est une mesure de perte utilisée pour maximiser le chevauchement entre les pixels de contour prédits et les pixels de contour de vérité terrain. Elle intègre à la fois l'information de chevauchement local et global, permettant d'obtenir une segmentation précise des contours.

Cependant, dans notre quête d'améliorer les résultats, nous avons également exploré une autre fonction de perte appelée Dice Cross Entropy Loss. Cette fonction de perte combine les avantages de la mesure de similarité Dice et de la Cross Entropy Loss. La Dice Cross Entropy Loss est calculée en prenant la somme pondérée de la Dice Loss et de la Cross Entropy Loss. La cross-entropie (CE) est dérivée de la divergence de Kullback-Leibler (KL), qui est une mesure de la dissimilarité entre deux distributions. Dans le cadre des tâches courantes d'apprentissage automatique, la distribution des données est donnée par l'ensemble d'entraînement. Ainsi, minimiser la divergence de KL revient à minimiser la CE [33] [34]. La cross-entropie est définie par :

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C g_i^c \log s_i^c$$

Dans cette équation, N représente le nombre total d'échantillons ou de pixels, C représente le nombre de classes, g_i^c est un indicateur binaire qui vaut 1 si l'étiquette de classe c est la classification correcte pour le pixel i , et s_i^c est la probabilité prédite correspondante pour la classe c .

Chapitre 4

Validation expérimentale

4.1 Détails des bases de données et annotations

La base de données sur laquelle nous nous sommes appuyés comprend à la fois des images tomodensitométriques (todométriques) et des images par résonance magnétique (IRM). Les images tomodensitométriques sont fournies avec des contours au format RT Struct, qui correspond à une carte de labels.

Étant donné que les tumeurs peuvent se localiser à différents endroits du crâne, les organes à risque annotés varient d'un patient à l'autre. Cela représente un défi pour nous car la base de données est hétérogène. Bien que 51 sujets ne soit pas un nombre élevé, il est déjà acceptable en imagerie médicale, car l'un des principaux défis en imagerie médicale est le manque de données.

Dans notre étude, la base de données comprend 51 sujets, dont 30 hommes et 21 femmes, nous avons utilisé un sous-ensemble de la base de données comprenant 37 sujets (21 hommes et 16 femmes) pour l'entraînement de notre modèle de segmentation. Il convient de noter que le déséquilibre de genre pourrait potentiellement influencer les performances de segmentation, avec une tendance à une meilleure segmentation chez les sujets masculins en raison de différences anatomiques. Nous prendrons en compte cette variable dans notre analyse et évaluation des résultats. Nous tenterons également de gérer la limitation du nombre de sujets en utilisant des techniques telles que l'augmentation de données, dont nous discuterons ultérieurement.

L'ensemble de données d'IRM provenant de trois établissements médicaux distincts. Chacun de ces endroits a utilisé des protocoles d'acquisition différents, ce qui a entraîné la disponibilité de séquences d'IRM différentes entre les patients. Cette diversité dans les séquences d'acquisition est bénéfique car elle représente une opportunité intéressante pour évaluer l'efficacité de différents algorithmes de segmentation dans des conditions réelles à fournir des résultats fiables et généralisables à des séquences différentes.



FIGURE 4.2 – Lieux et séquences IRM

En revanche, en ce qui concerne les images de tomodensitométrie (CT), Nous avons pu bénéficier d'une plus grande uniformité, car elles ont été réalisées au même endroit et avec la même machine, ce qui peut faciliter la comparaison et la segmentation plus précise des structures anatomiques.

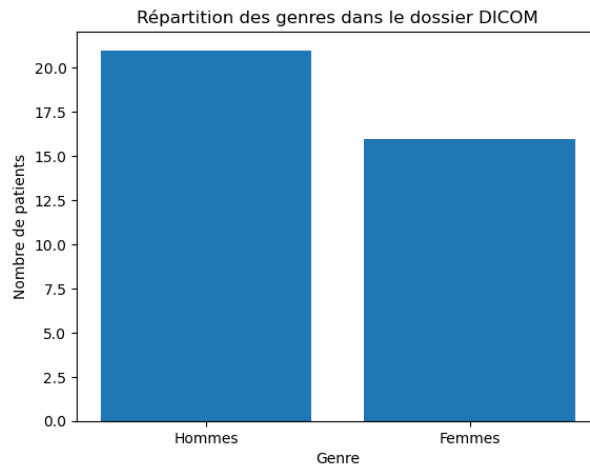


FIGURE 4.1 – Déséquilibre hommes/femmes de la base de données.

4.2 Application de la méthode à nos données

4.2.1 Données d'entrée et de sortie du modèle et prétraitement

Après les pré-traitements et le choix des organes à risque à segmenter il ne nous reste plus que 41 images .

L'architecture que nous avons utilisée nécessite le traitement de l'image d'en-

trée sous forme d'images 3D. Les résolutions spatiales étaient d'environ 512x512x254, mais elles pouvaient varier, principalement pour la troisième dimension. Une fois redimensionnées à 128, les intensités des voxels, qui étaient précédemment comprises entre 0 et 10 000, sont maintenant comprises entre 0 et 1.

La segmentation des organes à risques du crâne est formulée comme une tâche de multisegmentation. Elle reçoit en entrée un volume IRM ou CT prétraité et fournit en sortie 7 canaux, comprenant les 6 organes à risques et l'arrière-plan. Pendant l'entraînement, l'utilisation de l'architecture UNET nécessite également les cartes de labels (ou labels-maps) des organes annotés manuellement. Les labels-maps sont des matrices où chaque voxel est associé à un numéro compris entre 0 et 6 (l'arrière-plan étant représenté par 0).

En ce qui concerne les données de sortie du modèle, comme illustré dans la figure, le modèle est censé renvoyer une carte de labels complète des organes prédits, tout comme les labels-maps.

4.2.2 Augmentation de données

L'augmentation de données, connue sous le nom de "Data augmentation" en anglais, est une approche largement utilisée pour accroître le volume de données d'apprentissage en générant des variantes légèrement modifiées des données existantes ou en créant de nouvelles données synthétiques basées sur les données disponibles. Son objectif principal est de lutter contre le surapprentissage en introduisant une variabilité supplémentaire dans l'ensemble de données, ce qui permet au modèle d'apprentissage automatique de généraliser de manière plus efficace.

Dans le cadre de notre étude portant sur des données anatomiques réelles, il est crucial de veiller à ce que les augmentations de données restent cohérentes avec les caractéristiques anatomiques humaines et les plages de mouvement possibles dans nos images. Ainsi, nous avons appliqué des transformations spécifiques tout en respectant les limites physiologiques et anatomiques.

Ces transformations incluent principalement des rotations aléatoires et des retournements afin de simuler différentes orientations spatiales et perspectives d'observation. Les figures 3.2 et 3.3 illustrent respectivement l'opération de retournement (flip) effectuée pour la data augmentation.

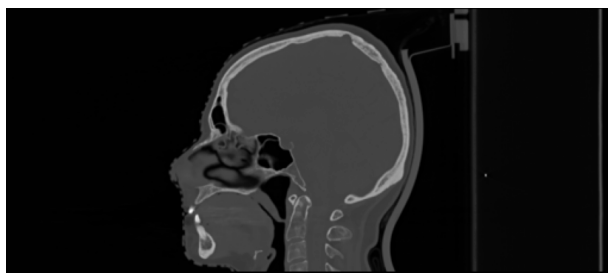


FIGURE 4.3 – Sujet numéro 44 dans sa position originale .



FIGURE 4.4 – Sujet numéro 44 dans sa position retournée.

4.2.3 Séparation de la base de données

Sur une base de données comprenant 51 sujets, seuls 41 ont été annotés pour les organes à risque qui nous intéressent, et présentés sous format RT struct. Comme expliqué dans la section 2.3.1, nous avons veillé à les convertir en cartes d'étiquettes.

Nous avons donc 41 CT/IRM corrects que nous avons choisi de répartir selon le ratio de 80/10/10 train/val/Test.

Nous avons divisé notre base de données en 33 sujets pour l'ensemble d'entraînement (train). Ce dernier a été spécifiquement utilisé pour permettre

au modèle de s'entraîner sur les CT/IRM des organes à risque du crâne et les comparer avec leurs annotations.

L'ensemble de validation (validation) se compose de 4 sujets. Cette base de données a été utilisée pour les validations effectuées tout au long de l'entraînement. La validation était réalisée sur les CT/IRM et utilisait leurs segmentations manuelles associées à des métriques pour enregistrer le modèle présentant les meilleurs résultats. Cette étape était effectuée tous les 5 epochs d'entraînement.

Enfin, l'ensemble de test (test) est constitué de 4 sujets, comprenant deux hommes et deux femmes. Nous avons choisi de conserver ces données afin de tester le modèle formé sur des données qu'il ne connaissait pas. Ces données ont été utilisées pour valider notre modèle.

4.2.4 Entraînement et validation

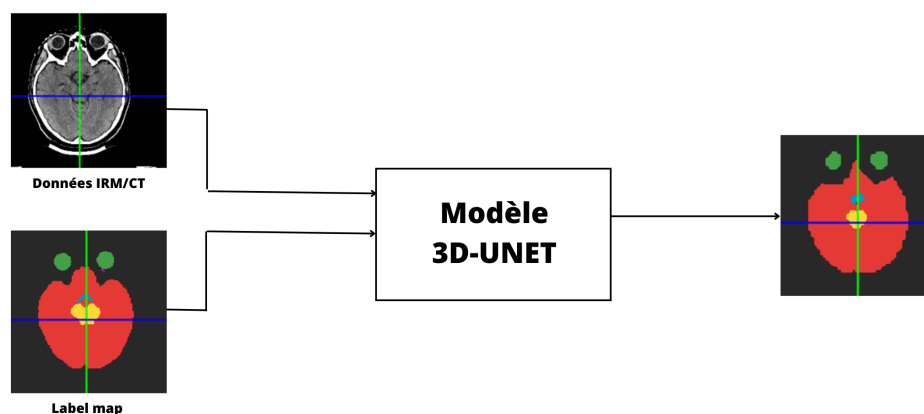


FIGURE 4.5 – Principe de l'entraînement du modèle.

4.2.4.1 Entraînement sur CT

Nous avons effectué la segmentation d'images médicales à l'aide du réseau de neurones convolutif UNet. Notre travail s'est concentré sur les images CT dont les intensités étaient normalisées entre 0 et 1 et redimensionnées à une résolution de 128. Le processus d'entraînement a été réalisé sur une durée

de 150 epochs afin de permettre une convergence adéquate du modèle. Pour prévenir les problèmes de surapprentissage, nous avons utilisé une fonction de perte basée sur la mesure Dice ainsi que la Dice Cross Entropy et implémenté un batch size de 1. Le taux d'apprentissage initial a été fixé à 0,001, et nous avons utilisé l'optimiseur Adam pour ajuster les poids du réseau. Notre modèle a été entraîné pour reconnaître sept classes distinctes, comprenant le fond et six structures d'intérêt spécifiques. Afin d'évaluer la performance de notre modèle de segmentation, nous avons effectué une validation régulière tout au long de l'entraînement. Pour ce faire, nous avons utilisé un ensemble de données distinct comprenant des images CT de quatre sujets différents. La validation a été réalisée toutes les cinq epochs, ce qui nous a permis de surveiller l'évolution des métriques de performance et d'ajuster notre modèle si nécessaire.

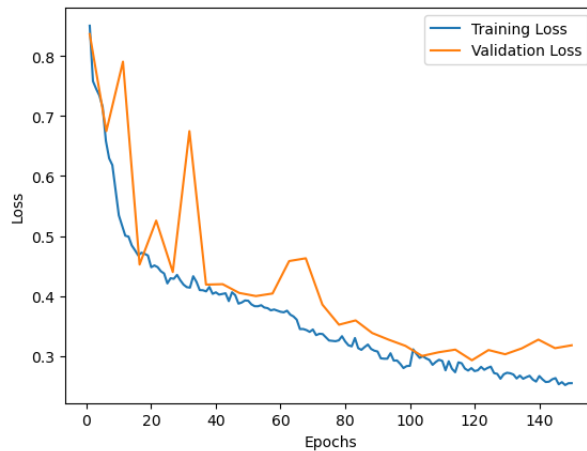


FIGURE 4.6 – Courbes de métrique Dice lors de l'entraînement et de la validation sur une base de données d'entraînement sur CT de 37 sujets.

L'analyse des courbes de validation et d'entraînement nous permet d'obtenir des indications précieuses sur la performance et la progression de notre modèle de segmentation. Dans un premier temps, nous observons que ces courbes présentent une décroissance continue, ce qui témoigne de l'apprentissage efficace du modèle. Cependant, à partir de la 100ème epoch, nous remarquons un comportement distinct entre les deux courbes. En effet, la

courbe de validation se stabilise, tandis que la courbe d'entraînement continue de décroître. Cette observation est significative, car elle suggère l'amorce d'un phénomène de surapprentissage, où le modèle apprend spécifiquement les caractéristiques des données d'entraînement au détriment de sa généralisation sur de nouvelles données. À ce stade, nous prenons la décision d'interrompre l'apprentissage afin de prévenir toute dégradation de la performance du modèle sur des données inconnues.

4.2.4.2 Entraînement sur IRM

Nous avons également appliqué la méthode de segmentation d'images médicales à l'aide du réseau de neurones convolutif UNet, mais cette fois-ci sur des images IRM. Pour garantir la comparabilité avec les résultats obtenus avec les images CT, nous avons redimensionné les images IRM de manière à avoir la même résolution de 128 pixels. De plus, nous avons normalisé les intensités des images IRM entre 0 et 1, tout comme nous l'avions fait pour les images CT. Le processus d'entraînement a été mené sur une durée de 150 epochs, de manière à assurer une comparabilité des deux modalités. Nous avons appliqué les mêmes principes, méthodes fonctions coût et hyperparamètres utilisés pour les images CT, mais cette fois-ci sur des images IRM.

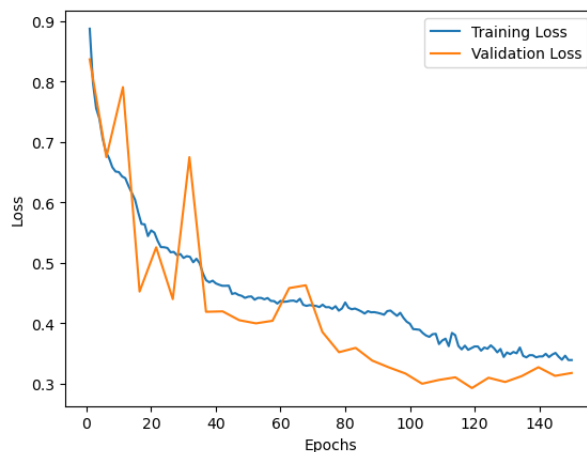


FIGURE 4.7 – Courbes de métrique Dice lors de l'entraînement et de la validation sur une base de données d'entraînement sur IRM de 37 sujet.

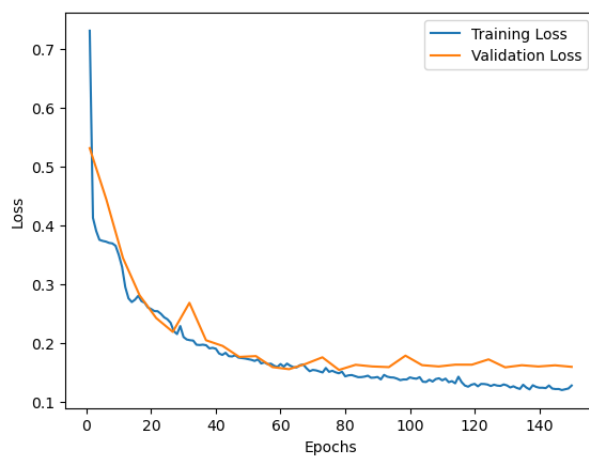


FIGURE 4.8 – Courbes de métrique Dice Cross Entropy lors de l’entraînement et de la validation sur une base de données d’entraînement sur CT de 37 sujet

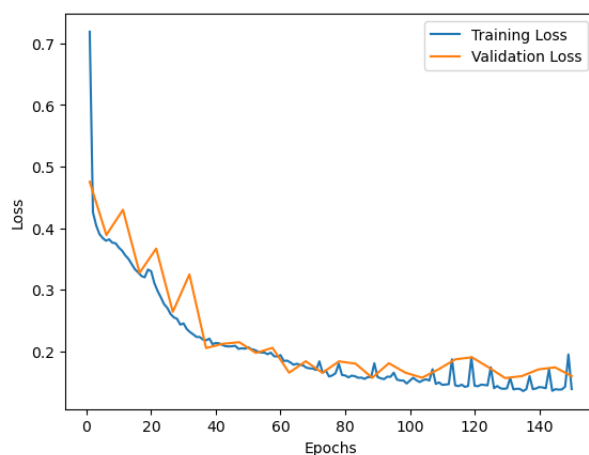


FIGURE 4.9 – Courbes de métrique Dice Cross Entropy lors de l’entraînement et de la validation sur une base de données d’entraînement sur IRM de 37 sujet

4.2.5 Entraînement sur les deux modalités

Dans le but d’explorer les capacités d’apprentissage du modèle à partir de différentes modalités de manière simultanée, nous avons entrepris un entraînement combiné en utilisant à la fois des images CT et des images IRM, accompagnées de leurs cartes de référence respectives. L’aspect distinctif de cet entraînement réside dans le fait que l’encodeur est exposé à des modalités d’entrée différentes. Les images CT et les images IRM présentent des

caractéristiques distinctes telles que le contraste, la résolution spatiale et la représentation des tissus anatomiques. Cependant, malgré ces différences, le décodeur, qui est chargé de générer les prédictions de segmentation, est identique pour les deux modalités. L'objectif sous-jacent était d'évaluer la capacité du modèle à acquérir une représentation générale des structures anatomiques en exploitant à la fois les similitudes et les différences entre les images CT et les images IRM.

Dans cette perspective, nous avons procédé à un entraînement utilisant 17 CT et 18 IRM, regroupant ainsi un total de 37 sujets, conformément aux deux autres entraînements précédents utilisant la fonction de perte Dice cross entropy qui a démontré une meilleure capacité d'apprentissage. La courbe de perte obtenue présente les caractéristiques suivantes.

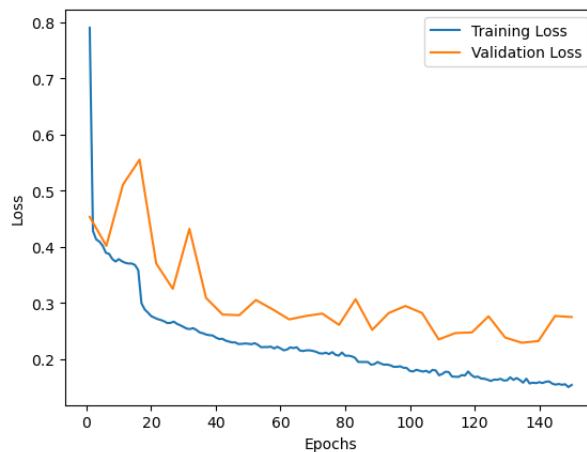


FIGURE 4.10 – Courbes de métrique Dice Cross Entropy lors de l'entraînement et de la validation sur une base de données d'entraînement sur un mélange de 37 IRM ET CT

4.2.6 Test

Nous avons effectué plusieurs tests en utilisant deux modalités d'imagerie, à savoir la tomodensitométrie (CT) et l'imagerie par résonance magnétique (IRM). L'objectif de ces tests est de réaliser un entraînement sur les données CT, puis de les tester sur des images CT, un autre entraînement est réalisé sur les données IRM et testé sur la même modalité.

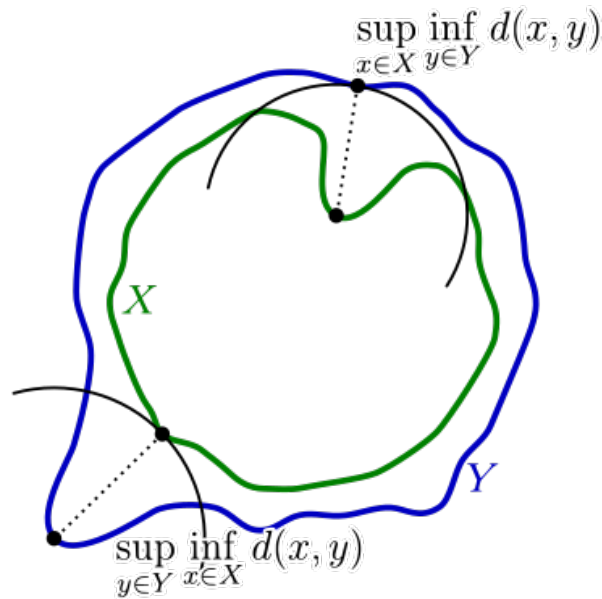
L'objectif de ces expérimentations est de déterminer quelle modalité d'imagerie fournit le plus d'informations pour la segmentation, en supposant que les données sont normalisées à la même échelle.

Un dernier test a été réalisé en évaluant le modèle formé sur les deux modalités simultanément, puis testé séparément sur chaque modalité distincte. L'objectif de cette étude était d'évaluer la capacité du modèle à généraliser lorsque les données d'entrée étaient inhomogènes tandis que les données de sortie restaient identiques.

4.2.7 Métriques utilisées

Dans le domaine de la segmentation d'images médicales, l'évaluation quantitative des résultats obtenus est cruciale pour évaluer l'efficacité et la précision des algorithmes. Pour mesurer la performance des méthodes de segmentation, il est courant d'utiliser différentes métriques qui permettent de quantifier objectivement la qualité des résultats obtenus. Ces métriques fournissent des informations essentielles sur la similarité entre la segmentation obtenue et la vérité terrain.

Une métrique utilisée fréquemment dans la segmentation d'images médicales est la distance de Hausdorff moyenne (Average Hausdorff Distance en anglais). [35] Cette métrique évalue la dissimilitude spatiale entre deux ensembles de points, dans notre cas, la segmentation automatique et la vérité terrain. La distance de Hausdorff mesure la plus grande distance entre chaque point d'un ensemble à son point le plus proche de l'autre ensemble, et la distance de Hausdorff moyenne est simplement la moyenne de ces distances pour tous les points de l'ensemble.



$$d_H(X, Y) = \max \left\{ \sup_{x \in X} d(x, Y), \sup_{y \in Y} d(X, y) \right\},$$

FIGURE 4.11 – Composants du calcul de la distance de Hausdorff entre la courbe verte X et la courbe bleue Y.

La distance de Hausdorff moyenne est particulièrement importante dans de nombreux contextes médicaux car elle permet d'évaluer la correspondance spatiale entre la segmentation automatique et la vérité terrain de manière globale. Une faible valeur de cette métrique indique une segmentation précise, où les contours des structures d'intérêt sont bien alignés avec ceux de la vérité terrain. En revanche, une valeur élevée indique une discordance spatiale importante, suggérant des erreurs significatives dans la segmentation.

Cependant, il est essentiel de noter qu'utiliser une seule métrique pour évaluer la segmentation peut être limitatif. Les images médicales sont complexes et présentent souvent différentes structures anatomiques, chacune nécessitant une évaluation spécifique. Par conséquent, il est recommandé d'utiliser plusieurs métriques différentes et complémentaires pour obtenir une évaluation plus complète et fiable des performances de segmentation. Il est important de calculer différentes métriques telles que le Dice coefficient, la précision, la

sensibilité et la spécificité. [36] Chacune de ces métriques apporte une information spécifique sur la performance de notre algorithme de segmentation, ce qui est essentiel pour évaluer sa précision et son efficacité.

Les métriques présentées reposent toutes sur le calcul d'une matrice de confusion pour un masque de segmentation binaire. Cette matrice comprend les nombres de prédictions de vrais positifs (VP), de faux positifs (FP), de vrais négatifs (VN) et de faux négatifs (FN). Les plages de valeurs de toutes les métriques présentées s'étendent de zéro (le pire) à un (le meilleur).

		Ground truth	
		FTU (1)	Background (0)
Prediction	FTU (1)	TP	FP
	Background (0)	FN	TN

FIGURE 4.12 – Matrice de confusion.

Dice coefficient Le Dice coefficient mesure la similarité entre la segmentation automatique et la vérité terrain en calculant le rapport entre la taille de l'intersection et la taille des ensembles. Il fournit une mesure globale de la précision de la segmentation, où une valeur proche de 1 indique une segmentation précise et une valeur proche de 0 indique une mauvaise segmentation. Le Dice coefficient est particulièrement utile pour évaluer l'alignement des contours des structures d'intérêt entre la segmentation automatique et la vérité terrain.

$$DICE = \frac{2 |S_g^1 \cap S_t^1|}{|S_g^1| + |S_t^1|} = \frac{2TP}{2TP + FP + FN}$$

Précision La précision mesure la proportion de vrais positifs parmi tous les pixels ou voxels étiquetés comme positifs par la segmentation automatique. Une valeur élevée de la précision indique que la segmentation est capable d'identifier correctement les régions d'intérêt sans trop de faux positifs. Cela est particulièrement important dans les applications médicales, où une fausse identification de régions d'intérêt peut avoir des conséquences graves. [35]

Sensibilité (rappel ou taux de vrais positifs) La sensibilité mesure la proportion de vrais positifs identifiés par rapport à tous les pixels ou voxels réellement positifs présents dans la vérité terrain. Une sensibilité élevée indique que la segmentation est capable de détecter correctement les régions d'intérêt et qu'elle est sensible aux détails les plus fins.

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

Spécificité La spécificité mesure la proportion de vrais négatifs identifiés par rapport à tous les pixels ou voxels réellement négatifs présents dans la vérité terrain. Une spécificité élevée indique que la segmentation est capable de différencier efficacement les régions d'intérêt des régions non pertinentes, réduisant ainsi le nombre de faux positifs.

$$Specificity = TNR = \frac{TN}{TN + FP}$$

4.3 Résultats

4.3.1 Résultats quantitatifs

Cette section présente les résultats quantitatifs, c'est-à-dire les résultats des métriques appliquées à nos prédictions lors du test. Ainsi, les tableaux

de la figure 3.2 présentent les résultats des métriques obtenues à l’aide d’un modèle entraîné sur 37 sujets.

D’abord sur les IRM, nous avons observé que les trois organes les mieux segmentés en terme de métrique de Dice étaient le cerveau, le tronc cérébral et les yeux, qui sont les structures les plus grandes. En revanche, les structures plus petites, telles que le chiasma optique et les nerfs oculaires, présentaient des défis particuliers en termes de segmentation ce qui est compréhensible parce que le Dice a tendance à être biaisé par la taille de l’organe, plus il est petit, moins on a de chance d’avoir un Dice score important. Leur petite taille rendait donc leur détection et leur segmentation plus difficiles. De plus, ces petites structures pouvaient présenter une forme et une texture plus complexes, ce qui compliquait davantage leur segmentation.

Organes	dice	hausdorff	avg_surface_dis	precision	sensitivity	specificity
Cerveau	0,94	98,35	4,19	0,95	0,92	0,998847
Chiasma optique	0,19	12,37	1,16	0,56	0,12	0,999998
Tronc cérébrale	0,77	11,58	1,76	0,88	0,68	0,999962
Yeux	0,81	5,48	1,33	0,85	0,77	0,999968
Nerfs oculaires	0,00	170,19	78,88	0,00	0,77	0,985124
Oreille interne	0,70	3,74	0,94	0,67	0,75	0,999990

FIGURE 4.13 – Tableau des résultats des métriques obtenus lors de l’entraînement sur 37 sujets en utilisant la loss Dice sur des images IRM.

Les résultats obtenus sur les images CT étaient également bons, voire meilleurs, pour les grandes structures. Surtout, nous avons constaté une amélioration notable pour les petites structures. Nous pensons que cela est dû au fait que les segmentations initiales étaient réalisées sur les images CT, et donc plus précises par rapport aux IRM, et que malgré un recalage rigide, il y a pu avoir des erreurs de localisation pour les petites structures, ce qui a induit notre modèle en erreur lors de la segmentation sur IRM.

Il est également important de noter que les métriques utilisées sont complémentaires et fournissent des informations différentes. Bien que le coefficient

de Dice pour les petits organes soit relativement plus faible que celui des grandes structures, la Average Hausdorff Distance en voxel demeure meilleure pour les premiers. Ce qui peut vouloir dire que les segmentations ont bien été faites, mais qu’il y a un léger décalage entre l’emplacement. Les valeurs élevées de sensibilité observées pour le cerveau, le tronc cérébral, les yeux et l’oreille interne témoignent de la capacité précise du modèle à identifier les régions positives correspondant à ces organes dans l’image.

Cependant, des valeurs de sensibilité moins bonnes sont observées pour le chiasma optique et les nerfs optiques, ce qui indique que le modèle éprouve des difficultés à retrouver ces régions spécifiques dans l’image.

En ce qui concerne la spécificité, on constate que les valeurs sont très élevées pour tous les organes étudiés. Cela suggère que notre modèle a tendance à sous-segmenter les régions d’intérêt, c’est-à-dire qu’il parvient à identifier correctement les vrais négatifs, c’est-à-dire les régions ou les pixels qui n’appartiennent pas à la classe d’intérêt.

Organes	dice	hausdorff	avg_surface_dis	precision	sensitivity	specificity
Cerveau	0,93	64,67	2,02	0,98	0,89	0,99
Chiasma optique	0,37	11,75	0,85	0,88	0,23	0,99
Tronc cérébrale	0,74	12,00	2,05	0,77	0,71	0,99
Yeux	0,77	7,00	1,47	0,88	0,68	0,99
Nerfs oculaires	0,19	20,62	0,93	0,61	0,12	0,99
Oreille interne	0,58	8,31	0,91	0,79	0,45	0,99

FIGURE 4.14 – Tableau des résultats des métriques obtenus lors de l’entraînement sur 37 sujets en utilisant la loss Dice sur des images CT.

Afin de vérifier cette hypothèse, nous avons essayé de modifier notre approche en utilisant une fonction de coût qui est un mélange de l’entropie croisée et de l’indice de Dice. Selon [33] cette fonction de perte est plus adaptée à la segmentation d’organes de tailles différentes. Nous avons remarqué une nette amélioration pour les petits organes, tant sur les images CT que sur les images IRM, tandis que les performances pour les grandes structures sont restées inchangées. Cette constatation représente déjà un résultat prometteur

dans notre étude.

Organes	dice	hausdorff	avg_surface_dis	precision	sensitivity	specificity
Cerveau	0,96	25,79	1,31	0,96	0,96	0,999
Chiasma optique	0,50	12,45	0,90	0,73	0,39	0,99
Tronc cérébrale	0,76	10,68	1,74	0,82	0,71	0,99
Yeux	0,77	7,00	1,47	0,89	0,68	0,99
Nerfs oculaires	0,48	13,75	0,87	0,57	0,41	0,99
Oreille interne	0,60	5,20	0,82	0,82	0,48	0,99

FIGURE 4.15 – Tableau des résultats des métriques obtenus lors de l’entraînement sur 37 sujets en utilisant la loss Dice Cross Entropy sur des images IRM.

Organes	dice	hausdorff	avg_surface_dis	precision	sensitivity	specificity
Cerveau	0,95	25,34	1,40	0,96	0,95	1,00
Chiasma optique	0,53	12,37	0,89	0,71	0,42	1,00
Tronc cérébrale	0,75	9,49	2,03	0,75	0,75	1,00
Yeux	0,77	7,28	1,37	0,94	0,65	1,00
Nerfs oculaires	0,17	12,85	0,80	0,66	0,10	1,00
Oreille interne	0,59	7,35	0,89	0,78	0,47	1,00

FIGURE 4.16 – Tableau des résultats des métriques obtenus lors de l’entraînement sur 37 sujets en utilisant la loss Dice Cross Entropy sur des images CT.

Nous avons également effectué des tests en utilisant le modèle entraîné sur les IRM pour les CT, ainsi que le modèle entraîné sur les CT pour les IRM, et il n’est pas surprenant de constater que le modèle n’était pas capable de généraliser efficacement d’une modalité à une autre sans une exposition préalable. Cependant, en utilisant un modèle entraîné sur les deux modalités simultanément, nous avons observé une capacité de généralisation assez bonne (en fonction du nombre de données d’entrée) lorsqu’il était testé sur les deux modalités différentes.

Organes	dice	hausdorff	avg_surface_dis	precision	sensitivity	specificity
Cerveau	0,91	28,09	2,85	0,98	0,85	0,999578
Chiasma optique	0,37	12,08	0,90	0,71	0,25	0,999998
Tronc cérébrale	0,74	11,92	1,53	0,96	0,61	0,999989
Yeux	0,45	10,49	2,99	0,99	0,29	0,999999
Nerfs oculaires	0,05	22,18	3,06	0,08	0,04	0,999990
Oreille interne	0,49	8,06	1,10	0,73	0,36	0,999996

FIGURE 4.17 – Tableau des résultats des métriques obtenus lors de l’entraînement sur 37 sujets en utilisant la loss Dice Cross Entropy sur IRM et CT testé sur IRM.

Organes	dice	hausdorff	avg_surface_dis	precision	sensitivity	specificity
Cerveau	0,94	24,54	1,72	0,98	0,91	0,9994770288
Chiasma optique	0,28	16,03	1,17	0,54	0,19	0,9999970198
Tronc cérébrale	0,76	12,37	1,71	0,81	0,73	0,9999312758
Yeux	0,71	8,06	1,64	0,96	0,56	0,9999949336
Nerfs oculaires	0,36	16,76	1,03	0,47	0,29	0,9999927878
Oreille interne	0,00	inf	inf	nan	0,00	1

FIGURE 4.18 – Tableau des résultats des métriques obtenus lors de l’entraînement sur 37 sujets en utilisant la loss Dice Cross Entropy sur IRM et CT testé sur CT.

4.3.2 Résultats qualitatifs

Tout d’abord, il est important de souligner que les résultats de segmentation prédits par le modèle d’apprentissage profond demeurent cohérents avec l’anatomie humaine.

Les performances de segmentation étaient globalement satisfaisantes pour les structures anatomiques de grande taille, mais moins précises pour les structures plus petites, telles que le chiasma optique.

Afin d’améliorer ces performances, nous avons exploré différentes fonctions de coût (loss). Nous avons observé qu’en utilisant une loss différente, Nous avons pu obtenir une amélioration notable des résultats.

Lors de l’évaluation des résultats obtenus sur les images CT en utilisant la première fonction de coût, nous avons constaté la présence d’« îles » dans certaines régions segmentées. Les « îles » font référence à des régions isolées qui ne sont pas directement connectées au reste de la structure segmentée.

D’autre part, une analyse plus approfondie des autres images et fonctions de coût a révélé la présence récurrente d’un problème de sous-segmentation. La sous-segmentation se manifeste lorsque certaines parties de la structure ne sont pas correctement identifiées ou incluses dans la segmentation finale. Il est important de noter que ce problème de sous-segmentation était plus prévalent que celui des « îles ».

Ces observations peuvent être expliquées par le fait que, sur les images d’entrées, certains tissus peuvent présenter des niveaux de contraste simi-

lares, ce qui rend difficile la distinction entre les différentes structures. Cela peut conduire à des sur-segmentations ou à des sous-segmentations des régions d'intérêt.

Ces résultats soulignent la complexité de la tâche de segmentation des structures anatomiques dans les images médicales et mettent en évidence la nécessité de poursuivre les recherches pour développer des approches plus avancées afin de résoudre ces problèmes.

Il est aussi important de souligner que nous avons également effectué des tests en utilisant le modèle préalablement entraîné sur les images CT sur les images IRM, et vice versa. Comme prévu, le modèle a rencontré des difficultés significatives dans la segmentation des images d'une modalité différente de celle sur laquelle il avait été entraîné.

Cependant, en utilisant le modèle entraîné sur le mélange d'IRM et de CT, dont nous avons mentionné précédemment, des résultats encourageants ont été observés. Malgré la nature non homogène de la base de données, le modèle a démontré une capacité notable à généraliser et à identifier des similitudes entre les deux modalités, permettant ainsi une segmentation relativement précise.

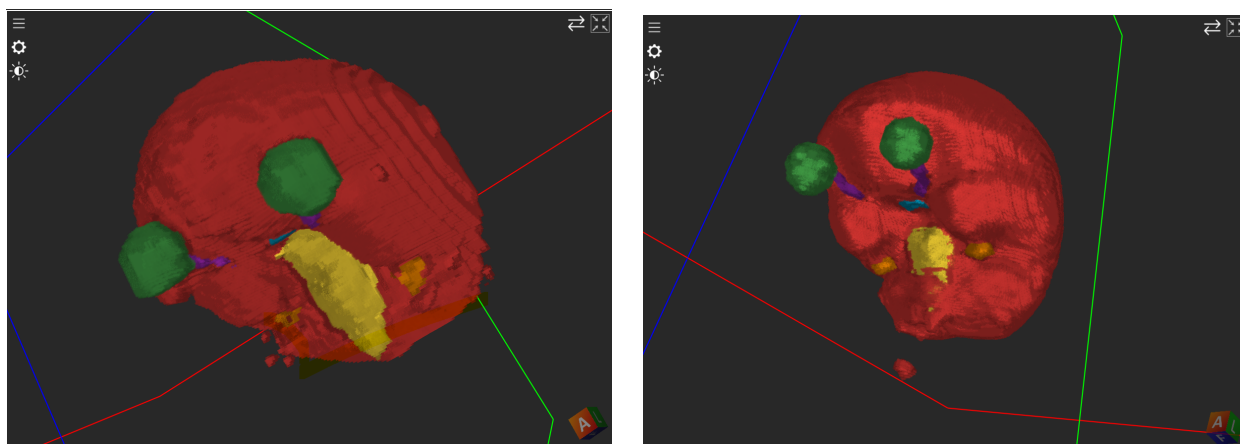


FIGURE 4.19 – Présence d'îles sur les segmentations des CT

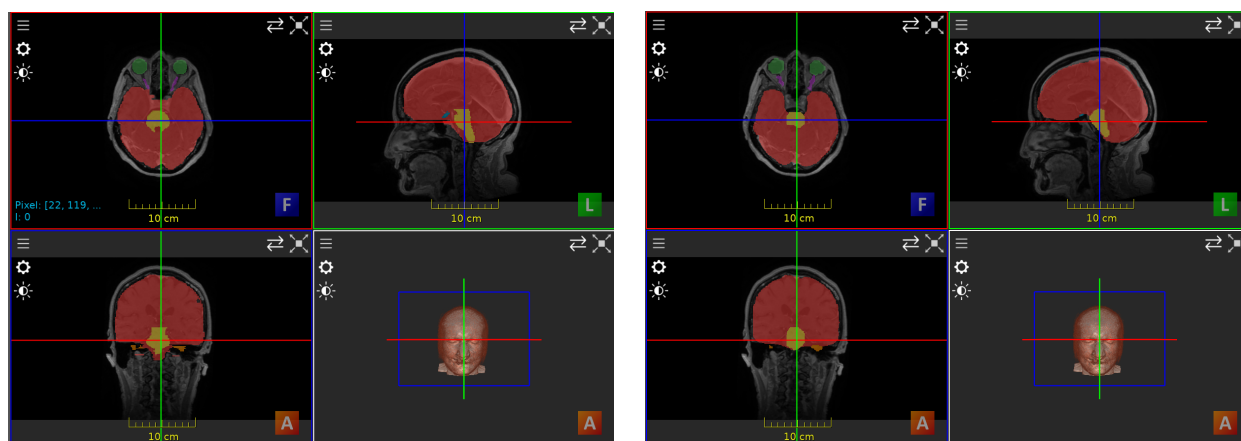
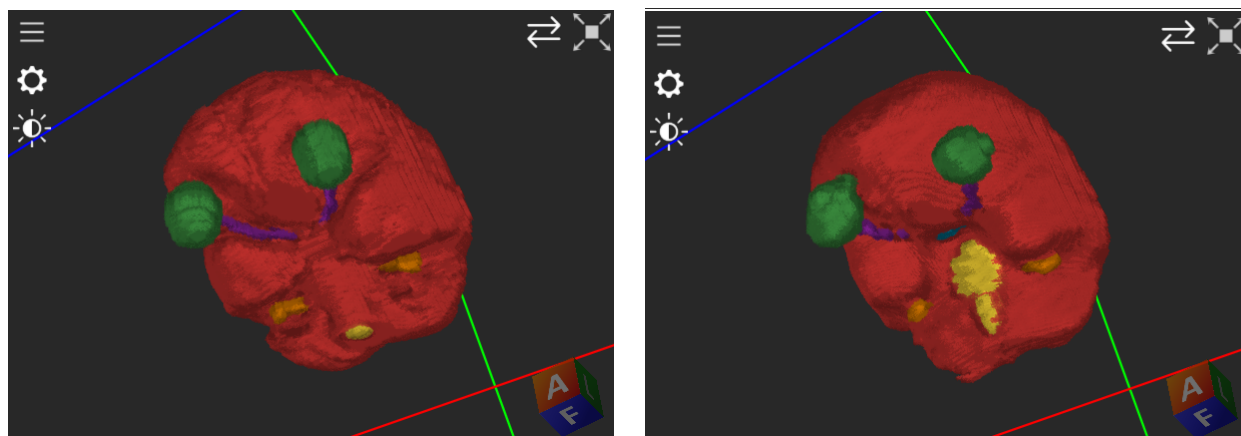


FIGURE 4.20 – A gauche superposition de la vérité terrain et de l'image CT, à droite la superposition de notre prédiction (Dice Cross Entropy) et de l'image CT



(a)

(b)

FIGURE 4.21 – Exemple de sous-segmentation (a) Véritée terrain (b) Prédiction

4.3.3 Discussion et perspectives d'amélioration

Les résultats obtenus dans le cadre de notre projet de segmentation d'images médicales sont prometteurs, bien qu'ils demeurent imparfaits. Au cours des sections précédentes, nous avons constaté que les segmentations étaient souvent fidèles à la réalité anatomique humaine. Cependant, nous avons également observé des problèmes de sous-segmentation et de volume inférieur par rapport à la vérité terrain, ce qui a conduit à des résultats moins satisfaisants selon les métriques utilisées.

De plus, nous avons fréquemment identifié la présence de petits "îlots" ou de discontinuités dans les segmentations des organes. Ces îlots peuvent être causés par divers facteurs, tels que les variations de contraste. Pour résoudre ce problème, il est courant d'appliquer l'algorithme du "plus grand composant connexe" (largest connected component) aux résultats de la segmentation.

Nous émettons l'hypothèse que ces problèmes pourraient être partiellement résolus en utilisant une fonction de coût qui gère mieux le déséquilibre entre les classes, étant donné que la plupart des organes moins bien segmentés sont de petite taille. Il serait également bénéfique d'avoir une base de données d'entraînement plus vaste et diversifiée.

Il est important de noter que l'utilisation de l'algorithme du plus grand composant connexe pour traiter les îlots dans les segmentations peut poser un défi supplémentaire pour les organes symétriques. Étant donné que ces organes sont souvent présents par paires, l'application de l'algorithme pourrait potentiellement entraîner la suppression d'un des organes s'il est considéré comme un îlot distinct.

Cependant, il convient de souligner qu'il reste encore beaucoup de travail à accomplir pour comprendre et analyser les résultats obtenus.

Il serait aussi judicieux de considérer l'utilisation d'un modèle plus complexe que UNet. une approche intéressante à explorer consisterait à intégrer des transformers dans notre modèle, ce qui pourrait améliorer les segmentations.

Mais nous avons opté pour l'utilisation de l'architecture U-Net plutôt que U-NetR en raison de contraintes liées au temps de calcul et au coût computationnel. Notre objectif était d'établir une ligne de base solide en utilisant une architecture plus simple, ce qui nous a permis de nous concentrer davantage sur les pré-traitements et les améliorations spécifiques à notre tâche, avant de passer éventuellement à une architecture plus complexe.

Cette approche nous a permis de maximiser l'efficacité et l'efficacité de nos expérimentations. Une autre possibilité à envisager serait d'entraîner le modèle en utilisant conjointement l'ensemble des données, étant donné sa ca-

pacité à généraliser. De plus, il est regrettable de constater que la base de données d'entraînement utilisée n'était pas très large et pas équilibrée en termes de répartition entre les femmes et les hommes. Il serait important de remédier à ce déséquilibre afin d'obtenir des résultats plus généralisables et représentatifs de la population.

En conclusion, bien que nos résultats préliminaires soient encourageants, il reste encore beaucoup de travail à faire pour comprendre et analyser les résultats obtenus. L'utilisation d'un modèle plus avancé et la correction du déséquilibre de la base de données pourraient contribuer à améliorer les segmentations dans le cadre de notre projet de fin d'études sur la segmentation d'images médicales.

4.4 Conclusion

Dans ce chapitre nous avons vu l'environnement de l'expérimentation, les résultats de prédiction des images que nous avons obtenus ainsi que l'évaluation de la performance de prédiction.

Conclusion Générale

En conclusion, la segmentation joue un rôle crucial dans le domaine médical et la radiothérapie assistée par ordinateur en permettant une précision et une personnalisation essentielles pour des traitements plus efficaces et moins invasifs. Dans ce stage de fin d'études en segmentation d'images sur IRM et CT des six organes à risques, à savoir le cerveau, le tronc cérébral, le chiasma optique, les yeux, les nerfs optiques et les oreilles internes, nous avons réalisé un travail approfondi de prétraitement des données afin d'obtenir des cartes de labels uniformes et de normaliser et redimensionner les images de CT et d'IRM dans le même espace. Cela nous a permis d'appliquer les mêmes expériences sur les deux modalités.

Nous avons entraîné un modèle UNet sur les deux modalités séparément ainsi que conjointement en utilisant deux fonctions de coût différentes, dans le but de déterminer dans quelle mesure une Cross Entropy de Dice et d'entropie croisée pouvait aider à segmenter les parties difficiles à identifier et également de déterminer dans quelle mesure le modèle arrivait à généraliser s'il était entraîné sur deux modalités différentes conjointement. Nos résultats ont clairement démontré que l'utilisation de cette perte améliorait significativement les performances de segmentation. De plus, nous avons constaté que la modalité d'origine des segmentations, à savoir le CT, produisait de meilleurs résultats par rapport à l'IRM. Enfin, nos résultats ont démontré que le modèle était capable de généraliser de manière satisfaisante sur les deux modalités distinctes lorsqu'il était entraîné sur les deux simultanément.

Pour nos perspectives d'amélioration, nous envisageons l'adoption d'une

fonction de perte encore plus performante comme la TopKLoss afin de continuer à améliorer les résultats de segmentation. De plus, nous prévoyons de passer à une architecture plus complexe, notamment l'utilisation d'UNet avec des transformers, qui est connue pour sa capacité à capturer des informations contextuelles plus riches. Nous considérons également l'augmentation de la taille de la base de données, si possible, afin d'enrichir et de diversifier les exemples d'apprentissage.

En résumé, ce stage de fin d'études nous a permis de réaliser une segmentation précise des six organes à risques sur les images de CT et d'IRM. Les résultats obtenus ont souligné l'importance de la fonction de perte dans l'amélioration des performances de segmentation, ainsi que l'avantage initial du CT en termes de qualité de segmentation. Ces résultats encouragent de futures recherches visant à explorer des architectures plus avancées et à étendre la taille de la base de données.

Bibliographie

- [1] Computed tomography (ct or cat) scan of the brain. Jogn Hopkins Medicine <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/computed-tomography-ct-or-cat-scan-of-the-brain>.
- [2] How mri head scans work. How MRI Head Scans Work <https://www.vista-health.co.uk/blog/how-mri-head-scans-work>.
- [3] Md Motiur Rahman Sagar and Martin Dyrba. Learning shape features and abstractions in 3d convolutional neural networks for detecting alzheimer's disease. *arXiv preprint arXiv :2009.05023*, 2020.
- [4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net : Convolutional networks for biomedical image segmentation. 2015.
- [5] Ali Hatamizadeh, Yucheng Tang, Vishwesh Nath, Dong Yang, Andriy Myronenko, Bennett Landman, Holger R Roth, and Daguang Xu. Unetr : Transformers for 3d medical image segmentation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 574–584, 2022.
- [6] École centrale de nantes. Wikipédia https://fr.wikipedia.org/wiki/%C3%89cole_centrale_de_Nantes.
- [7] Ls2n - équipe sims (signal, image et son). Wikipédia <https://hal.archives-ouvertes.fr/LS2N-SIMS>.
- [8] Champakamala Sundar Rao¹ K. Karunakara¹. A comprehensive review on brain tumor segmentation and classification of mri images. 2021.

- [9] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net : learning dense volumetric segmentation from sparse annotation. In *MICCAI*, pages 424–432. Springer, 2016.
- [10] Matthieu Lê, Hervé Delingette, Jayashree Kalpathy-Cramer, Elizabeth R Gerstner, Tracy Batchelor, Jan Unkelbach, and Nicholas Ayache. Personalized radiotherapy planning based on a computational tumor growth model. *IEEE transactions on medical imaging*, 2016.
- [11] Chiasma. Wikipédia https://fr.wikipedia.org/wiki/Chiasma_optique.
- [12] Tronc cérébral. Wikipédia https://fr.wikipedia.org/wiki/Tronc_c%C3%A9r%C3%A9bral.
- [13] Nerf optique. Wikipédia https://fr.wikipedia.org/wiki/Nerf_oculomoteur, url = https://fr.wikipedia.org/wiki/Nerf_oculomoteur.
- [14] Oreille interne. Wikipédia https://fr.wikipedia.org/wiki/Oreille_interne.
- [15] Laura A Dawson and Michael B Sharpe. Image-guided radiotherapy : rationale, benefits, and limitations. *The lancet oncology*, 2006.
- [16] Ct scan vs. mri. Healthline <https://www.healthline.com/health/ct-scan-vs-mri>.
- [17] Jiang Hsieh. Computed tomography : principles, design, artifacts, and recent advances. 2003.
- [18] Magnetic resonance imaging. The free Dictionary <https://medical-dictionary.thefreedictionary.com/magnetic+resonance+imaging>.
- [19] What’s the difference between dicom and nifti? <https://encord.com/blog/whats-the-difference-between-dicom-and-nifti/>.
- [20] Michele Larobina and Loredana Murino. Medical image file formats. *Journal of digital imaging*, 2014.

- [21] Informatics in radiology : Dicom-rt and its utilization in radiation therapy. radiographics. *International journal of radiation biology*, 2009.
- [22] Maria YY Law and Brent Liu. Dicom-rt and its utilization in radiation therapy. *Radiographics*, 2009.
- [23] Wiki slicer documentation. Wiki <https://www.slicer.org/wiki/Documentation>.
- [24] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. 2016.
- [25] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. 2002.
- [26] Image segmentation - 3d slicer. 3D Slicer https://slicer.readthedocs.io/en/latest/user_guide/image_segmentation.html.
- [27] Mohammad Hesam Hesamian, Wenjing Jia, Xiangjian He, and Paul Kennedy. Deep learning techniques for medical image segmentation : Achievements and challenges. *Journal of Digital Imaging*, 2019.
- [28] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net : Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. IEEE, 2016.
- [29] Trevor Darrell, Evan Shelhamer, and Jonathan Long. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.
- [30] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words : Transformers for image recognition at scale. *International Conference on Learning Representations (ICLR)*, 2020.

- [31] Isabelle Bloch. Recalage et fusion d'images médicales.
- [32] Image registration. Wikipedia https://en.wikipedia.org/wiki/Image_registration.
- [33] Jun Ma, Jianan Chen, Matthew Ng, Rui Huang, Yu Li, Chen Li, Xiaoping Yang, and Anne L Martel. Loss odyssey in medical image segmentation. 2021.
- [34] Shruti Jadon. A survey of loss functions for semantic segmentation. In *2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB)*, pages 1–7. IEEE, 2020.
- [35] Abdel Aziz Taha and Allan Hanbury. Metrics for evaluating 3d medical image segmentation : analysis, selection, and tool. *BMC medical imaging*, 2015.
- [36] Nicholas Reinke, William Grathwohl, and Dan Hendrycks. Common limitations of image processing metrics : A picture story. ArXiv <https://arxiv.org/abs/2106.04460>, 2021.
- [37] Imfusion. ImFusion <https://www.imfusion.com/>.
- [38] About 3d slicer. 3D Slicer https://slicer.readthedocs.io/en/latest/user_guide/about.html#what-is-3d-slicer.
- [39] Project monai. MONAI, <https://docs.monai.io/en/stable/>.

Annexe

A.1 Imfusion

ImFusion Suite est un puissant outil logiciel polyvalent spécialement conçu pour la visualisation et le traitement de données médicales. Avec son accent sur le calcul haute performance à la fois sur le CPU et le GPU, ImFusion Suite propose une gamme complète d’algorithmes de pointe pour des tâches telles que la reconstruction d’images, l’enregistrement, la segmentation et l’apprentissage automatique.

Ce qui distingue ImFusion Suite, c’est son interface conviviale, qui facilite les interactions intuitives qui m’a permis d’analyser et de manipuler efficacement les ensembles de données d’images médicales complexes.

De plus, le système de plugins d’ImFusion Suite permet une intégration transparente de fonctionnalités supplémentaires et l’incorporation d’algorithmes personnalisés.

Ce logiciel a joué un rôle crucial dans ma recherche en fournissant des outils avancés et des fonctionnalités qui améliorent ma capacité à effectuer des tâches précises et efficaces de segmentation d’images médicales. [37]

A.2 3Dslicer

Il s’agit d’une application logicielle à code source ouvert spécialement conçue pour la visualisation et l’analyse de jeux de données d’images médicales. Il prend en charge tous les types de données couramment utilisés, tels que les images, les segmentations, les surfaces, les annotations, les transformations, etc., en 2D, 3D et 4D. L’analyse réalisée avec 3D Slicer comprend le recalage et diverses mesures quantitatives. En tant que plateforme de recherche, 3D Slicer permet le développement et l’évaluation rapide de nouvelles méthodes, ainsi que leur distribution aux utilisateurs cliniques. De plus, un environnement Python complet est également disponible avec ce logiciel. [38] [26]

A.3 LS2N Infra Cloud Computing for AI (LICCIA)

hébergée par OS-BIRD, c'est le serveur qui m'a été mis à disposition pour exécuter les algorithmes développés.

A.4 Pytorch

PyTorch est un framework de calcul scientifique utilisé dans de nombreux domaines, y compris la vision par ordinateur et l'analyse d'images médicales. Dans mon projet de fin d'études portant sur la segmentation d'images médicales, Nous avons choisi d'utiliser PyTorch en raison de sa flexibilité, de sa puissance de calcul et de sa riche collection de fonctionnalités. PyTorch permet de créer des modèles de segmentation sophistiqués en combinant des couches de neurones, des opérations mathématiques et des fonctions d'optimisation. Grâce à sa nature dynamique, PyTorch facilite le processus de développement, de débogage et de modification des modèles. De plus, sa prise en charge du calcul sur GPU accélère significativement les calculs, permettant ainsi de traiter des volumes importants de données médicales en un temps raisonnable. En résumé, PyTorch s'est révélé être un outil essentiel et puissant dans la réalisation de mon projet de fin d'études, me permettant d'explorer et d'expérimenter des techniques avancées de segmentation d'images médicales avec précision et efficacité. [?]

A.5 MONAI

MONAI (Medical Open Network for AI) est une bibliothèque de traitement d'images médicales développée spécifiquement pour la recherche et l'application de l'intelligence artificielle en imagerie médicale. Dans mon projet de fin d'études portant sur la segmentation d'images médicales, Nous avons choisi d'utiliser MONAI en raison de sa nature spécialisée et de sa richesse en

fonctionnalités dédiées à l'analyse et à la manipulation d'images médicales. MONAI offre une vaste collection d'outils et d'algorithmes avancés, tels que la préparation des données, la transformation des images, la construction de modèles d'apprentissage automatique et l'évaluation des performances. Grâce à son intégration harmonieuse avec PyTorch, MONAI permet une mise en œuvre efficace et flexible des pipelines de segmentation d'images médicales, en fournissant des modules prêts à l'emploi pour le prétraitement des données, l'augmentation des données, la définition des architectures de réseau et l'évaluation des résultats. En utilisant MONAI, Nous avons pu bénéficier d'un environnement de développement complet et optimisé pour la réalisation de mon projet, me permettant d'explorer et de mettre en œuvre des approches de pointe en segmentation d'images médicales avec rigueur et précision. [39]