



**HAL**  
open science

## Segmentation automatisée de structures cérébrales profondes à partir d'IRMs à Inversion-Récupération

Aigerim Dautkulova, Omar Ait Aider, Céline Teulière, Jerome Coste, Rémi Chaix, Omar Ouachikh, Bruno Pereira, Jean-Jacques Lemaire

► **To cite this version:**

Aigerim Dautkulova, Omar Ait Aider, Céline Teulière, Jerome Coste, Rémi Chaix, et al.. Segmentation automatisée de structures cérébrales profondes à partir d'IRMs à Inversion-Récupération. RFIAP 2024 - Reconnaissance des Formes, Image, Apprentissage et Perception, Jul 2024, Lille, France. hal-04616150

**HAL Id: hal-04616150**

**<https://hal.science/hal-04616150>**

Submitted on 18 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Segmentation automatisée de structures cérébrales profondes à partir d'IRM à Inversion-Récupération

A. Dautkulova<sup>1</sup>  
C. Teulière<sup>1</sup>  
R. Chaix<sup>1,2</sup>  
B. Pereira<sup>3</sup>

O. Ait Aider<sup>1</sup>  
J. Coste<sup>1,2</sup>  
O. Ouachik<sup>1</sup>  
J.-J. Lemaire<sup>1,2</sup>

<sup>1</sup> Université Clermont Auvergne, Clermont Auvergne INP, CNRS, Institut Pascal, F-63000 Clermont-Ferrand, France

<sup>2</sup> Université Clermont Auvergne, CNRS, CHU Clermont-Ferrand, Clermont Auvergne INP, Institut Pascal, F-63000 Clermont-Ferrand, France

<sup>3</sup> Direction de la Recherche et de l'Innovation, CHU Clermont-Ferrand, F-63000 Clermont-Ferrand, France

aigerim.dautkulova@uca.fr

## Résumé

La segmentation automatique des structures cérébrales est un sujet majeur en recherche médicale. Les petites structures du cerveau profond ont reçu peu d'attention, notamment en raison de l'absence de contourage manuel par des experts. Dans cette étude, nous avons évalué la segmentation automatique par un réseau nnU-Net entraîné grâce à un nouveau jeu de données cliniques contenant des images IRM WAIR (White Matter Attenuated Inversion-Recovery) représentant cinq structures segmentées manuellement chez 53 patients atteints d'une maladie de Parkinson sévère. Des images T1 et DTI ont également été utilisées. Nous avons évalué l'impact de la réorientation des vecteurs de diffusion DTI par rapport à la ligne ACPC. Les sous-ensembles d'entraînement et de test comprenaient respectivement 38 et 15 patients. Comme critères d'évaluation, nous avons utilisé le coefficient de similarité de Dice (DSC), la distance de Hausdorff à 95% (95HD) ainsi que la similarité volumétrique (VS). Des modèles à effets aléatoires ont été utilisés afin de comparer statistiquement les performances. Les résultats montrent notamment que WAIR est nettement plus performant que T1 pour DSC ( $0,739 \pm 0,073$ ), 95HD ( $1,739 \pm 0,398$ ) et VS ( $0,892 \pm 0,044$ ). Les valeurs DSC pour la segmentation automatisée de MB, RN, SN, STN et MT-fa ont plus ou moins diminué selon la complexité de la segmentation manuelle. Par ailleurs, la réorientation des vecteurs DTI a amélioré la segmentation automatisée.

## Mots Clef

Segmentation automatisée, nnU-Net, images multimodales, WAIR, T1, DTI, ACPC.

## Abstract

Methods for automatic segmentation of brain structures are a major subject of medical research. The small structures of the deep brain have received little attention, partly due to the lack of manual delineation by medical experts. In this study, we evaluated the automatic segmentation of a novel clinical dataset containing White Matter Attenuated Inversion-Recovery (WAIR) MRI images and five manually segmented structures in 53 patients with severe Parkinson's disease. T1 and DTI images were also used. We also evaluated the reorientation of the DTI diffusion vectors relative to the ACPC line. A state-of-the-art nnU-Net method was trained and tested on subsets of 38 and 15 image sets respectively. We used Dice Similarity Coefficient (DSC), 95% Hausdorff Distance (95HD), and Volumetric Similarity (VS) as metrics to evaluate the network's effectiveness in reproducing manual contours. Random effects models statistically compared values across structures, accounting for variability between and within participants. The results show that WAIR significantly outperforms T1 for DSC ( $0.739 \pm 0.073$ ), 95HD ( $1.739 \pm 0.398$ ), and VS ( $0.892 \pm 0.044$ ). DSC values for automated segmentation of MB, RN, SN, STN, and MT-fa decreased in that order, consistent with the increasing complexity observed in manual segmentation. According to the training results, reorientation of DTI vectors improved automated segmentation.

## Keywords

Automated segmentation, nnU-Net, multi-modal images, WAIR, T1, DTI, ACPC.

# 1 Introduction

Le cerveau profond (DB) est une région hautement complexe [6, 20] et très importante pour les maladies neuropsychiatriques telles que Parkinson, Alzheimer, ou les désordres compulsifs (voir par exemple [15]). Dans les cas de maladie de Parkinson sévère, des traitements médicamenteux combinés avec une stimulation du cerveau (DBS) permettent d’améliorer la qualité de vie [11]. Cependant, l’efficacité de cette bithérapie repose notamment sur la bonne localisation des contacts dans la région supéro-latérale du noyau subthalamique (STN) [16], en repérant la cible STN dans le sous-thalamus.

La cible est généralement représentée sur des coupes d’imagerie par résonance magnétique (IRM), sur lesquelles sont superposés des repères stéréotaxiques, par exemple le point final de la trajectoire et/ou le contour du STN, supervisés par le neurochirurgien. Les atlas stéréotaxiques, tels que le célèbre atlas de Schaltenbrand et Bailey [29], peuvent être mis en correspondance en utilisant les points de référence des commissures ventriculaires (commissures blanches antérieures et postérieures, AC et PC). La plupart des cliniciens interprètent les contrastes IRM en fonction du type de séquence afin d’optimiser le ciblage. Dans de précédents travaux, nous avons développé une séquence d’inversion-récupération dédiée, nommée WAIR (White Matter Attenuated Inversion-Recovery) pour faciliter l’identification des structures du DB directement sur les coupes coronales [10, 34]. En bref, (i) plus la cellularité est élevée (par exemple, la substantia nigra, notamment la pars compacta), plus le signal est fort, (ii) plus la densité axonale est élevée (par exemple, le fascicule mammillo-thalamique), plus le signal est faible, et (iii) le noyau rouge, une structure de matière grise, a un signal faible parce qu’il contient un contingent important d’axones. Cette séquence est utilisée en routine clinique depuis 2006 avec un appareil de 1.5 tesla, et le STN et ses structures voisines sont délimités manuellement, ce qui permet une planification 3D détaillée pour l’implantation des électrodes. Cependant, cette cartographie individualisée nécessite 1h 30 min de temps de travail pour les deux hémisphères en raison de l’architecture 3D complexe et de l’interprétation difficile du contraste de l’IRM au niveau individuel. L’objectif du présent travail est donc d’automatiser la segmentation des structures sous-thalamiques pour simplifier la personnalisation de la cartographie, avec un temps de calcul raisonnable de quelques dizaines de minutes. Pour cela, nous avons choisi une approche d’apprentissage profond (DL) appliquée à notre base de données clinique de planification d’électrodes.

L’un des modèles de segmentation d’images basés sur le DL les plus utilisés est U-Net [28], qui repose sur un réseau neuronal convolutif en forme de U. Plusieurs versions modifiées de cette architecture ont été proposées, telles que 3D U-Net [4], Modified U-Net [30] et Dense-UNet [2]. [17] considère qu’un U-Net classique bien configuré est difficile à surpasser et a introduit nnU-Net, un réseau

auto-configurant utilisant un U-Net 3D standard modifié. Le modèle nnU-Net configure automatiquement toutes les étapes nécessaires du pipeline de segmentation automatisé, y compris le prétraitement, l’architecture du réseau, les hyperparamètres et les étapes de post-traitement. Il a donné des résultats de pointe dans diverses tâches, notamment la segmentation de l’hippocampe dans le cadre du Medical Segmentation Decathlon Challenge [31]. Pour notre travail, nous avons choisi la configuration 3D de nnU-Net avec un U-Net 3D standard comme architecture de base (*backbone*).

La segmentation précise des structures du DB représente un défi, notamment en raison de la petite taille de la plupart des structures, qui sont peu connues et souvent coalescentes, fusionnant les corps cellulaires et les fascicules d’axones, ce qui entraîne des contrastes IRM importants. Par conséquent, les jeux de données de référence, basés sur des atlas ou contourés manuellement sur des images IRM, sont très rares et de taille limitée. La plupart des méthodes de segmentation par DL ont utilisé des jeux de données d’images T1 [8, 18, 26] (Tableau 1). Dans notre travail, en plus du WAIR, nous avons également pris en compte deux autres séquences IRM acquises lors de la chirurgie, à savoir 3D T1 MPRAGE, et l’image du tenseur de diffusion (DTI) rarement utilisée [25]) pour la planification complète des électrodes en 3D et le contrôle postopératoire.

Nous avons évalué la segmentation DL de cinq structures du DB à l’aide de nnU-Net sur un nouvel ensemble de données cliniques avec une séquence d’inversion-récupération et sa fusion avec les imageries T1 et DTI. Nous avons évalué l’efficacité du nnU-Net 3D auto-configuré entraîné ( $n = 38$  patients) sur un ensemble de test ( $n = 15$  patients) pour reproduire le contour de cinq structures sous-thalamiques dans les hémisphères droit et gauche. Nous avons étudié l’efficacité des séquences WAIR et T1, de différentes formulations dérivées du DTI et de leurs combinaisons. Nous avons également évalué l’impact de la réorientation des vecteurs de diffusion principaux par rapport au système de coordonnées du patient afin de réduire la variabilité inter-individuelle due aux légères différences de positionnement de la tête dans le cadre stéréotaxique de l’appareil d’IRM. En résumé, nos contributions sont les suivantes :

- Première utilisation de l’IRM WAIR pour la segmentation d’image basée sur l’apprentissage profond,
- Nouveau jeu de données cliniques des structures du DB segmentées manuellement par un clinicien expert (GT),
- Évaluation de l’amélioration des performances grâce à la multimodalité (WAIR + T1, WAIR + DTI),
- Impact de la réorientation du vecteur DTI par rapport à la ligne ACPC de chaque patient.

TABLE 1 – Méthodes de l'état de l'art pour la segmentation des structures cérébrales profondes (n.d. non disponible)

Auteurs (année) et méthode	Référence de l'ensemble de données	Séquence IRM, champ magnétique (nombre de cas)	Méthode d'annotation	Accès gratuit	Structures
[8] LiviaNET	IBSR <sup>a</sup>	T1, n.d. (18)	Manuel	Oui	Thalamus, Noyau Caudé, Putamen, Pallidum
	ABIDE <sup>b</sup>	T1, n.d. (1112)	Basé sur l'Atlas	No	
[18] 2.5D CNN	MICCAI 2012 <sup>c</sup>	T1, n.d. (35)	Manuel	Oui	Thalamus, Noyau Caudé, Putamen, Pallidum, Amygdale, Hippocampe, Accumbens
	IBSR	T1, n.d. (18)	Manuel	Oui	
[26] SubForkNet	NAMIC <sup>d</sup>	T1, n.d. (18)	Semi-automatique	Non	Thalamus, Noyau Caudé, Putamen, Pallidum, Amygdale, Hippocampe, Accumbens
	MICCAI 2012	T1, n.d. (35)	Manuel	Oui	
[22] Modified U-Net	Interne	T1 MPRAGE, 3T (105), 7T (40)	Manuel (côté gauche)	Non	Thalamus, 11 noyaux thalamiques, voie mammillothalamique
[1] nnU-Net	10 publiques et 4 internes	T1, mixte, 1.5T, 3T (1028)	Basé sur l'Atlas	Non	Caudé, Globus Pallidus (Externus, Internus), Noyaux Habénuclaires, Capsule interne, Accumbens, Putamen, Noyau rouge, Substantia nigra (pars compacta, pars reticulata), Noyau subthalamique, Thalamus, Noyau latéral postérieur ventral du Thalamus, Ventricule latéral, Noyau ventro-intermédiaire du Thalamus
[12] 3D U-Net	FSM <sup>e</sup> ADNI <sup>f</sup>	T1, 3T (29) T1, n.d. (10)	Manuel	Sur demande	Hypothalamus, corps mamillaires, noyaux septaux, Fornix, Accumbens, cerveau antérieur basal
[25] 2D U-Net	Interne	T1, 3T, DWI (121)	Manuel	Non	Ventricule, Noyau caudé, Putamen, Thalamus, Globus pallidus, Hippocampe, Amygdale et Accumbens

<sup>a</sup> <https://www.nitrc.org/projects/ibsr> (consulté le 05-February-2024)

<sup>b</sup> [https://econ\\_1000.projects.nitrc.org/indi/abide/](https://econ_1000.projects.nitrc.org/indi/abide/) (consulté le 05-February-2024)

<sup>c</sup> [http://www.neuromorphometrics.com/2012\\_MICCAI\\_Challenge\\_Data.html](http://www.neuromorphometrics.com/2012_MICCAI_Challenge_Data.html) (consulté le 05-February-2024)

<sup>d</sup> [https://www.na-mic.org/w/index.php?title=Main\\_Page&oldid=98830](https://www.na-mic.org/w/index.php?title=Main_Page&oldid=98830) (consulté le 27-June-2023)

<sup>e</sup> <https://surfer.nmr.mgh.harvard.edu/fswiki/FsData> (consulté le 27-June-2023)

<sup>f</sup> <https://adni.loni.usc.edu> (consulté le 05-February-2024)

## 2 Matériels et méthodes

Nous présentons ici le matériel clinique utilisé dans les travaux (Section 2.1), le prétraitement de l'espace géométrique (Section 2.2), ainsi que la représentation et la transformation du DTI (Section 2.2). La section 2.2 décrit la réorientation des vecteurs DTI. Dans la section 2.3, nous détaillons l'entraînement et la mise en œuvre de la segmentation automatisée. Dans la section 2.4, nous décrivons les critères d'évaluation. Enfin, dans la section 2.5, nous fournissons des détails sur l'analyse statistique.

### 2.1 Matériel clinique

Le jeu de données est constitué d'images de 53 patients atteints de la maladie de Parkinson sévère ( $61 \pm 6$  ans, 26 femmes), implantés bilatéralement dans le STN, dans un seul établissement (Centre expert pour la maladie de Parkinson), entre novembre 2011 et novembre 2022. Les autorisations éthiques, institutionnelles et individuelles ont été obtenues (programme Ptolémée<sup>®</sup> : IRB 5921, 30 mars 2018; CNIL réf. MR 004, M200702 10 novembre 2020). Sur les 110 patients opérés sur la période et ayant accepté de participer, 42 ont été exclus car ils présentaient d'autres troubles (dont 32 tremblements essentiels), 12 car les électrodes étaient implantées dans d'autres cibles, et 3 car les jeux de données DTI étaient corrompus. Chaque ensemble de données individuel se composait (i) des ensembles de

données WAIR, T1 et DTI, co-registrés lors de l'acquisition (cadre stéréotaxique fixé à la tête et verrouillé dans la bobine de tête) (Fig. 1), et (ii) des objets construits à partir des contours des structures, exportés en format de stéréolithographie (stl) (Iplan 3.0<sup>®</sup>, BrainLab, Allemagne). Pour une analyse plus approfondie, les objets stl ont été convertis au format de fichier NIfTI (Neuroimaging Informatics Technology Initiative) (logiciel 3D Slicer<sup>1</sup>) [24]). Les trois séquences ont été acquises sur un appareil d'IRM de 1,5 tesla (Aera et Sonata; Siemens, Allemagne). Toutes les données ont été rendues anonymes et les images T1 ont été défigurées à l'aide du logiciel pydeface [13].

La séquence WAIR a été développée pour augmenter les contrastes tissulaires spontanés inter et intra structures du DB [21, 33, 34]. Les paramètres de la séquence WAIR étaient les suivants : 2D coronal, temps de répétition 4500 ms, temps d'écho 13 ms, temps d'inversion 160 ms, angle de bascule 170°, taille de pixel  $0,527 \times 0,527 \text{ mm}^2$ , épaisseur de coupe 2 mm, matrice  $512 \times 512 \times 30$ . L'orientation coronale a été utilisée pour faciliter l'identification des structures sous-thalamiques dans le plan perpendiculaire à l'axe rostro-caudal du système nerveux central. La délimitation manuelle de toutes les structures sur les images WAIR a été réalisée par un expert clinique (JJL) à l'aide du même logiciel chirurgical (Iplan 3.0<sup>®</sup> BrainLab, Alle-

1. <https://www.slicer.org/>

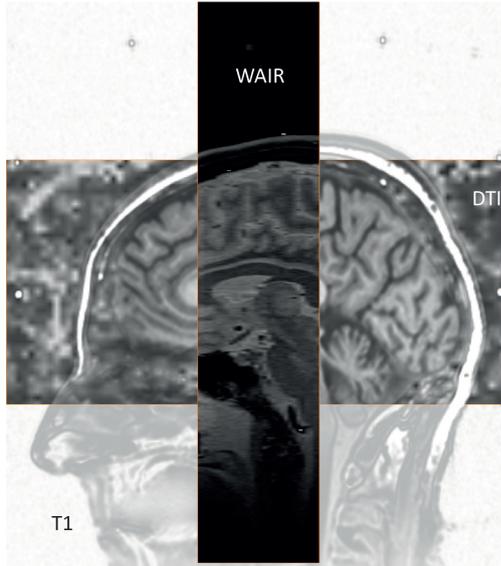


FIGURE 1 – Superposition d’images WAIR, T1 et DTI recalées.

magne) (Fig. 2). Nous avons sélectionné cinq structures caractéristiques dans chaque hémisphère : la substantia nigra (SN), le noyau subthalamique (STN), le noyau rouge (RN), le corps mammillaire (MB) et le faisceau mammillothalamique (MT-fa) (Fig. 2). Les paramètres T1 et DTI étaient les suivants : T1, 3D MPRAGE, sagittal, temps de répétition 1660 ms, temps d’écho 3,35 ms, angle de bascule 15°, taille de pixel 0,625 × 0,625 mm<sup>2</sup>, épaisseur de la tranche 2 mm, matrice 480 × 512 × 80, DTI, épi axial, 20 directions,  $b_0 = 750$  s/mm<sup>2</sup>, temps de répétition 5300 ms, temps d’écho 82 ms, taille du pixel 3 × 3 mm<sup>2</sup>, épaisseur 3 mm, matrice 80 × 80 × 40. Nous avons calculé les volumes de ces structures en mm<sup>3</sup> à partir des volumes NIFTI.

## 2.2 Prétraitement

**Espace géométrique de référence.** Les ensembles d’images T1 et DTI ont été recalés sur le WAIR, sur lequel les structures ont été contourées, et enregistrés à l’aide de 3D Slicer (T1, enregistrement général (outil BRAINS), affine, 12 degrés de liberté; DTI, module de rééchantillonnage du volume DTI). La taille de l’image résultante était de 512 × 512 × 30 avec une taille de voxel de 0,53 × 0,53 × 2,0 mm<sup>3</sup>.

**Représentation et transformation DTI.** Le voxel DTI contient un tenseur de diffusion de 3 × 3 qui peut être décomposé en trois valeurs propres ( $\lambda_1, \lambda_2, \lambda_3$ ) et trois vecteurs propres ( $v_1, v_2, v_3$ ), qui représentent respectivement la quantité et l’orientation de la diffusion des protons. Le vecteur propre principal  $v_1$  indique la direction de la diffusion maximale avec la valeur propre la plus élevée  $\lambda_1$  [36]. Dans nos expériences, nous avons considéré quatre dérivés du DTI :

- **Carte d’anisotropie fractionnelle (FA).** FA est une valeur scalaire indépendante de l’orientation

qui décrit le degré d’anisotropie. Elle est comprise entre 0 (isotrope) et 1 (anisotrope).

- **Carte vectorielle de la direction de diffusion principale (pdv).** Il s’agit d’une image 3D avec trois canaux représentant les trois composantes du vecteur propre principal  $v_1$  à chaque voxel.
- **Color-coded fractional anisotropy map (ccFA).** La carte d’anisotropie fractionnelle codée en couleur est générée en multipliant le vecteur propre principal du tenseur de diffusion par la carte d’anisotropie fractionnelle. Nous suivons le codage couleur général [23], où le ccFA indique à la fois l’intensité directionnelle de la diffusion et sa direction (rouge signifie gauche-droite, bleu signifie supérieur-inférieur et vert signifie antérieur-postérieur). La formule couramment utilisée est la suivante :

$$R = |v_{1x}| \cdot FA, G = |v_{1y}| \cdot FA, B = |v_{1z}| \cdot FA \quad (1)$$

où  $v_{1x}, v_{1y}, v_{1z}$  sont les composantes x, y et z du vecteur  $v_1$ .

- **Carte de direction de diffusion principale pondérée (pdvFA)** Pour cette étude, nous avons créé une carte pdvFA qui contient à la fois le degré d’anisotropie et l’information sur l’orientation du vecteur. Contrairement à la ccFA, qui utilise les valeurs absolues du vecteur, la pdvFA préserve le signe de  $v_1$  (c’est-à-dire toutes les informations directionnelles) en multipliant directement le vecteur  $v_1$  par FA, ce qui permet une réorientation cohérente des données DTI :

$$\begin{aligned} \text{pdvFA}_x &= v_{1x} \cdot FA, \text{pdvFA}_y = v_{1y} \cdot FA, \\ \text{pdvFA}_z &= v_{1z} \cdot FA \end{aligned} \quad (2)$$

FA, ccFA et pdv sont les dérivés couramment utilisés du DTI, où pdv est généralement visualisé comme une image RGB, mais nous avons utilisé la représentation vectorielle originale de ce dérivé. L’estimation du tenseur de diffusion et le rééchantillonnage du DTI par rapport à la référence WAIR ont été effectués avec 3D Slicer. La commande TV-tool de la boîte à outils DTI-TK<sup>2</sup> [35] a été utilisée pour calculer les images de diffusion FA, ccFA et pdv dérivées des données DTI.

### Réorientation des vecteurs DTI selon la ligne ACPC.

Nous avons systématiquement réorienté les gradients de direction bruts de la machine le long de la ligne ACPC afin d’optimiser la cohérence des données de diffusion entre les patients, même si les acquisitions ont été réalisées dans des conditions très similaires (cadre stéréotaxique aligné manuellement sur les plans orbito-méatal et de la ligne médiane). L’orientation de la diffusion a donc été semi-standardisée par rapport au plan horizontal du DB. Les points AC, PC et la ligne médiane ont été positionnés manuellement et la matrice de rotation a été calculée à partir

2. <http://dti-tk.sourceforge.net/>

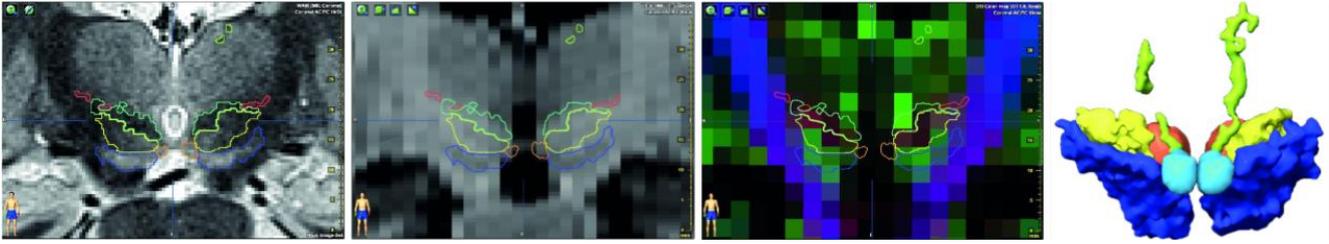


FIGURE 2 – Contours des cinq structures sous-thalamiques sélectionnées pour ce travail, superposés à la carte WAIR (première image à gauche), T1 (deuxième image), DTI codée en couleur (troisième image) et vue antérieure en 3D des cinq structures (quatrième image) (pas de filtrage d’image; Iplan 3.0<sup>®</sup>, BrainLab, Allemagne), le long de la même coupe coronale du STN : noyau subthalamique, jaune; substantia nigra, bleu; champs de Forel, vert; noyau de l’ansa lenticularis, orange; zona incerta, rouge; faisceau mammillo-thalamique, vert pâle.

des coordonnées. Les dérivés de la DTI ont été calculés avec et sans l’application de la transformation. Dans les sections suivantes, nous désignons les images avec des vecteurs de direction transformés et non transformés comme réorientés (<sup>R</sup>) et non réorientés (<sup>NR</sup>).

### 2.3 Entraînement et mise en œuvre

Nous avons utilisé la configuration 3D de nnU-Net (Fig. 3). Nous avons modifié le nombre d’époques de 1000 à 300 afin de réduire le temps d’entraînement. Nous avons divisé l’ensemble de données ( $n = 53$ ) en sous-ensembles d’entraînement ( $n = 38$ , 15 femmes, âge moyen 61) et de test (ou d’inférence) ( $n = 15$ , 11 femmes, âge moyen 61). Nous avons veillé à une répartition équitable des dates d’acquisition des données, c’est pourquoi le rapport de masculinité entre les groupes était différent, avec davantage de femmes dans le groupe de test (73% contre 39% dans l’ensemble d’entraînement). Les étapes de prétraitement par défaut ont été appliquées aux images WAIR, T1, ccFA et FA. Nous avons ignoré l’étape de normalisation de l’intensité pour les images vectorielles (pdv et pdvFA) car elles sont déjà normalisées (vecteurs unitaires). Pour le processus d’inférence, nous avons utilisé le meilleur point de contrôle du modèle DL (nnU-Net entraîné), au lieu du dernier point de contrôle obtenu. La taille de lot de 2 a été sélectionnée par nnU-Net pour toutes les expériences. nnU-Net a été implémenté en Python 3.8.5 en utilisant PyTorch 1.6.0. Nos calculs ont été effectués sur un superordinateur à l’aide d’un GPU doté de 16 Go de mémoire.

### 2.4 Métriques d’évaluation

Pour évaluer l’efficacité de la segmentation automatisée des cinq structures cérébrales profondes, nous avons utilisé trois mesures de similarité :

- **Coefficient de Dice (DSC)** [7]. Il s’agit de la mesure la plus largement utilisée pour valider les segmentations de volumes médicaux en vue d’une comparaison directe entre les segmentations automatiques et les segmentations manuelles. Le DSC mesure le chevauchement spatial entre deux segmentations. Il est compris entre 0 et 1, 1 indiquant

un chevauchement parfait.

- **95% distance de Hausdorff (95HD)**. Il s’agit d’une mesure de la dissimilarité entre deux ensembles de points [32]. La 95HD est similaire à la HD maximale. Cependant, elle est basée sur le 95e centile des distances au lieu du maximum. Une valeur de 95HD plus petite indique un meilleur alignement et un meilleur accord entre les segmentations. Elle est mesurée en millimètres.
- **Similitude volumétrique (VS)**. Elle calcule la similarité entre deux segments sur la base de leurs volumes [32]. La VS va de 0 à 1, où 1 indique la plus grande similarité volumétrique entre la segmentation automatisée et les labels GT.
- Nous considérons le DSC comme la principale mesure d’évaluation et 95HD et VS comme des mesures supplémentaires.

### 2.5 Analyse statistique

La performance de nnU-Net 3D auto-configuré a été analysée sur le sous-ensemble de test ( $n = 15$  patients), en fonction des valeurs des indices de similarité (DSC, 95HD et VS) des paires de contours (manuels et automatiques) des cinq structures (SN, STN, RN, MB et MT-fa). L’analyse a été effectuée sur la base d’ensembles de données d’images (WAIR, T1, WAIR + T1, WAIR + ccFA<sup>NR, R</sup>, WAIR + pdv<sup>NR, R</sup>, WAIR + pdvFA<sup>NR, R</sup>, WAIR + FA, WAIR + T1 + FA; WAIR comme référence, car il s’agit de la modalité utilisée pour la vérité terrain), par structure (SN, STN, RN, MB et MT-fa; STN comme référence, car il s’agit de l’objectif clinique du contourage manuel). Nous avons également évalué spécifiquement l’impact de la réorientation des gradients de diffusion sur les ensembles de données WAIR + ccFA<sup>NR, R</sup>, WAIR + pdv<sup>NR, R</sup>, WAIR + pdvFA<sup>NR, R</sup>. Les données sont présentées sous forme de moyenne et d’écart-type. L’hypothèse de normalité a été vérifiée à l’aide du test de Shapiro-Wilk. Des modèles à effets aléatoires (c’est-à-dire des modèles mixtes linéaires) ont été exécutés pour comparer le DSC, le 95HD et le VS entre les structures en tenant compte de la variabilité inter et intra participants (le sujet étant considéré comme un effet aléa-

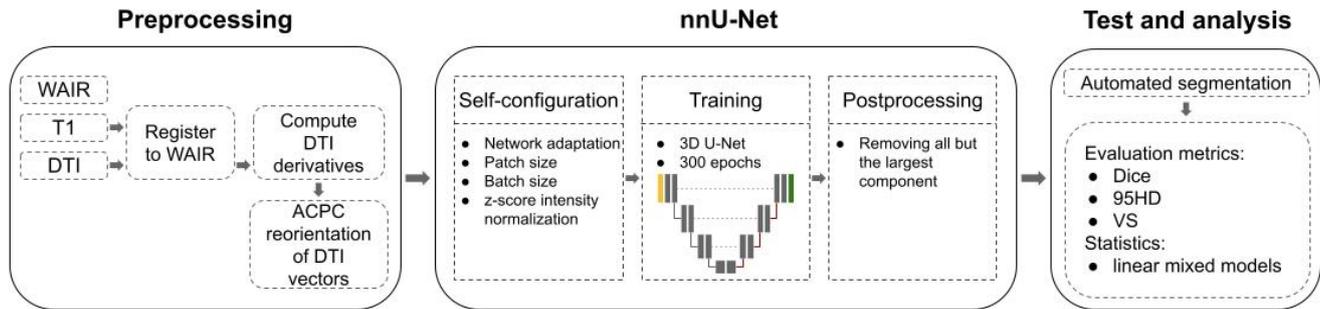


FIGURE 3 – Workflow : préparation des données ; entraînement du réseau de segmentation 3D nnU-Net ; inférence et évaluation.

toire). Les effets fixes suivants ont été analysés : ensembles de données d’images, structures, réorientation du gradient de diffusion ( $R^{NR}$ ) et leurs interactions. La normalité des résidus a été étudiée comme indiqué ci-dessus et une transformation logarithmique a été effectuée pour atteindre la normalité le cas échéant. Toutes les analyses ont été effectuées avec un ajustement sur les volumes. Les analyses statistiques ont été réalisées avec le logiciel statistique Stata (version 15, StataCorp, College Station, USA). Les tests étaient bilatéraux, avec une erreur de type 1 fixée à 5%. Une erreur de type 1 de Sidak a été appliquée pour prendre en compte les comparaisons multiples deux à deux. Les résultats ont été exprimés à l’aide des tailles d’effet de Hedge et des IC à 95% (disponibles sur demande).

### 3 Résultats

Le temps de calcul moyen de chaque expérience de test était d’environ 5 minutes. Le temps de calcul moyen de chaque optimisation d’entraînement était d’environ 30 heures pour WAIR et T1, d’environ 35 heures pour WAIR + T1 et WAIR + FA, et d’environ 60 heures pour les autres expériences.

**Comparaison globale des modèles DL.** Le Tableau 2 présente les résultats moyens des DSC, 95HD et VS pour les modèles DL formés. Le Tableau 3 présente les résultats de chaque structure pour plusieurs expériences. Nous avons constaté que les résultats obtenus avec WAIR étaient significativement supérieurs à ceux sur T1 en termes de valeurs DSC, 95HD et VS avec des valeurs moyennes de  $0,739 \pm 0,073$ ,  $1,739 \pm 0,398$  et  $0,892 \pm 0,044$  respectivement ( $p < 0,001$  pour DSC et 95HD et  $p = 0,025$  pour VS) (Fig. 4). Bien que la combinaison WAIR + T1 ait montré des valeurs plus élevées (avec DSC =  $0,747 \pm 0,073$ , 95HD =  $1,712 \pm 0,517$ , et VS =  $0,901 \pm 0,039$ ), cette différence n’était pas significative ( $p > 0,05$ ) par rapport à WAIR seul. Il n’y avait pas de différence significative entre WAIR et les combinaisons de WAIR avec DTI ( $p > 0,05$ ).

*Vecteurs DTI réorientés et non réorientés.* Dans le cas de WAIR + pdvFA, la valeur DSC s’est améliorée, passant de  $0,723 \pm 0,075$  à  $0,737 \pm 0,073$  (les valeurs 95HD et VS se sont également améliorées), mais cette différence n’est

pas statistiquement significative ( $p > 0,05$ ). Sur la base des résultats de l’entraînement, la valeur 95HD s’est statistiquement améliorée, passant de  $2,142 \pm 0,489$  à  $1,844 \pm 0,280$  lors de l’utilisation de vecteurs réorientés pour l’expérience WAIR + ccFA ( $p = 0,006$ ).

**Comparaison des modèles DL par structures.** WAIR est statistiquement plus performant que T1 sur la base du DSC pour les cinq structures ( $p < 0,05$ ), sur la base de 95HD pour STN ( $p = 0,004$ ), MT-fa ( $p < 0,001$ ), SN ( $p < 0,001$ ), et RN ( $p < 0,001$ ), et sur la base de VS pour STN ( $p = 0,009$ ). Pour la MB, sur la base des valeurs DSC, WAIR a nettement surpassé WAIR + pdv<sup>NR</sup> ( $p = 0,024$ ) et WAIR + pdv<sup>R</sup> ( $p = 0,028$ ). La combinaison WAIR + T1 a été significativement plus performante que l’expérience WAIR pour le MB sur la base de la valeur 95HD ( $p = 0,046$ ).

### 4 Discussion

Dans ce travail, nous avons utilisé un ensemble de données cliniques de référence comprenant les contours manuels du SN, RN, STN, MB et MT-fa, réalisés sur des images IRM d’une séquence WAIR dédiée à l’analyse de l’architecture du DB. Le contourage automatique basé sur WAIR a donné un bon résultat, avec une valeur DSC moyenne de  $0,739 \pm 0,073$ . Les valeurs DSC du MB, du RN, du SN, du STN et du MT-fa ont diminué dans cet ordre, reflétant la difficulté croissante rencontrée par un expert médical pour identifier ces structures, au moins pour les noyaux. Le sexe des patients ne semble pas affecter les performances du réseau DL.

La valeur DSC moyenne de 0,74 suggère que le nnU-Net 3D auto-configuré peut être appliqué cliniquement, et ce avec un temps de calcul raisonnable. Des valeurs moyennes de DSC supérieures à 0,85 ont été rapportées pour de grandes structures bien connues, facilement identifiables en T1, telles que le caudé, le putamen et le thalamus, [8, 18, 26]. Des résultats similaires ont également été obtenus en utilisant les cartes T1 et scalaires de la DTI [25]. Des structures moins connues, telles que le noyau accumbens ([18, 26], les noyaux thalamiques [22], les structures limbiques [1, 12] et diverses structures, notamment la capsule interne, l’habénula et les ventricules [1], ont

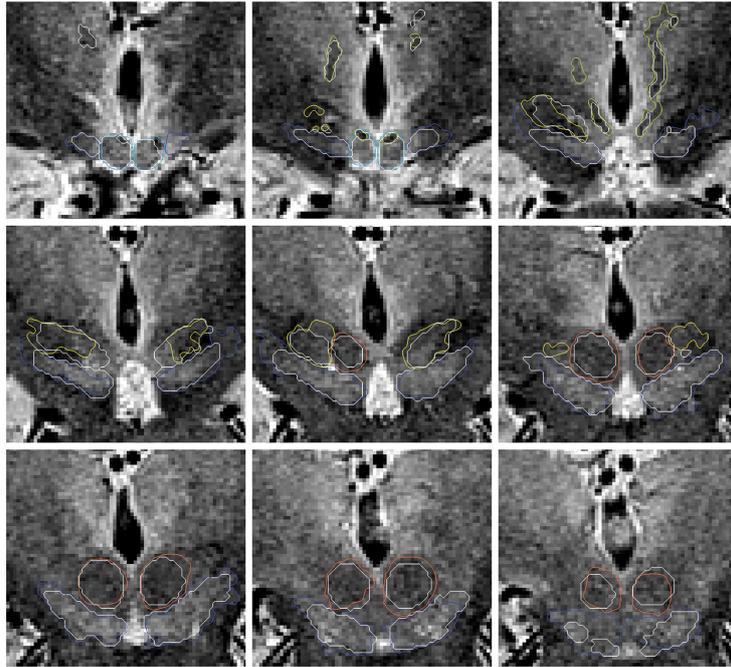


FIGURE 4 – Exemple d'un patient avec des contours manuels (ligne colorée) et automatiques (ligne blanche) basés sur WAIR : noyau sous-thalamique, jaune ; substantia nigra, bleu foncé ; fascicule mammillo-thalamique, vert pâle ; noyau rouge, rouge ; corps mammillaire, bleu pâle (même code de couleur que celui utilisé avec le logiciel chirurgical Iplan).

été analysées et ont montré une hétérogénéité des valeurs DSC moyennes. La diversité des protocoles expérimentaux pourrait expliquer le large éventail des valeurs DSC, comme l'utilisation de séries de cas mixtes (sujets sains, maladie dégénérative, autisme) [1, 8, 12, 18, 22, 26], les différentes segmentations de vérité terrain, manuelle avec atlas (aide visuelle ou recalage) [1, 8], semi-automatique avec tissu de classe et "gold standard" [26], manuelle sans atlas [3]. Malgré cela, le potentiel du DL dans la segmentation des structures cérébrales profondes à partir d'images IRM est confirmé.

Le contourage automatique en fonction de T1 uniquement a donné de moins bons résultats (valeur DSC moyenne,  $0,670 \pm 0,106$ ) que WAIR, en cohérence avec les détails de contraste plus faibles observés dans le sous-thalamus sur T1 (Fig. 2). D'autres équipes ont proposé des séquences IRM non T1 pour faciliter l'identification des structures du DB, notamment le noyau sous-thalamique [19]. Les séquences d'inversion-récupération, comme le WAIR, seraient de bonnes candidates, seules ou combinées avec d'autres modalités, pour segmenter le DB [27].

Nous avons également observé que même si les indices de similarité étaient plus élevés pour les ensembles d'images WAIR + T1 et WAIR + FA, ils n'étaient pas statistiquement significatifs. La dérivée pdvFA proposée pour le DTI a obtenu de meilleurs résultats que la ccFA, ce qui montre que l'utilisation de la valeur absolue du vecteur n'est pas optimale pour les tâches de segmentation d'images. En ce qui concerne la réorientation, les résultats des expériences WAIR + ccFA et WAIR + pdvFA se sont améliorés après

l'application de la réorientation aux vecteurs (sans différence statistiquement significative). La différence significative dans les valeurs 95HD pendant l'entraînement indique qu'il existe des différences entre les expériences réorientées et non réorientées en termes de localisation des contours.

Les résultats obtenus en utilisant des combinaisons avec des images dérivées de la DTI ont montré des valeurs DSC, 95HD et VS inférieures à celles de la combinaison WAIR + T1. Cela pourrait s'expliquer par les caractéristiques spécifiques des voxels DTI. La différence de taille des voxels a pu entraîner une perte d'information importante lors de l'enregistrement. La complexité des données DTI, pdvFA, pdv et ccFA ont été enregistrées sous forme d'images 4D avec trois volumes, où chaque volume correspond à la composante du vecteur ou à la composante de la couleur RVB (ccFA). La DTI s'intéresse principalement aux mouvements de l'eau dans les axones, tandis que T1 et WAIR fournissent des informations structurelles plus détaillées sur le tissu cérébral. Enfin, la simple concaténation des volumes d'entrée utilisée dans ce travail est une fusion naïve de modalités différentes [9]. D'autres méthodes de fusion prenant en compte la complexité des relations inter-modalités seront envisagées dans nos futurs travaux [38]. En résumé, la cartographie du DB basée sur l'apprentissage profond peut être cliniquement pertinente, parallèlement à d'autres segmentations automatiques des structures sous-thalamiques, telles que les méthodes mixtes avec segmentation automatique des tissus sur des ensembles d'IRM individuels couplés à un contour basé sur un atlas [27] ou des

TABLE 2 – Scores de Dice moyens (DSC), distance de Hausdorff à 95% (95HD) et similarité volumétrique (VS) obtenus par des expériences avec différentes combinaisons de dérivés WAIR, T1 et DTI. Les résultats sont donnés avec (<sup>R</sup>) et sans (<sup>NR</sup>) réorientation du vecteur. La flèche pointant vers le bas (↓) indique que l’expérience a donné des résultats significativement moins bons que l’expérience WAIR. Les meilleurs résultats sont en gras.

Entrée	DSC ± STD	95HD ± STD	VS ± STD
WAIR	0,739 ± 0,073	1,739 ± 0,398	0,892 ± 0,044
T1	0,670 ± 0,106 ↓	2,516 ± 0,849 ↓	0,871 ± 0,048 ↓
WAIR + T1	<b>0,747 ± 0,073</b>	<b>1,712 ± 0,517</b>	<b>0,901 ± 0,039</b>
WAIR + ccFA <sup>NR</sup>	0,727 ± 0,078	1,872 ± 0,515	0,884 ± 0,045
WAIR + ccFA <sup>R</sup>	0,730 ± 0,078	1,807 ± 0,465	0,886 ± 0,046
WAIR + pdv <sup>NR</sup>	0,728 ± 0,071	1,864 ± 0,480	0,884 ± 0,041
WAIR + pdv <sup>R</sup>	0,728 ± 0,071	1,821 ± 0,423	0,886 ± 0,044
WAIR + pdvFA <sup>NR</sup>	0,723 ± 0,075	1,891 ± 0,464	0,885 ± 0,043
WAIR + pdvFA <sup>R</sup>	0,737 ± 0,073	1,831 ± 0,451	0,891 ± 0,049
WAIR + FA	0,740 ± 0,072	1,763 ± 0,476	0,899 ± 0,039
WAIR + T1 + FA	0,736 ± 0,075	1,778 ± 0,515	0,890 ± 0,043

TABLE 3 – Valeurs moyennes du DSC et écart-type pour les cinq structures segmentées (hémisphères gauche et droit) pour plusieurs expériences. L’indice R signifie que les gradients vectoriels ont été réorientés. La flèche pointant vers le bas (↓) indique que l’expérience a donné des résultats significativement moins bons pour chaque étiquette sur chaque hémisphère que l’expérience WAIR. Les meilleurs résultats sont en gras.

Label	Hem.	WAIR	T1	WAIR + T1	WAIR + pdvFA <sup>R</sup>	WAIR + FA	WAIR + T1 + FA
SN	G	0,793 ± 0,048	0,731 ± 0,060 ↓	<b>0,796 ± 0,044</b>	0,787 ± 0,045	0,782 ± 0,048	0,782 ± 0,042
	D	0,789 ± 0,046	0,701 ± 0,155 ↓	<b>0,791 ± 0,051</b>	0,776 ± 0,048	0,774 ± 0,049	0,777 ± 0,048
RN	G	0,756 ± 0,067	0,709 ± 0,122 ↓	<b>0,772 ± 0,055</b>	0,761 ± 0,059	0,757 ± 0,066	0,760 ± 0,062
	D	0,778 ± 0,059	0,732 ± 0,093 ↓	<b>0,784 ± 0,046</b>	0,778 ± 0,063	0,783 ± 0,064	0,776 ± 0,054
STN	G	0,698 ± 0,068	0,625 ± 0,168 ↓	<b>0,717 ± 0,049</b>	0,715 ± 0,049	0,711 ± 0,066	0,698 ± 0,072
	D	0,715 ± 0,068	0,661 ± 0,103 ↓	<b>0,722 ± 0,071</b>	0,719 ± 0,079	0,713 ± 0,076	0,693 ± 0,089
MB	G	0,807 ± 0,069	0,786 ± 0,091	0,816 ± 0,072	0,801 ± 0,068	0,810 ± 0,066	<b>0,821 ± 0,060</b>
	D	0,827 ± 0,083	0,792 ± 0,108 ↓	<b>0,832 ± 0,081</b>	0,819 ± 0,094	0,831 ± 0,089	0,827 ± 0,100
MT-fa	G	0,636 ± 0,138	0,476 ± 0,156 ↓	0,640 ± 0,135	0,644 ± 0,125	<b>0,647 ± 0,129</b>	0,634 ± 0,141
	D	0,593 ± 0,109	0,486 ± 0,139 ↓	<b>0,600 ± 0,113</b>	0,575 ± 0,128	0,587 ± 0,105	0,593 ± 0,107
<b>Moyenne</b>		0,739 ± 0,073	0,670 ± 0,106 ↓	<b>0,747 ± 0,073</b>	0,737 ± 0,073	0,740 ± 0,072	0,736 ± 0,075

méthodes basées modèle [14]. L’incorporation de connaissances anatomiques au processus d’apprentissage pourrait aider le réseau à segmenter des structures complexes, en particulier en présence d’un ensemble de données limité, par exemple en utilisant un atlas du cerveau pour guider la segmentation de la substance blanche [37] ou en utilisant des connaissances a priori de l’atlas comme entrées supplémentaires pour le réseau [5, 18].

## 5 Conclusion

Dans ce travail, nous avons segmenté les structures cérébrales profondes selon différentes modalités d’IRM. Notre contribution comprend un jeu de données multimodales avec la segmentation manuelle de cinq petites structures cérébrales profondes par un expert. À notre connaissance, c’est la première fois qu’un tel jeu de données a été construit avec une vérité terrain de contourage de ces structures dans un contexte clinique réellement pertinent. Les séquences d’inversion-récupération, comme le WAIR,

semblent être un bon choix dans ce domaine. En effet, elles offrent suffisamment de contraste pour un contourage manuel précis des structures, qui a servi de base pour l’entraînement d’un réseau de segmentation automatique. L’utilisation d’un entraînement multi-modal a donné des résultats encourageants (bien que non statistiquement significatifs), les meilleures valeurs DSC, 95HD et VS étant obtenues en utilisant la combinaison WAIR + T1. En outre, nous avons montré que la réorientation correcte des vecteurs DTI (par rapport à la ligne ACPC) améliorerait le processus de segmentation.

## Remerciements

Ce travail a été cofinancé par le Fonds européen de développement régional (FEDER ; UNION EUROPEENNE ; projet AIM), l’Agence nationale de la recherche française (bourse de doctorat ANR-IA) et la Fondation de l’Avenir (programme Ptolémée, 21\_0099\_01 ; financement de doctorat).

Les calculs ont été effectués dans les installations de supercalculateurs du Mésocentre Clermont-Auvergne à l'Université Clermont Auvergne.

## Références

- [1] M. Baniyadi, M. V. Petersen, J. Gonçalves, A. Horn, V. Vlasov, F. Hertel, and A. Husch. DBSegment : Fast and robust segmentation of deep brain structures considering domain generalization. *Human Brain Mapping*, 44(2) :762–778, 2023.
- [2] S. Cai, Y. Tian, H. Lui, H. Zeng, Y. Wu, and G. Chen. Dense-UNet : a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network. *Quantitative Imaging in Medicine and Surgery*, 10(6), 2020.
- [3] J. Chen, H. Xu, B. Xu, Y. Wang, Y. Shi, and L. Xiao. Automatic Localization of Key Structures for Subthalamic Nucleus–Deep Brain Stimulation Surgery via Prior-Enhanced Multi-Object Magnetic Resonance Imaging Segmentation. *World Neurosurgery*, 178 :e472–e479, 2023.
- [4] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3D U-Net : Learning Dense Volumetric Segmentation from Sparse Annotation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, pages 424–432, Cham, 2016. Springer International Publishing.
- [5] P. Coupé, B. Mansencal, M. Clément, R. Giraud, B. Denis de Senneville, V.-T. Ta, V. Lepetit, and J. V. Manjon. AssemblyNet : A large ensemble of CNNs for 3D whole brain MRI segmentation. *NeuroImage*, 219 :117026, 2020.
- [6] J. Dejerine. *Anatomie des centres nerveux*, volume 1. Rueff (Paris), 1901.
- [7] L. R. Dice. Measures of the Amount of Ecologic Association Between Species. *Ecology*, 26(3) :297–302, 1945.
- [8] J. Dolz, C. Desrosiers, and I. Ben Ayed. 3d fully convolutional networks for subcortical segmentation in MRI : A large-scale study. *NeuroImage*, 170 :456–470, 2018. Segmenting the Brain.
- [9] J. Dolz, K. Gopinath, J. Yuan, H. Lombaert, C. Desrosiers, and I. Ben Ayed. HyperDense-Net : A Hyper-Densely Connected CNN for Multi-Modal Image Segmentation. *IEEE Transactions on Medical Imaging*, 38(5) :1116–1126, 2019.
- [10] Y. El Ouidih, A. Marques, B. Pereira, M. Luisoni, B. Claise, J. Coste, A. Sontheimer, R. Chaix, B. Debilly, P. Derost, D. Morand, F. Durif, and J.-J. Le-maire. Deep brain stimulation of the subthalamic nucleus in severe Parkinson’s disease : relationships between dual-contact topographic setting and 1-year worsening of speech and gait. *Acta Neurochirurgica*, 165(12) :3927–3941, Dec 2023.
- [11] S. H. Fox, R. Katzenschlager, S.-Y. Lim, B. Barton, R. M. A. de Bie, K. Seppi, M. Coelho, C. Sampaio, and on behalf of the Movement Disorder Society Evidence-Based Medicine Committee. International Parkinson and movement disorder society evidence-based medicine review : Update on treatments for the motor symptoms of Parkinson’s disease. *Movement Disorders*, 33(8) :1248–1266, 2018.
- [12] D. N. Greve, B. Billot, D. Cordero, A. Hoopes, M. Hoffmann, A. V. Dalca, B. Fischl, J. E. Iglesias, and J. C. Augustinack. A deep learning toolbox for automatic segmentation of subcortical limbic structures from MRI images. *NeuroImage*, 244 :118610, 2021.
- [13] O. F. Gulban, D. Nielson, John Lee, R. Poldrack, C. Gorgolewski, Vanessasaurus, and C. Markiewicz. poldracklab/pydeface : PyDeface v2.0.2, July 2022.
- [14] C. Haegelen, P. Coupé, V. Fonov, N. Guizard, P. Jannin, X. Morandi, and D. L. Collins. Automated segmentation of basal ganglia and deep brain structures in MRI of Parkinson’s disease. *International Journal of Computer Assisted Radiology and Surgery*, 8(1) :99–110, Jan 2013.
- [15] R. Hannah and A. R. Aron. Towards real-world generalizability of a circuit for action-stopping. *Nature Reviews Neuroscience*, 22(9) :538–552, Sep 2021.
- [16] I. E. Harmsen, G. J. Elias, M. E. Bejn, A. Boutet, A. Pancholi, J. Germann, A. Mansouri, C. S. Lozano, and A. M. Lozano. Clinical trials for deep brain stimulation : Current state of affairs. *Brain Stimulation*, 13(2) :378–385, 2020.
- [17] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein. nnU-Net : A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2) :203–211, 2020.
- [18] K. Kushibar, S. Valverde, S. González-Villà, J. Bernal, M. Cabezas, A. Oliver, and X. Lladó. Automated sub-cortical brain structure segmentation combining spatial and deep convolutional features. *Medical Image Analysis*, 48 :177–186, 2018.
- [19] M. Lefranc, S. Derrey, P. Merle, M. Tir, J.-M. Constans, D. Montpellier, J. M. Macron, D. Le Gars, J. Peltier, O. Baledent, and P. Krystkowiak. High-resolution 3-dimensional T2\*-weighted angiography (HR 3-D SWAN) : an optimized 3-T magnetic resonance imaging sequence for targeting the subthalamic nucleus. *Neurosurgery*, 74(6) :615–26; discussion 627, June 2014.

- [20] J.-J. Lemaire, A. De Salles, G. Coll, Y. El Oquadhi, R. Chaix, J. Coste, F. Durif, N. Makris, and R. Kikinis. MRI Atlas of the Human Deep Brain. *Frontiers in Neurology*, 10, 2019.
- [21] J.-J. Lemaire, A. Sontheimer, B. Pereira, J. Coste, S. Rosenberg, C. Sarret, G. Coll, J. Gabrillargues, B. Jean, T. Gillart, A. Coste, B. Roche, A. Kelly, B. Pontier, and F. Feschet. Deep brain stimulation in five patients with severe disorders of consciousness. *Annals of Clinical and Translational Neurology*, 5(11) :1372–1384, 2018.
- [22] M. S. Majdi, M. B. Keerthivasan, B. K. Rutt, N. M. Zahr, J. J. Rodriguez, and M. Saranathan. Automated thalamic nuclei segmentation using multi-planar cascaded convolutional neural networks. *Magnetic Resonance Imaging*, 73 :45–54, 2020.
- [23] S. Pajevic and C. Pierpaoli. Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data : application to white matter fiber tract mapping in the human brain. *Magnetic Resonance in Medicine*, 43 :526, 1999.
- [24] S. D. Pieper, M. W. Halle, and R. Kikinis. 3D Slicer. *2004 2nd IEEE International Symposium on Biomedical Imaging : Nano to Macro (IEEE Cat No. 04EX821)*, pages 632–635 Vol. 1, 2012.
- [25] G. R. Pinheiro, D. S. Carmo, C. Yasuda, R. A. Lotufo, and L. Rittner. Convolutional Neural Network on DTI Data for Sub-cortical Brain Structure Segmentation. In *Computational Diffusion MRI*, pages 135–146, Cham, 2020. Springer International Publishing.
- [26] E. A. Rashed, J. Gomez-Tames, and A. Hirata. End-to-end semantic segmentation of personalized deep brain structures for non-invasive brain stimulation. *Neural Networks*, 125 :233–244, 2020.
- [27] P. Reinacher, B. Várkuti, M. Krüger, T. Piroth, K. Egger, R. Roelz, and V. Coenen. Automatic Segmentation of the Subthalamic Nucleus : A Viable Option to Support Planning and Visualization of Patient-Specific Targeting in Deep Brain Stimulation. *Operative Neurosurgery*, 17, 03 2019.
- [28] O. Ronneberger, P. Fischer, and T. Brox. U-Net : Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.
- [29] G. Schaltenbrand and P. Bailey. *Einführung in die stereotaktischen Operationen : mit einem Atlas des menschlichen Gehirns. Introduction to stereotaxis, with an atlas of the human brain*. Number vol. 1 in Einführung in die stereotaktischen Operationen. G. Thieme, 1959.
- [30] H. Seo, C. Huang, M. Bassenne, R. Xiao, and L. Xing. Modified u-net (mu-net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in ct images. *IEEE Transactions on Medical Imaging*, 39(5) :1316–1325, 2020.
- [31] A. L. Simpson, M. Antonelli, S. Bakas, M. Bilello, K. Farahani, B. van Ginneken, A. Kopp-Schneider, B. A. Landman, G. Litjens, B. Menze, O. Ronneberger, R. M. Summers, P. Bilic, P. F. Christ, R. K. G. Do, M. Gollub, J. Golia-Pernicka, S. H. Heckers, W. R. Jarnagin, M. K. McHugo, S. Napel, E. Vorontsov, L. Maier-Hein, and M. J. Cardoso. The Medical Segmentation Decathlon. *Nature Communications*, 13(1), July 2022.
- [32] A. A. Taha and A. Hanbury. Metrics for evaluating 3d medical image segmentation : analysis, selection, and tool. *BMC Medical Imaging*, 15(1) :29, Aug 2015.
- [33] F. Vassal, J. Coste, P. Derost, V. Mendes, J. Gabrillargues, C. Nuti, F. Durif, and J.-J. Lemaire. Direct stereotactic targeting of the ventrointermediate nucleus of the thalamus based on anatomic 1.5-T MRI mapping with a white matter attenuated inversion recovery (WAIR) sequence. *Brain Stimulation*, 5(4) :625–633, 2012.
- [34] A. Zerroug, J. Gabrillargues, G. Coll, F. Vassal, B. Jean, E. Chabert, B. Claise, T. Khalil, L. Sakka, F. Feschet, F. Durif, L. Boyer, J. Coste, and J.-J. Lemaire. Personalized mapping of the deep brain with a white matter attenuated inversion recovery (WAIR) sequence at 1.5-tesla : Experience based on a series of 156 patients. *Neurochirurgie*, 62(4) :183–189, 2016.
- [35] H. Zhang, B. B. Avants, P. A. Yushkevich, J. H. Woo, S. Wang, L. F. McCluskey, L. B. Elman, E. R. Melhem, and J. C. Gee. High-Dimensional Spatial Normalization of Diffusion Tensor Images Improves the Detection of White Matter Differences : An Example Study Using Amyotrophic Lateral Sclerosis. *IEEE Transactions on Medical Imaging*, 26(11) :1585–1597, 2007.
- [36] S. Zhang, D. H. Laidlaw, and G. Kindlmann. 16 - Diffusion Tensor MRI Visualization. In C. D. Hansen and C. R. Johnson, editors, *Visualization Handbook*, pages 327–340. Butterworth-Heinemann, Burlington, 2005.
- [37] Z. Zhang, K. Powell, C. Yin, S. Cao, D. Gonzalez, Y. Hannawi, and P. Zhang. Brain Atlas Guided Attention U-Net for White Matter Hyperintensity Segmen-

tation. *AMIA Jt Summits Transl Sci Proc*, 2021 :663–671, 2021.

- [38] T. Zhou, S. Ruan, and S. Canu. A review : Deep learning for medical image segmentation using multi-modality fusion. *Array*, 3-4 :100004, 2019.