Automated segmentation of deep brain structures from Inversion-Recovery MRI

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

Computerized Medical Imaging and Graphics 120 (2025) 102488

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Automated segmentation of deep brain structures from Inversion-Recovery MRI

Aigerim Dautkulova [a,*], Omar Ait Aider [a], Céline Teulière [a], Jérôme Coste [a,b], Rémi Chaix [a,b], Omar Ouachik [a], Bruno Pereira [c], Jean-Jacques Lemaire [a,b]

[a] Université Clermont Auvergne, Clermont Auvergne INP, CNRS, Institut Pascal, F-63000 Clermont-Ferrand, France
[b] Université Clermont Auvergne, CNRS, CHU Clermont-Ferrand, Clermont Auvergne INP, Institut Pascal, F-63000 Clermont-Ferrand, France
[c] Direction de la Recherche et de l'Innovation, CHU Clermont-Ferrand, F-63000 Clermont-Ferrand, France

ARTICLE INFO

Keywords:
Deep brain
Automated segmentation
WAIR
DTI
ACPC

ABSTRACT

Methods for the automated segmentation of brain structures are a major subject of medical research. The small structures of the deep brain have received scant attention, notably for lack of manual delineations by medical experts. In this study, we assessed an automated segmentation of a novel clinical dataset containing White Matter Attenuated Inversion-Recovery (WAIR) MRI images and five manually segmented structures (substantia nigra (SN), subthalamic nucleus (STN), red nucleus (RN), mammillary body (MB) and mammillothalamic fascicle (MT-fa)) in 53 patients with severe Parkinson's disease. T1 and DTI images were additionally used. We also assessed the reorientation of DTI diffusion vectors with reference to the ACPC line. A state-of-the-art nnU-Net method was trained and tested on subsets of 38 and 15 image datasets respectively. We used Dice similarity coefficient (DSC), 95% Hausdorff distance (95HD), and volumetric similarity (VS) as metrics to evaluate network efficiency in reproducing manual contouring. Random-effects models statistically compared values according to structures, accounting for between- and within-participant variability. Results show that WAIR significantly outperformed T1 for DSC (0.739 ± 0.073), 95HD (1.739 ± 0.398), and VS (0.892 ± 0.044). The DSC values for automated segmentation of MB, RN, SN, STN, and MT-fa decreased in that order, in line with the increasing complexity observed in manual segmentation. Based on training results, the reorientation of DTI vectors improved the automated segmentation.

1. Introduction

The deep brain (DB) is a highly sophisticated region (Dejerine, 1901; Haber et al., 2012; Lemaire et al., 2019; Mai et al., 2016; Nieuwenhuys et al., 2008; Riley, 1943) of clinical importance in many neuropsychiatric diseases, such as Parkinson's disease, Alzheimer's disease, and compulsive disorders, supporting complex behavior and motor controls (see for example Bohnen et al. (2022) and Hannah and Aron (2021)). In Parkinson's disease, when severe motor conditions occur, medication adjustments can be combined with chronic deep brain stimulation (DBS) to improve quality of life (Fox et al., 2018). The efficacy of this dual therapy relies notably on the correct location of contacts in the superolateral region of the subthalamic nucleus (STN) (Harmsen et al., 2020), by spotting the STN target in the subthalamus.

The target is commonly represented on magnetic resonance imaging (MRI) slices, onto which are superimposed stereotactic landmarks, for example the endpoint of trajectory and/or the contour of STN, supervised by the neurosurgeon. Stereotactic atlases, such as the well-known Schaltenbrand and Bailey atlas (Schaltenbrand and Bailey, 1959) can be matched using proportionality according to ventricular commissural points (anterior and posterior white commissures, AC and PC). Most clinicians interpret the MRI contrasts according to type of sequence to optimize targeting. In our institution, we developed a dedicated inversion-recovery sequence, named WAIR (White Matter Attenuated Inversion-Recovery) to facilitate the identification of DB structures directly on coronal slices (El Ouadih et al., 2023; Zerroug et al., 2016). Briefly, (i) the higher the cellularity (e.g., the substantia nigra, notably the pars compacta), the stronger the signal, (ii) the higher the axonal density (e.g., the mammillo-thalamic fascicle), the weaker the signal, and (iii) the red nucleus, a gray matter structure, has a weak signal because it contains a significant contingent of axons. This sequence has been used in clinical routine since 2006 with a 1.5 tesla machine, and the STN and its neighboring structures are manually contoured,

* Corresponding author.
E-mail address: aigerim.dautkulova@uca.fr (A. Dautkulova).

https://doi.org/10.1016/j.compmedimag.2024.102488
Received 21 February 2024; Received in revised form 5 September 2024; Accepted 30 December 2024
Available online 3 January 2025

enabling detailed 3D planning for electrode implantation (see Supplementary Material 1 Fig. S.2). However, this precise individualized mapping costs 1 h 30 min of worktime for the two hemispheres because of the awkward 3D architecture and the difficult interpretation of spontaneous MRI contrast at the individual level. Over time, we decided to automate the segmentation of subthalamic structures to simplify the personalization of mapping, with a reasonable computational time of a few tens of minutes. For this we chose a deep learning (DL) approach applied to our clinical electrode planning database.

One of the most widely used DL-based image segmentation models is U-Net (Ronneberger et al., 2015), which relies on a U-shaped convolutional neural network. Several modified versions of this architecture have been proposed, such as 3D U-Net (Çiçek et al., 2016), V-Net (Milletari et al., 2016), Modified U-Net (Seo et al., 2020), and Dense-UNet (Cai et al., 2020). Isensee et al. (2020) consider that a well-configured plain U-Net is still difficult to surpass, and introduced an nnU-Net, a self-configuring network using a modified standard 3D U-Net. The nnU-Net automatically configures all the necessary steps of the automated segmentation pipeline, including the preprocessing, architecture of the network, training variables, and postprocessing steps. It yielded state-of-the-art results in various tasks, including segmentation of the hippocampus of the Medical Segmentation Decathlon Challenge (Antonelli et al., 2022). For our purpose, we selected the 3D configuration of nnU-Net with a standard 3D U-Net as a backbone, as 3D would be most appropriate (Avesta et al., 2023; Shivdeo et al., 2021).

Precise delineation of the DB architecture is challenging, notably because of the small sizes of most structures, which are little known and often coalescent, merging cell bodies and axon fascicles, leading to ranging MRI contrasts. Consequently, the ground truth datasets, atlas-based or manually contoured on MRI images are very limited. Most of the DL segmentation methods have used T1 image datasets (Dolz et al., 2018; Kushibar et al., 2018; Rashed et al., 2020) (Table 1). In our work, along with the WAIR, we also considered two other MRI sequences acquired at surgery, namely 3D T1 MPRAGE, and the seldom used diffusion tensor image (DTI) (Pinheiro et al., 2020) for the full 3D electrode planning and postoperative control.

We assessed the DL segmentation of five DB structures using nnU-Net on a new clinical dataset with an Inversion-Recovery sequence and its combinations with T1 and DTI. We evaluated the efficiency of the trained self-configured 3D nnU-Net (n = 38 patients) on a test set (n = 15 patients) to reproduce the manual contouring of five subthalamic structures in the right and left hemispheres. The dataset with ground truth labels (i.e., the five structures manually contoured on WAIR images by a clinical expert) is publicly available.[1] We investigated the effectiveness of the WAIR and T1 sequences, and the DTI derivatives and their combinations for the task. We also evaluated the impact of reorienting principal diffusion vectors with respect to the patient coordinate system to reduce interindividual variability due to slight differences in head positioning in the stereotactic frame in the MRI machine.

In summary, our contributions are as follows:

- First use of WAIR MRI for deep learning-based image segmentation,
- New clinical dataset with ground truth labels of DB structures manually segmented by an expert clinician,
- Evaluation of performance improvement by multi-modality (WAIR + T1, WAIR + DTI),
- Impact of DTI vector reorientation with respect to the ACPC line of each patient.

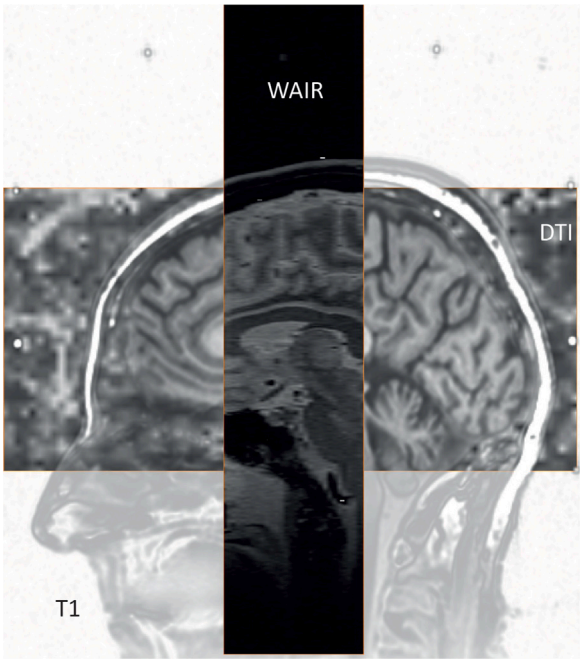

Fig. 1. Overlay of co-registered WAIR, T1, and DTI image datasets.

2. Materials and methods

Here we present the clinical material used in the work (Section 2.1), the geometric space preprocessing (Section 2.2.1), and the representation and transformation of the DTI (Section 2.2.2). In Section 2.2.3 we describe the reorientation of DTI vectors. In Section 2.3 we detail the training and implementation of the automated segmentation. In Section 2.4 we describe the evaluation metrics. Lastly, in Section 2.5 we provide details of the statistical analysis.

2.1. Clinical material

The dataset consisted of images of 53 patients with severe Parkinson's disease (61 ± 6 years, 26 female), implanted bilaterally in the STN, in a single institution (Expert Center for Parkinson's Disease), between November 2011 and November 2022. Ethical, institutional, and individual approvals were obtained (Ptolémée® program: IRB 5921, March 30, 2018; CNIL ref. MR 004, M200702 November 10, 2020). Of the 110 patients operated over the period who agreed to participate, 42 were excluded because they had other disorders (of which 32 essential tremor), 12 because the electrodes were implanted in other targets, and 3 because the DTI datasets were corrupted. Each individual dataset consisted of (i) WAIR, T1, and DTI datasets, co-registered at acquisition (stereotactic frame secured to the head and locked in the head coil) (Fig. 1), and (ii) the objects built from the contours of structures, exported in stereolithography (stl) format (Iplan 3.0®, BrainLab, Germany). For further analysis, stl objects were converted to Neuroimaging Informatics Technology Initiative (NIfTI) file format (3D Slicer software[2] Pieper et al., 2012). The three sequences were acquired on a 1.5-tesla MRI machine (Aera and Sonata; Siemens, Germany). All the data was anonymized and T1 images were defaced using the pydeface software package (Gulban et al., 2022).

The WAIR sequence was developed to increase the spontaneous tissue contrasts between and within the DB structures (Lemaire et al., 2018; Vassal et al., 2012; Zerroug et al., 2016). The WAIR settings

[1] https://osf.io/49arj/?view_only=b319d34151dd4905ad9052d2998e9ecd.

[2] https://www.slicer.org/.

Table 1
State-of-the-art methods in deep brain structure segmentation (n.a. not available).

Authors (year) and method	Dataset reference	MRI sequence, magnetic field (number of cases)	Annotation method	Free access	Structures
Dolz et al. (2018) LiviaNET	IBSR[a] ABIDE[b]	T1, n.a. (18) T1, n.a. (1112)	Manual Atlas-based	Yes No	Thalamus, Caudate, Putamen, Pallidum
Kushibar et al. (2018) 2.5D CNN	MICCAI 2012[c] IBSR	T1, n.a. (35) T1, n.a. (18)	Manual Manual	Yes Yes	Thalamus, Caudate, Putamen, Pallidum, Amygdala, Hippocampus, Accumbens
Rashed et al. (2020) SubForkNet	NAMIC[d] MICCAI 2012	T1, n.a. (18) T1, n.a. (35)	Semi-automatic Manual	No Yes	Thalamus, Caudate, Putamen, Pallidum, Amygdala, Hippocampus, Accumbens
Majdi et al. (2020) Modified U-Net	Internal	T1 MPRAGE, 3T (105), 7T (40)	Manual (left side)	No	Thalamus, 11 thalamic nuclei, Mammillothalamic tract
Baniasadi et al. (2023) nnU-Net	10 public and 4 internal	T1, mixed, 1.5T, 3T (1028)	Atlas-based	No	Caudate, Globus Pallidus (Externus, Internus), Habenular Nuclei, Internal capsule, Accumbens, Putamen, Red nucleus, Substantia nigra (pars compacta, pars reticulata), Subthalamic nucleus, Thalamus, Ventral lateral posterior nucleus of Thalamus, Lateral ventricle, Ventrointermediate nucleus of Thalamus
Greve et al. (2021) 3D U-Net	FSM[e] ADNI[f]	T1, 3T (29) T1, n.a. (10)	Manual	On request	Hypothalamus, Mammillary bodies, Septal nuclei, Fornix, Accumbens, Basal forebrain
Pinheiro et al. (2020) 2D U-Net	Internal	T1, 3T, DWI (121)	Manual	No	Ventricle, Caudate, Putamen, Thalamus, Globus pallidus, Hippocampus, Amygdala, and Accumbens
Our study nnU-Net	CLT Subthalamus Dataset	WAIR, T1, DTI (53)	Manual	Yes	Substantia nigra, Subthalamic nucleus, Red nucleus, Mammillary body, and Mammillothalamic fascicle

[a] https://www.nitrc.org/projects/ibsr (accessed 05-February-2024).
[b] https://fcon_1000.projects.nitrc.org/indi/abide/ (accessed 05-February-2024).
[c] http://www.neuromorphometrics.com/2012_MICCAI_Challenge_Data.html (accessed 05-February-2024).
[d] https://www.na-mic.org/w/index.php?title=Main_Page&oldid=98830 (accessed 27-June-2023).
[e] https://surfer.nmr.mgh.harvard.edu/fswiki/FsmData (accessed 27-June-2023).
[f] https://adni.loni.usc.edu (accessed 05-February-2024).

were as follows: 2D coronal, repetition time 4500 ms, echo time 13 ms, inversion time 160 ms, flip angle 170°, pixel size 0.527×0.527 mm^2, slice thickness 2 mm, matrix 512 × 512 × 30 (Supplementary Material 1 Fig. S.1). The coronal orientation was used to facilitate the identification of subthalamic structures in the plane perpendicular to the rostro-caudal axis of the central nervous system. The manual delineation of all structures on WAIR images was performed by one clinical expert (JJL) using the same surgical software (Iplan 3.0®, BrainLab, Germany) (Fig. 2; Supplementary Material 1 Fig. S.2). For our purpose, we selected five characteristic structures in each hemisphere: the substantia nigra (SN), the subthalamic nucleus (STN), the red nucleus (RN), the mammillary body (MB), and the mammillothalamic fascicle (MT-fa) (Fig. 3). We computed volumes for these structures in mm^3 from NIfTI volumes (Supplementary Material 2 Tables S.1 and S.2). T1 and DTI settings were as follows: T1, 3D MPRAGE, sagittal, repetition time 1660 ms, echo time 3.35 ms, flip angle 15°, pixel size 0.625 × 0.625 mm^2, slice thickness 2 mm, matrix 480 × 512 × 80, DTI, epi axial, 20 directions, b0 = 750 s/mm^2, repetition time 5300 ms, echo time 82 ms, pixel size 3 × 3 mm^2, thickness 3 mm, matrix 80 × 80 × 40.

2.2. Preprocessing

2.2.1. Geometric space

The T1 and DTI image datasets were aligned to the WAIR, on which the structures were contoured, and registered using 3D Slicer. The General Registration (BRAINS) tool was used for T1 images (ITK toolkit, McCormick et al., 2014). We used an affine transformation with 12 degrees of freedom, suitable for intra-subject registration.

For the DTI images, firstly, the baseline (non-diffusion-weighted) volume was obtained using the Diffusion Brain Masking module of the SlicerDMRI (Zhang et al., 2020) extension of the 3D Slicer. Then by registering a baseline volume to WAIR using the General Registration (BRAINS) tool (the same parameters as T1), we obtained a transformation matrix. Finally, using the obtained matrix, the Resample DTI Volume module (based on the ITK toolkit) was employed to align the DTI to WAIR. Where the WAIR serves as a reference volume and sets the origin, spacing, orientation, and dimension for the resampled image. The resulting image size was 512 × 512 × 30 with a voxel size of 0.53 × 0.53 × 2.0 mm^3.

2.2.2. DTI representation and transformation

The DTI voxel contains a 3 × 3 diffusion tensor that can be decomposed into three eigenvalues (λ_1, λ_2, λ_3) and three eigenvectors (v_1, v_2, v_3), which respectively represent the amount and orientation of proton diffusion. The principal eigenvector v_1 shows the direction of the maximal diffusion with the largest eigenvalue λ_1 (Zhang et al., 2005). In our experiments, we mapped four derivatives of DTI (Fig. 4):

- **Fractional anisotropy map (FA).** FA is an orientation-invariant scalar value that describes the degree of anisotropy. It ranges between 0 (isotropic) and 1 (anisotropic).
- **Principal diffusion direction vector map (pdv).** This is a 3D image with three channels representing the three components of the principal eigenvector v_1 at each voxel.
- **Color-coded fractional anisotropy map (ccFA).** The color-coded FA map is generated by multiplying the principal eigenvector of the diffusion tensor, with the FA map. We follow the general color coding (Pajevic and Pierpaoli, 1999), where the ccFA shows both how directionally intense the diffusion is and its direction (red is left–right, blue is superior–inferior, and green is anterior–posterior). The commonly used formula is:

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

  where v_{1x}, v_{1y}, v_{1z} are the x, y, and z components of the vector v_1.

- **Weighted principal diffusion direction map (pdvFA).** For this study, we created a pdvFA map that contains both the degree of anisotropy and vector orientation information. Unlike ccFA, where the absolute values of the vector are used, pdvFA preserves

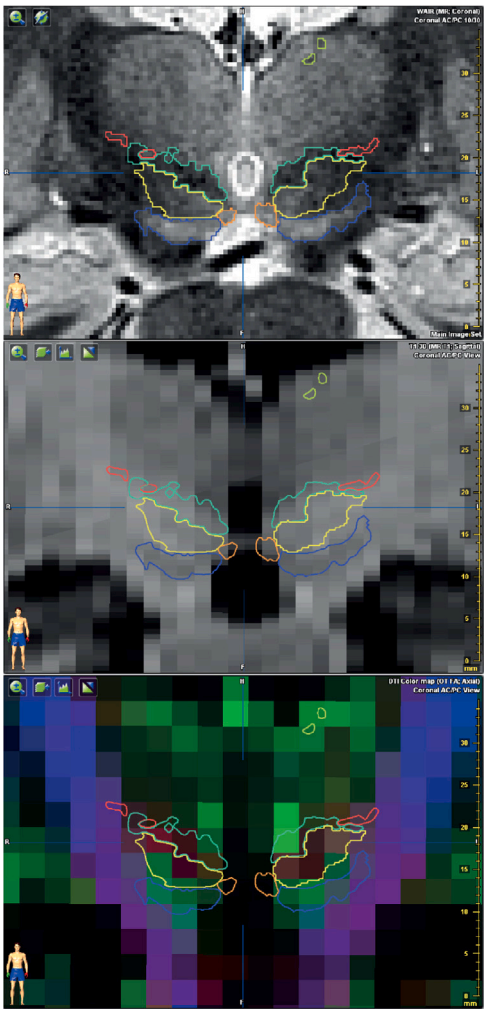

Fig. 2. Contourings of five subthalamic structures selected for this work overlaid on WAIR (top), T1 (intermediate), and color-coded map (bottom) (no image filtering; Iplan 3.0®, BrainLab, Germany), along the same coronal slice through the STN: subthalamic nucleus, yellow; substantia nigra, blue; Forel's fields, green; nucleus of ansa lenticularis, orange; zona incerta, red; mammillo-thalamic fascicle, pale green.

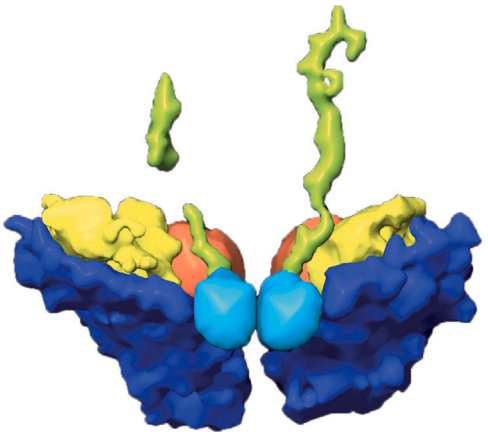

Fig. 3. 3D anterior view of the five DB structures selected for this work (same patient as Fig. 1): subthalamic nucleus, yellow; substantia nigra, dark blue; mammillo-thalamic fascicle, pale green; red nucleus, red; mammillary body, pale blue (Iplan 3.0®, BrainLab, Germany).

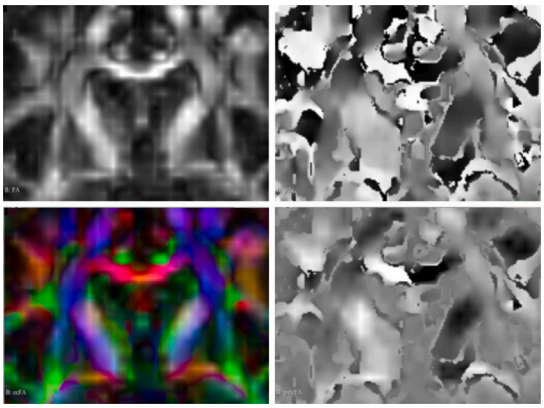

Fig. 4. Same coronal slice of four DTI derivatives, without gradient reorientation: fractional anisotropy map (FA, top left), principal diffusion direction vector map (pdv, top right), color-coded fractional anisotropy map (ccFA, bottom left), and weighted principal diffusion direction map (pdvFA, bottom right).

the sign of v_1, (i.e., all directional information) by directly multiplying vector v_1 by FA, thus enabling consistent reorientation of DTI data:

$$\mathrm{pdvFA_x} = v_{1x} \cdot FA, \mathrm{pdvFA_y} = v_{1y} \cdot FA, \quad \mathrm{pdvFA_z} = v_{1z} \cdot FA \tag{2}$$

We chose FA and ccFA because they are common and standardized derivatives of DTI (O'Donnell and Westin, 2011). pdv is usually visualized as an RGB image, but we used the original vector representation of this derivative. Diffusion tensor estimation and resampling of the DTI to the reference of WAIR was carried out with 3D Slicer. TVtool command from DTI-TK toolbox[3] (Zhang et al., 2007) was used to compute diffusion images derived from DTI (FA, ccFA, pdv).

2.2.3. ACPC reorientation of DTI vectors

We systematically reoriented the raw direction of diffusion gradients of the machine along the ACPC line to optimize the consistency of diffusion data across the patients. Before acquisition, in routine clinical conditions, the head positioning was optimized in the head coil using standard external landmarks (orbito-meatal and midline planes). Then, afterward, we reoriented the table of gradients according to AC and PC landmarks (manual positioning; double control; JJL, OO) enabling to compute a 3 × 3 matrix further applied to the DTI derivatives. Details on ACPC transform and computation are available in Supplementary Material 3. DTI derivates were computed with and without the application of the transform. In the following sections, we refer to the images with transformed and non-transformed direction gradients as reoriented (^{R}) and non-reoriented (^{NR}).

2.3. Training and implementation

We used the 3D configuration of nnU-Net (Fig. 5), where 3D U-Net was used as an architecture of their DL method. We changed the number of epochs to optimize the training time from 1000 to 300. We split the dataset (n = 53) into training (n = 38, 15 female, mean age 61) and test (or inference) (n = 15, 11 female, mean age 61) subsets. We ensured fair distribution of the acquisition dates of image datasets (see Supplementary Material 1 Fig. S.3), aiming to minimize the impact of the learning curve of the manual drawing. Consequently, the sex ratio between the two subsets was different, with more females in the test subset (73% versus 39% in the training subset). For training multimodal models with WAIR, T1, and DTI sequences, we use early fusion

[3] http://dti-tk.sourceforge.net/.

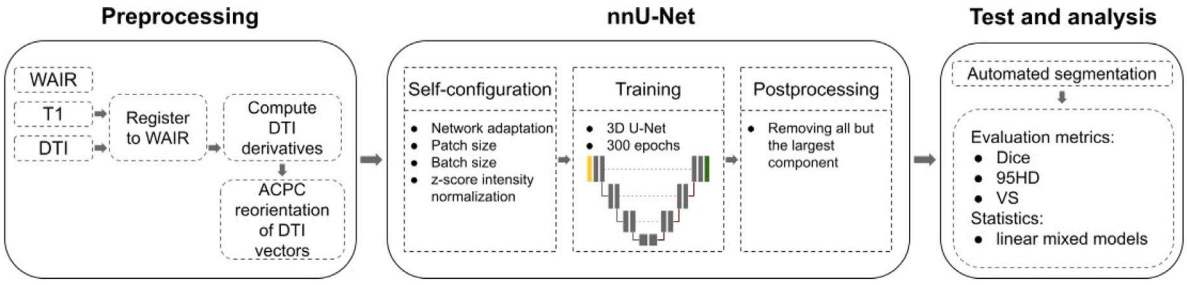

Fig. 5. Workflow: dataset preparation included preprocessing of T1 and DTI data; 3D nnU-Net method; testing and evaluation.

by concatenating these inputs into a single multi-channel image. The default preprocessing steps were applied to WAIR, T1, ccFA, and FA images. We skipped the intensity normalization step for vector images (pdv and pdvFA) as they are already normalized (unit vectors). For the inference process, we used the best-obtained checkpoint of the DL model (trained nnU-Net), instead of the last-obtained checkpoint. Details of the network configuration for each experiment are shown in Supplementary Material Table S.1. The batch size of 2 was selected by nnU-Net for all experiments. Originally, nnU-Net was implemented in Python 3.8.5 using PyTorch framework 1.6.0. Our computations were performed in a supercomputer facility using a GPU with 16 GB memory.

2.4. Evaluation metrics

To assess the efficiency of the automated segmentation of the five deep brain structures, we used three similarity metrics:

- **Dice coefficient (DSC)** (Dice, 1945)**.** This is the metric most widely used to validate medical volume segmentations for the direct comparison between automatic and ground truth segmentations. The DSC measures the spatial overlap between two segmentations. The DSC ranges between 0 and 1, with 1 indicating perfect overlap.
- **95% Hausdorff distance (95HD)**. This is a measure of dissimilarity between two point sets (Taha and Hanbury, 2015). 95HD is similar to maximum HD. However, it is based on the 95th percentile of the distances instead of the maximum. A smaller 95% Hausdorff distance indicates better alignment and agreement between the segmentations. It is measured in millimeters.
- **Volumetric similarity (VS)**. This computes the similarity between two segments based on their volumes (Taha and Hanbury, 2015). The VS ranges from 0 to 1, where 1 indicates the highest volumetric similarity between the automated segmentation and GT.

  – We consider the DSC as the main evaluation metric and 95HD and VS as additional metrics.

2.5. Statistical analysis

The performance of our trained self-configured 3D nnU-Net was analyzed on the test subset (n = 15 patients), according to the values of indices of similarity (DSC, 95HD, and VS) of the pairs of contours (manual and automatic) of the five structures (SN, STN, RN, MB, and MT-fa). The analysis was conducted by image datasets (WAIR, T1, WAIR + T1, WAIR + ccFA$^{NR,\ R}$, WAIR + pdv$^{NR,\ R}$, WAIR + pdvFA$^{NR,\ R}$, WAIR + FA, WAIR + T1 + FA; WAIR as a reference because this was the ground truth condition), by structure (SN, STN, RN, MB, and MT-fa; STN as reference, because it was the clinical goal of manual contouring) and hemisphere (right and left). We also specifically evaluated the impact of reorientating the diffusion gradients on WAIR + ccFA$^{NR,\ R}$, WAIR + pdv$^{NR,\ R}$, WAIR + pdvFA$^{NR,\ R}$ datasets.

Continuous data are presented as mean and standard deviation. The assumption of normality was checked using the Shapiro–Wilk test. Random-effects models (i.e., linear mixed models) were run to compare DSC, 95HD, and VS between structures considering between- and within-participant variability (subject as random-effect). The following fixed effects were analyzed: image datasets, structures, diffusion gradient reorientation ($^{R}/^{NR}$), and their interactions. The normality of residuals was studied as stated above and logarithmic transformation was performed to achieve normality when appropriate. A stepwise approach was proposed to simultaneously test participants, imaging sessions, structures and diffusion gradient reorientation as random effects, first testing each random effect separately and then combining significant effects, nested or crossed. The intra-class correlation coefficient was calculated for each random effect and comparisons of the performance of each model (log-likelihood) were performed. All analyses were conducted with an adjustment on volumes. Statistical analyses were conducted with Stata statistical software (version 15, StataCorp, College Station, USA). The tests were two-sided, with a type-1 error set at 5%. A Sidak type-1 error was applied to take into account two-by-two multiple comparisons. The results were expressed using Hedge's effect sizes and 95% CIs (available on demand).

3. Results

The average computation time of each test experiment was around 5 min. The average computation time of each training optimization was around 30 h for WAIR and T1, around 35 h for WAIR + T1 and WAIR + FA, and around 60 h for the other experiments.

Overall DL model comparison. Table 2 shows the average DSC, 95HD, and VS results for the trained DL models. Table 3 shows the results for each structure for several experiments (see Supplementary Material 1 Tables S.2, S.3, and S.4 for additional results). We found that the WAIR experiment significantly outperformed the T1 experiment in terms of DSC, 95HD, and VS values with mean values of 0.739 ± 0.073, 1.739 ± 0.398, and 0.892 ± 0.044 respectively ($p < 0.001$ for DSC and 95HD and $p = 0.025$ for VS) (Fig. 6). Although the WAIR + T1 combination showed higher values (with the DSC = 0.747 ± 0.073, 95HD = 1.712 ± 0.517, and VS = 0.901 ± 0.039), the increase was not significant ($p > 0.05$) compared to WAIR only. There was no significant difference between WAIR and combinations of WAIR with DTI ($p > 0.05$).

Reoriented and non-reoriented DTI vectors. In the case of WAIR + pdvFA, the DSC value improved from 0.723 ± 0.075 to 0.737 ± 0.073 (the 95HD and VS values also improved), but this difference was not statistically significant ($p > 0.05$). Based on the training results, the 95HD value improved statistically from 2.142 ± 0.489 to 1.844 ± 0.280 when using reoriented vectors for the WAIR + ccFA experiment ($p = 0.006$) (see Supplementary Material 1 Table S.6).

DL model comparison by structures. WAIR statistically outperformed T1 based on DSC for the five structures ($p < 0.05$), based on 95HD for STN ($p = 0.004$), MT-fa ($p < 0.001$), SN ($p < 0.001$), and RN ($p < 0.001$), and based on VS for STN ($p = 0.009$). For MB based on DSC values WAIR significantly outperformed WAIR + pdv^{NR} ($p = 0.024$) and

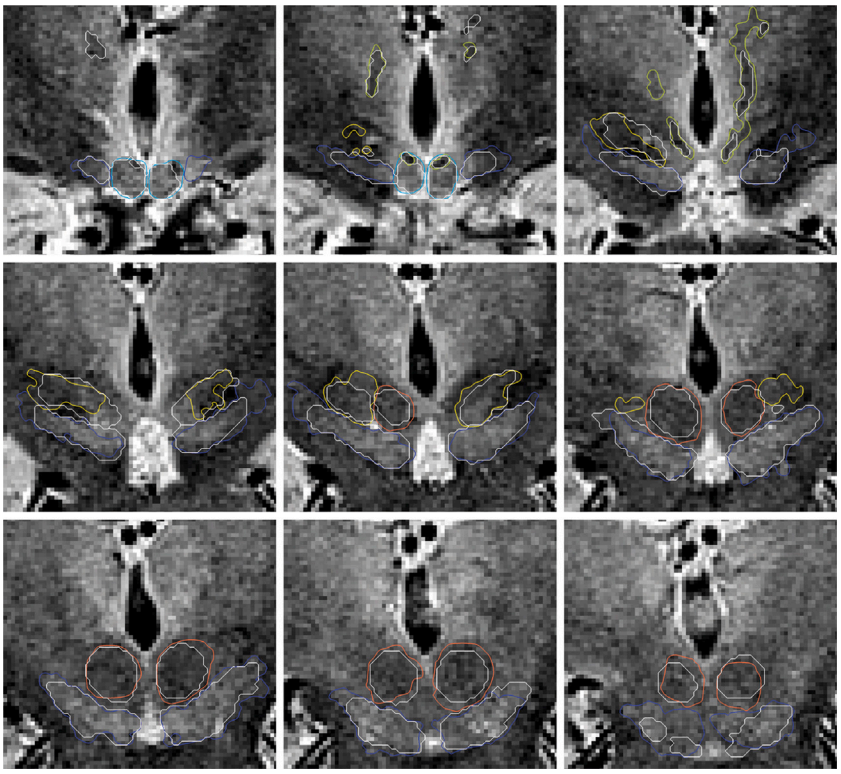

Fig. 6. One patient example of manual (colored line) and automatic (white line), WAIR-based contours: subthalamic nucleus, yellow; substantia nigra, dark blue; mammillo-thalamic fascicle, pale green; red nucleus, red; mammillary body, pale blue (same color code as used with the surgical software, Iplan).

Table 2
Mean Dice scores (DSC), 95% Hausdorff distance (95HD), and volumetric similarity (VS) obtained by experiments with different combinations of WAIR, T1, and DTI derivatives. Results for both with (^{R}) and without (^{NR}) vector reorientation are given. The down-pointing arrow (↓) indicates that the experiment yielded significantly poorer results than the WAIR experiment. The best results are in bold.

Input	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	**0.747 ± 0.073**	**1.712 ± 0.517**	**0.901 ± 0.039**
WAIR + ccFA^{NR}	0.727 ± 0.078	1.872 ± 0.515	0.884 ± 0.045
WAIR + ccFA^{R}	0.730 ± 0.078	1.807 ± 0.465	0.886 ± 0.046
WAIR + pdv^{NR}	0.728 ± 0.071	1.864 ± 0.480	0.884 ± 0.041
WAIR + pdv^{R}	0.728 ± 0.071	1.821 ± 0.423	0.886 ± 0.044
WAIR + pdvFA^{NR}	0.723 ± 0.075	1.891 ± 0.464	0.885 ± 0.043
WAIR + pdvFA^{R}	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

WAIR + pdv^{R} ($p = 0.028$). The WAIR + T1 combination significantly outperformed the WAIR experiment for the MB based on the 95HD value ($p = 0.046$).

The comparisons by hemisphere did not add to the results (Supplementary Material 4). The sex ratio of the training set did not affect the performance of the DL network applied to the test set. The results of the best-obtained DL model (trained nnU-Net) were comparable to the test results (see Supplementary Material 1 Tables S.5, S.6, S.7).

4. Discussion

In this work, we used a ground truth clinical dataset comprising the manual contourings of the SN, RN, STN, MB, and MT-fa, carried out on MRI images of a WAIR sequence dedicated to DB architecture analysis. WAIR-based automatic contouring yielded a good result, with a mean DSC value of 0.739 ± 0.073. The MB, RN, SN, STN, and MT-fa, DSC values decreased in that order, reflecting the increasing difficulty met by a medical expert identifying these structures, at least for the nuclei. The sex of the patients did not seem to affect the performance of the DL network.

The mean DSC value of 0.74 can be interpreted as substantial or excellent given the overall agreement derived from Kappa statistics (Watson and Petrie, 2010; Zijdenbos et al., 1994), suggesting that our trained self-configured 3D nnU-Net could be applied clinically with reasonable computation time. Higher mean DSC values above 0.85 were reported for large well-known structures, easily identified in T1, such as the caudate, the putamen, and the thalamus, Dolz et al. (2018), Kushibar et al. (2018) and Rashed et al. (2020). Similar results were also obtained by using T1 and scalar maps of DTI (Pinheiro et al., 2020). Less well-known structures, such as the nucleus accumbens (Kushibar et al., 2018; Rashed et al., 2020, thalamic nuclei Majdi et al., 2020, limbic structures Baniasadi et al., 2023; Greve et al., 2021 and miscellaneous structures including internal capsule, habenula, ventricles Baniasadi et al., 2023 were analyzed, and showed heterogeneity of mean DSC values. The diversity of experimental protocols could explain the wide range of DSC values, such as the use of mixed case series (healthy subjects, degenerative disease, autism) (Baniasadi et al., 2023; Dolz et al., 2018; Greve et al., 2021; Kushibar et al., 2018; Majdi et al., 2020; Rashed et al., 2020), the different ground truth segmentations, manual with atlas (aid or registration) (Baniasadi et al., 2023; Dolz et al., 2018; Greve et al., 2021; Kushibar et al., 2018; Majdi et al., 2020; Pinheiro et al., 2020; Manjón et al., 2020), semi-automatic with class tissue and "golden" (Rashed et al., 2020), manual without atlas (Baxter and Jannin, 2022; Chen et al., 2023). The interpretation of DSC value should also account for the specificities of structures. Thus, in our work, the nnU-Net performed differently according to structures (Fig. 6, Table 3), such as the MB, with the highest DSC value, a simple balloon shape highly contrasted from the environment, and the MT-fa, with the lowest DSC value, an elongated shape not continuous between adjacent slices, although highly contrasted from the environment (the thalamus). Even so, the potential of DL in segmenting deep brain structures from

Table 3
Mean DSC and standard deviation values for the five segmented structures (left and right hemispheres) for several experiments. Superscript R means that the vector gradients were reoriented. The down-pointing arrow (↓) indicates that the experiment yielded significantly poorer results for each label on each hemisphere than the WAIR experiment. The best results are in bold.

Label	Hem.	WAIR	T1	WAIR + T1	WAIR + pdvFA^R	WAIR + FA	WAIR + T1 + FA
SN	L	0.793 ± 0.048	0.731 ± 0.060 ↓	**0.796 ± 0.044**	0.787 ± 0.045	0.782 ± 0.048	0.782 ± 0.042
	R	0.789 ± 0.046	0.701 ± 0.155 ↓	**0.791 ± 0.051**	0.776 ± 0.048	0.774 ± 0.049	0.777 ± 0.048
RN	L	0.756 ± 0.067	0.709 ± 0.122 ↓	**0.772 ± 0.055**	0.761 ± 0.059	0.757 ± 0.066	0.760 ± 0.062
	R	0.778 ± 0.059	0.732 ± 0.093 ↓	**0.784 ± 0.046**	0.778 ± 0.063	0.783 ± 0.064	0.776 ± 0.054
STN	L	0.698 ± 0.068	0.625 ± 0.168 ↓	**0.717 ± 0.049**	0.715 ± 0.049	0.711 ± 0.066	0.698 ± 0.072
	R	0.715 ± 0.068	0.661 ± 0.103 ↓	**0.722 ± 0.071**	0.719 ± 0.079	0.713 ± 0.076	0.693 ± 0.089
MB	L	0.807 ± 0.069	0.786 ± 0.091	0.816 ± 0.072	0.801 ± 0.068	0.810 ± 0.066	**0.821 ± 0.060**
	R	0.827 ± 0.083	0.792 ± 0.108 ↓	**0.832 ± 0.081**	0.819 ± 0.094	0.831 ± 0.089	0.827 ± 0.100
MT-fa	L	0.636 ± 0.138	0.476 ± 0.156 ↓	0.640 ± 0.135	0.644 ± 0.125	**0.647 ± 0.129**	0.634 ± 0.141
	R	0.593 ± 0.109	0.486 ± 0.139 ↓	**0.600 ± 0.113**	0.575 ± 0.128	0.587 ± 0.105	0.593 ± 0.107
Average		0.739 ± 0.073	0.670 ± 0.106 ↓	**0.747 ± 0.073**	0.737 ± 0.073	0.740 ± 0.072	0.736 ± 0.075

MRI images is confirmed.

The automatic contouring according to T1 alone performed less well (mean DSC value, 0.670 ± 0.106) than according to WAIR, in line with the lower contrast details observed in the subthalamus on T1 (Fig. 2). Other teams have proposed non-T1 MRI sequences to facilitate the identification of DB structures, notably the subthalamic nucleus (Lefranc et al., 2014; Li et al., 2016; O'Gorman et al., 2011; Oros-Peusquens et al., 2019). Inversion-recovery sequences, like the WAIR, would be good candidates, alone or in combination, to segment the DB (Reinacher et al., 2019).

We also observed that even though similarity indices were higher for WAIR + T1 and WAIR + FA image datasets, they did not reach statistical significance. The proposed pdvFA derivative of the DTI performed better than the ccFA, showing that the use of the absolute value of the vector was not optimal for image segmentation tasks. As regards the reorientation, the result of the WAIR + ccFA and WAIR + pdvFA experiments improved after applying the reorientation to the vectors (not statistically significant). The statistical difference in 95HD values during training indicates that there are differences between reoriented and non-reoriented experiments in terms of boundary localization.

Results obtained using combinations with images derived from DTI showed lower DSC, 95HD, and VS values compared to the WAIR + T1 combination. This could be explained by the specific features of DTI voxels. The discrepancy in voxel size could have caused a significant information loss during registration. The complexity of DTI data, pdvFA, pdv, and ccFA were saved as 4D images with three volumes, where each volume corresponded to the component of the vector or to the component of the RGB color (ccFA). DTI informed primarily on water movements within the axons (Le Bihan, 2003), whereas T1 and WAIR provided more detailed structural information on brain tissue. Finally, the concatenation of input volumes used in this work is considered an early fusion method, and it assumes a linear (simple) relationship between modalities (Dolz et al., 2019), which was not the case in our work. Accordingly, other ways of merging the WAIR and DTI information will be considered in our future work. For instance, the layer-level (intermediate) fusion methods, where the independently extracted features from different modalities are fused at an intermediate layer of a neural network and can thus exploit more complex features of different modalities (Zhou et al., 2019; Nie et al., 2016).

The progressive inclusion of patients resulted in limitations. The first being the small number of recruited patients (n = 53), although the largest reported concerning a clinical application. In addition, the manual contouring, time-consuming, necessitates advanced knowledge and informative structural MRI images, which still impede a fast and wide development. The ongoing progress in artificial intelligence, fed and controlled by researchers and clinicians, should speed up improvements.

In summary, deep learning-based DB mapping could be clinically relevant, alongside other automatic segmentations of subthalamic structures, such as mixed methods with automatic tissue segmentation on individual MRI sets coupled with atlas-based contouring (Polanski et al., 2020; Reinacher et al., 2019) or template-based methods (Haegelen et al., 2013). Adding anatomical priors to the learning process could help the network to segment complex structures, particularly in the presence of a limited dataset, for instance by using a brain atlas to guide the white matter segmentation (Zhang et al., 2021) or using atlas priors as additional inputs to the network (Coupé et al., 2020; Kushibar et al., 2018).

5. Conclusion

In this work, we segmented deep brain structures according to different MRI modalities. Our contribution includes a multi-modal dataset and manual segmentation of five small deep brain structures by a clinical expert. To the best of our knowledge, this is the first time that a ground truth data set was built a priori prospectively in a truly relevant clinical context, independently of DL training, and with an advanced analysis taking into account individual variability (mixed model). Inversion-recovery sequences, like the WAIR, seem a good first choice in the field, providing enough contrast to support precise manual contouring of DB structures, and now automatic contouring, as shown with our trained self-configured 3D nnU-Net. The use of multi-modality training yielded encouraging results (albeit not statistically significant), the best DSC, 95HD, and VS values being obtained using the WAIR + T1 combination. Moreover, we showed that the correct reorientation of DTI vectors (with respect to the ACPC line) improved the segmentation process. Finally, our results suggest that clinical trials could start as soon as secured pipelines meeting ethical criteria are implemented.

CRediT authorship contribution statement

Aigerim Dautkulova: Writing – original draft, Visualization, Software, Methodology, Investigation. **Omar Ait Aider:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Céline Teulière:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Conceptualization. **Jérôme Coste:** Resources, Investigation, Data curation. **Rémi Chaix:** Resources, Investigation, Data curation. **Omar Ouachik:** Resources, Investigation, Data curation. **Bruno Pereira:** Writing – original draft, Validation, Methodology, Formal analysis. **Jean-Jacques Lemaire:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Omar Ait Ader reports financial support was provided by European Regional Development Fund. Jean-Jacques Lemaire reports financial support was provided by Foundation de l'Avenir pour la Recherche Médicale Appliquée. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was co-financed by the European Regional Development Fund (FEDER; UNION EUROPEENNE; AIM project) with the French National Research Agency (French ANR-IA Ph.D. student grant), and the Fondation de l'Avenir (Ptolémée program, 21_0099_01; Ph.D. student funding). Computations were performed in the supercomputer facilities of the Mésocentre Clermont-Auvergne at the Université Clermont Auvergne.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.compmedimag.2024.102488.

Data availability

Data will be made available on request.

References

Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., Ronneberger, O., Summers, R.M., van Ginneken, B., Bilello, M., Bilic, P., Christ, P.F., Do, R.K.G., Gollub, M.J., Heckers, S.H., Huisman, H., Jarnagin, W.R., McHugo, M.K., Napel, S., Pernicka, J.S.G., Rhode, K., Tobon-Gomez, C., Vorontsov, E., Meakin, J.A., Ourselin, S., Wiesenfarth, M., Arbeláez, P., Bae, B., Chen, S., Daza, L., Feng, J., He, B., Isensee, F., Ji, Y., Jia, F., Kim, I., Maier-Hein, K., Merhof, D., Pai, A., Park, B., Perslev, M., Rezaiifar, R., Rippel, O., Sarasua, I., Shen, W., Son, J., Wachinger, C., Wang, L., Wang, Y., Xia, Y., Xu, D., Xu, Z., Zheng, Y., Simpson, A.L., Maier-Hein, L., Cardoso, M.J., 2022. The medical segmentation decathlon. Nature Commun. 13 (1), http://dx.doi.org/10.1038/s41467-022-30695-9.

Avesta, A., Hossain, S., Lin, M., Aboian, M., Krumholz, H.M., Aneja, S., 2023. Comparing 3D, 2.5D, and 2D approaches to brain image auto-segmentation. Bioengineering 10 (2), http://dx.doi.org/10.3390/bioengineering10020181, URL: https://www.mdpi.com/2306-5354/10/2/181.

Baniasadi, M., Petersen, M.V., Gonçalves, J., Horn, A., Vlasov, V., Hertel, F., Husch, A., 2023. DBSegment: Fast and robust segmentation of deep brain structures considering domain generalization. Hum. Brain Mapp. 44 (2), 762–778. http://dx.doi.org/10.1002/hbm.26097, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.26097.

Baxter, J.S.H., Jannin, P., 2022. Combining simple interactivity and machine learning: a separable deep learning approach to subthalamic nucleus localization and segmentation in MRI for deep brain stimulation surgical planning. J. Med. Imaging 9 (4), 045001. http://dx.doi.org/10.1117/1.JMI.9.4.045001.

Bohnen, N.I., Yarnall, A.J., Weil, R.S., Moro, E., Moehle, M.S., Borghammer, P., Bedard, M.-A., Albin, R.L., 2022. Cholinergic system changes in Parkinson's disease: emerging therapeutic approaches. Lancet Neurol. 21 (4), 381–392. http://dx.doi.org/10.1016/S1474-4422(21)00377-X, URL: https://www.sciencedirect.com/science/article/pii/S147444222100377X.

Cai, S., Tian, Y., Lui, H., Zeng, H., Wu, Y., Chen, G., 2020. Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network. Quant. Imaging Med. Surg. 10 (6), URL: https://qims.amegroups.org/article/view/43519.

Chen, J., Xu, H., Xu, B., Wang, Y., Shi, Y., Xiao, L., 2023. Automatic localization of key structures for subthalamic nucleus–deep brain stimulation surgery via prior-enhanced multi-object magnetic resonance imaging segmentation. World Neurosurg. 178, e472–e479. http://dx.doi.org/10.1016/j.wneu.2023.07.103, URL: https://www.sciencedirect.com/science/article/pii/S1878875023010410.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Springer International Publishing, Cham, pp. 424–432.

Coupé, P., Mansencal, B., Clément, M., Giraud, R., Denis de Senneville, B., Ta, V.-T., Lepetit, V., Manjon, J.V., 2020. AssemblyNet: A large ensemble of CNNs for 3D whole brain MRI segmentation. Neuroimage 219, 117026. http://dx.doi.org/10.1016/j.neuroimage.2020.117026, URL: https://www.sciencedirect.com/science/article/pii/S1053811920305127.

Dejerine, J., 1901. Anatomie Des Centres Nerveux, vol. 1, Rueff (Paris).

Dice, L.R., 1945. Measures of the amount of ecologic association between species. Ecology 26 (3), 297–302, URL: http://www.jstor.org/stable/1932409.

Dolz, J., Desrosiers, C., Ben Ayed, I., 2018. 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. Neuroimage 170, 456–470. http://dx.doi.org/10.1016/j.neuroimage.2017.04.039, URL: https://www.sciencedirect.com/science/article/pii/S1053811917303324, Segmenting the Brain.

Dolz, J., Gopinath, K., Yuan, J., Lombaert, H., Desrosiers, C., Ben Ayed, I., 2019. HyperDense-Net: A hyper-densely connected CNN for multi-modal image segmentation. IEEE Trans. Med. Imaging 38 (5), 1116–1126. http://dx.doi.org/10.1109/TMI.2018.2878669.

El Ouadih, Y., Marques, A., Pereira, B., Luisoni, M., Claise, B., Coste, J., Sontheimer, A., Chaix, R., Debilly, B., Derost, P., Morand, D., Durif, F., Lemaire, J.-J., 2023. 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 Neurochir. 165 (12), 3927–3941. http://dx.doi.org/10.1007/s00701-023-05843-9.

Fox, S.H., Katzenschlager, R., Lim, S.-Y., Barton, B., de Bie, R.M.A., Seppi, K., Coelho, M., Sampaio, C., on behalf of the Movement Disorder Society Evidence-Based Medicine Committee, 2018. International Parkinson and movement disorder society evidence-based medicine review: Update on treatments for the motor symptoms of Parkinson's disease. Mov. Disorders 33 (8), 1248–1266. http://dx.doi.org/10.1002/mds.27372, URL: https://movementdisorders.onlinelibrary.wiley.com/doi/abs/10.1002/mds.27372.

Greve, D.N., Billot, B., Cordero, D., Hoopes, A., Hoffmann, M., Dalca, A.V., Fischl, B., Iglesias, J.E., Augustinack, J.C., 2021. A deep learning toolbox for automatic segmentation of subcortical limbic structures from MRI images. Neuroimage 244, 118610. http://dx.doi.org/10.1016/j.neuroimage.2021.118610, URL: https://www.sciencedirect.com/science/article/pii/S1053811921008831.

Gulban, O.F., Nielson, D., john lee, Poldrack, R., Gorgolewski, C., Vanessasaurus, Markiewicz, C., 2022. Poldracklab/pydeface: PyDeface v2.0.2. http://dx.doi.org/10.5281/zenodo.6856482.

Haber, S.N., Adler, A., Bergman, H., 2012. Chapter 20 - The basal ganglia. In: Mai, J.K., Paxinos, G. (Eds.), The Human Nervous System (Third Edition), third ed. Academic Press, San Diego, pp. 678–738. http://dx.doi.org/10.1016/B978-0-12-374236-0.10020-3, URL: https://www.sciencedirect.com/science/article/pii/B9780123742360100203.

Haegelen, C., Coupé, P., Fonov, V., Guizard, N., Jannin, P., Morandi, X., Collins, D.L., 2013. Automated segmentation of basal ganglia and deep brain structures in MRI of Parkinson's disease. Int. J. Comput. Assist. Radiol. Surg. 8 (1), 99–110. http://dx.doi.org/10.1007/s11548-012-0675-8.

Hannah, R., Aron, A.R., 2021. Towards real-world generalizability of a circuit for action-stopping. Nature Rev. Neurosci. 22 (9), 538–552. http://dx.doi.org/10.1038/s41583-021-00485-1.

Harmsen, I.E., Elias, G.J., Beyn, M.E., Boutet, A., Pancholi, A., Germann, J., Mansouri, A., Lozano, C.S., Lozano, A.M., 2020. Clinical trials for deep brain stimulation: Current state of affairs. Brain Stimul. 13 (2), 378–385. http://dx.doi.org/10.1016/j.brs.2019.11.008, URL: https://www.sciencedirect.com/science/article/pii/S1935861X19304668.

Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H., 2020. NnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. Nature Methods 18 (2), 203–211. http://dx.doi.org/10.1038/s41592-020-01008-z.

Kushibar, K., Valverde, S., González-Villà, S., Bernal, J., Cabezas, M., Oliver, A., Lladó, X., 2018. Automated sub-cortical brain structure segmentation combining spatial and deep convolutional features. Med. Image Anal. 48, 177–186. http://dx.doi.org/10.1016/j.media.2018.06.006, URL: https://www.sciencedirect.com/science/article/pii/S1361841518303839.

Le Bihan, D., 2003. Looking into the functional architecture of the brain with diffusion MRI. Nature Rev. Neurosci. 4 (6), 469–480. http://dx.doi.org/10.1038/nrn1119.

Lefranc, M., Derrey, S., Merle, P., Tir, M., Constans, J.-M., Montpellier, D., Macron, J.M., Le Gars, D., Peltier, J., Baledentt, O., Krystkowiak, P., 2014. 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), 615–626, discussion 627.

Lemaire, J.-J., De Salles, A., Coll, G., El Ouadih, Y., Chaix, R., Coste, J., Durif, F., Makris, N., Kikinis, R., 2019. MRI atlas of the human deep brain. Front. Neurol. 10, http://dx.doi.org/10.3389/fneur.2019.00851, URL: https://www.frontiersin.org/articles/10.3389/fneur.2019.00851.

Lemaire, J.-J., Sontheimer, A., Pereira, B., Coste, J., Rosenberg, S., Sarret, C., Coll, G., Gabrillargues, J., Jean, B., Gillart, T., Coste, A., Roche, B., Kelly, A., Pontier, B., Feschet, F., 2018. Deep brain stimulation in five patients with severe disorders of consciousness. Ann. Clin. Transl. Neurol. 5 (11), 1372–1384. http://dx.doi.org/10.1002/acn3.648.

Li, B., Jiang, C., Li, L., Zhang, J., Meng, D., 2016. Automated segmentation and reconstruction of the subthalamic nucleus in Parkinson's disease patients. Neuromodulation 19 (1), 13–19. http://dx.doi.org/10.1111/ner.12350, URL: https://www.sciencedirect.com/science/article/pii/S1094715921048741.

Mai, J., Majtanik, M., Paxinos, G., 2016. Atlas of the Human Brain, fourth ed. Academic Press, Elsevier.

Majdi, M.S., Keerthivasan, M.B., Rutt, B.K., Zahr, N.M., Rodriguez, J.J., Saranathan, M., 2020. Automated thalamic nuclei segmentation using multi-planar cascaded convolutional neural networks. Magn. Reson. Imaging 73, 45–54. http://dx.doi.org/10.1016/j.mri.2020.08.005, URL: https://www.sciencedirect.com/science/article/pii/S0730725X20303118.

Manjón, J.V., Bertó, A., Romero, J.E., Lanuza, E., Vivo-Hernando, R., Aparici-Robles, F., Coupe, P., 2020. pBrain: A novel pipeline for Parkinson related brain structure segmentation. NeuroImage: Clin. 25, 102184. http://dx.doi.org/10.1016/j.nicl.2020.102184, URL: https://www.sciencedirect.com/science/article/pii/S221315822030022X.

McCormick, M.M., Liu, X., Ibanez, L., Jomier, J., Marion, C., 2014. ITK: enabling reproducible research and open science. Front. Neuroinform. 8, http://dx.doi.org/10.3389/fninf.2014.00013, URL: https://www.frontiersin.org/journals/neuroinformatics/articles/10.3389/fninf.2014.00013.

Milletari, F., Navab, N., Ahmadi, S.-A., 2016. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV). pp. 565–571. http://dx.doi.org/10.1109/3DV.2016.79.

Nie, D., Wang, L., Gao, Y., Shen, D., 2016. Fully convolutional networks for multi-modality isointense infant brain image segmentation. In: 2016 IEEE 13th International Symposium on Biomedical Imaging. ISBI, pp. 1342–1345. http://dx.doi.org/10.1109/ISBI.2016.7493515.

Nieuwenhuys, R., Voogd, J., Van Huijzen, C., 2008. The human central nervous system, 4th ed.. Am. J. Neuroradiol. 29 (5), http://dx.doi.org/10.3174/ajnr.A0991, e39–e39.

O'Donnell, L.J., Westin, C.-F., 2011. An introduction to diffusion tensor image analysis. Neurosurg. Clin. North Am. 22 (2), 185–196. http://dx.doi.org/10.1016/j.nec.2010.12.004, URL: https://www.sciencedirect.com/science/article/pii/S104236801000121X, Functional Imaging.

O'Gorman, R.L., Shmueli, K., Ashkan, K., Samuel, M., Lythgoe, D.J., Shahidiani, A., Wastling, S.J., Footman, M., Selway, R.P., Jarosz, J., 2011. Optimal MRI methods for direct stereotactic targeting of the subthalamic nucleus and globus pallidus. Eur. Radiol. 21 (1), 130–136. http://dx.doi.org/10.1007/s00330-010-1885-5.

Oros-Peusquens, A.-M., Loução, R., Abbas, Z., Gras, V., Zimmermann, M., Shah, N.J., 2019. A single-scan, rapid whole-brain protocol for quantitative water content mapping with neurobiological implications. Front. Neurol. 10, http://dx.doi.org/10.3389/fneur.2019.01333, URL: https://www.frontiersin.org/articles/10.3389/fneur.2019.01333.

Pajevic, S., Pierpaoli, C., 1999. Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: application to white matter fiber tract mapping in the human brain. Magn. Reson. Med. 43, 526, URL: https://api.semanticscholar.org/CorpusID:9288100.

Pieper, S.D., Halle, M.W., Kikinis, R., 2012. 3D slicer. In: 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821). Vol. 1, pp. 632–635.

Pinheiro, G.R., Carmo, D.S., Yasuda, C., Lotufo, R.A., Rittner, L., 2020. Convolutional neural network on DTI data for sub-cortical brain structure segmentation. In: Computational Diffusion MRI. Springer International Publishing, Cham, pp. 135–146.

Polanski, W.H., Zolal, A., Sitoci-Ficici, K.H., Hiepe, P., Schackert, G., Sobottka, S.B., 2020. Comparison of automatic segmentation algorithms for the subthalamic nucleus. Stereotact. Funct. Neurosurg. 98 (4), 256–262. http://dx.doi.org/10.1159/000507028.

Rashed, E.A., Gomez-Tames, J., Hirata, A., 2020. End-to-end semantic segmentation of personalized deep brain structures for non-invasive brain stimulation. Neural Netw. 125, 233–244. http://dx.doi.org/10.1016/j.neunet.2020.02.006, URL: https://www.sciencedirect.com/science/article/pii/S0893608020300514.

Reinacher, P., Várkuti, B., Krüger, M., Piroth, T., Egger, K., Roelz, R., Coenen, V., 2019. Automatic segmentation of the subthalamic nucleus: A viable option to support planning and visualization of patient-specific targeting in deep brain stimulation. Oper. Neurosurg. 17, http://dx.doi.org/10.1093/ons/opz015.

Riley, H.A., 1943. An Atlas of the Basal Ganglia, Brain Stem and Spinal Cord: Based on Myelin-Stained Material. Williams & Wilkins.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp. 234–241.

Schaltenbrand, G., Bailey, P., 1959. Einführung in die stereotaktischen operationen: mit einem atlas des menschlichen Gehirns. Introduction to stereotaxis, with an atlas of the human brain. Einführung in Die Stereotaktischen Operationen, G. Thieme, Number vol.1.

Seo, H., Huang, C., Bassenne, M., Xiao, R., Xing, L., 2020. Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images. IEEE Trans. Med. Imaging 39 (5), 1316–1325. http://dx.doi.org/10.1109/TMI.2019.2948320.

Shivdeo, A., Lokwani, R., Kulkarni, V., Kharat, A., Pant, A., 2021. Evaluation of 3D and 2D deep learning techniques for semantic segmentation in CT scans. In: 2021 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems. icABCD, pp. 1–8, URL: https://api.semanticscholar.org/CorpusID:231639325.

Taha, A.A., Hanbury, A., 2015. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med. Imaging 15 (1), 29. http://dx.doi.org/10.1186/s12880-015-0068-x.

Vassal, F., Coste, J., Derost, P., Mendes, V., Gabrillargues, J., Nuti, C., Durif, F., Lemaire, J.-J., 2012. 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 Stimul. 5 (4), 625–633. http://dx.doi.org/10.1016/j.brs.2011.10.007, URL: https://www.sciencedirect.com/science/article/pii/S1935861X1100163X.

Watson, P., Petrie, A., 2010. Method agreement analysis: A review of correct methodology. Theriogenology 73 (9), 1167–1179. http://dx.doi.org/10.1016/j.theriogenology.2010.01.003, URL: https://www.sciencedirect.com/science/article/pii/S0093691X10000233.

Zerroug, A., Gabrillargues, J., Coll, G., Vassal, F., Jean, B., Chabert, E., Claise, B., Khalil, T., Sakka, L., Feschet, F., Durif, F., Boyer, L., Coste, J., Lemaire, J.-J., 2016. 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. http://dx.doi.org/10.1016/j.neuchi.2016.01.009, URL: https://www.sciencedirect.com/science/article/pii/S0028377016300194.

Zhang, H., Avants, B.B., Yushkevich, P.A., Woo, J.H., Wang, S., McCluskey, L.F., Elman, L.B., Melhem, E.R., Gee, J.C., 2007. High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: An example study using amyotrophic lateral sclerosis. IEEE Trans. Med. Imaging 26 (11), 1585–1597. http://dx.doi.org/10.1109/TMI.2007.906784.

Zhang, S., Laidlaw, D.H., Kindlmann, G., 2005. 16 - Diffusion tensor MRI visualization. In: Hansen, C.D., Johnson, C.R. (Eds.), Visualization Handbook. Butterworth-Heinemann, Burlington, pp. 327–340. http://dx.doi.org/10.1016/B978-012387582-2/50018-6, URL: https://www.sciencedirect.com/science/article/pii/B9780123875822500186.

Zhang, F., Noh, T., Juvekar, P., Frisken, S.F., Rigolo, L., Norton, I., Kapur, T., Pujol, S., Wells, W., Yarmarkovich, A., Kindlmann, G., Wassermann, D., San Jose Estepar, R., Rathi, Y., Kikinis, R., Johnson, H.J., Westin, C.-F., Pieper, S., Golby, A.J., O'Donnell, L.J., 2020. SlicerDMRI: Diffusion MRI and tractography research software for brain cancer surgery planning and visualization. JCO Clin. Cancer Inform. (4), 299–309. http://dx.doi.org/10.1200/CCI.19.00141, arXiv:https://doi.org/10.1200/CCI.19.00141, PMID: 32216636.

Zhang, Z., Powell, K., Yin, C., Cao, S., Gonzalez, D., Hannawi, Y., Zhang, P., 2021. Brain atlas guided attention U-Net for white matter hyperintensity segmentation. AMIA Jt. Summits Transl. Sci. Proc. 2021, 663–671.

Zhou, T., Ruan, S., Canu, S., 2019. A review: Deep learning for medical image segmentation using multi-modality fusion. Array 3–4, 100004. http://dx.doi.org/10.1016/j.array.2019.100004, URL: https://www.sciencedirect.com/science/article/pii/S2590005619300049.

Zijdenbos, A., Dawant, B., Margolin, R., Palmer, A., 1994. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans. Med. Imaging 13 (4), 716–724. http://dx.doi.org/10.1109/42.363096.