Utilizing the TractSeg Tool for Automatic Corticospinal Tract Segmentation in Patients With Brain Pathology

Yael H. Moshe, BS[1,2], **Dafna Ben Bashat, PhD**[1,3], **Zeev Hananis, BS**[1,4], **Mina Teicher, PhD**[2,4,*], and **Moran Artzi, PhD**[1,3,*] (iD)

Abstract

Purpose: White-matter tract segmentation in patients with brain pathology can guide surgical planning and can be used for tissue integrity assessment. Recently, TractSeg was proposed for automatic tract segmentation in healthy subjects. The aim of this study was to assess the use of TractSeg for corticospinal-tract (CST) segmentation in a large cohort of patients with brain pathology and to evaluate its consistency in repeated measurements. **Methods:** A total of 649 diffusion-tensor-imaging scans were included, of them: 625 patients and 24 scans from 12 healthy controls (scanned twice for consistency assessment). Manual CST labeling was performed in all cases, and by 2 raters for the healthy subjects. Segmentation results were evaluated based on the Dice score. In order to evaluate consistency in repeated measurements, volume, Fractional Anisotropy (FA), and Mean Diffusivity (MD) values were extracted and correlated for the manual versus automatic methods. **Results:** For the automatic CST segmentation Dice scores of 0.63 and 0.64 for the training and testing datasets were obtained. Higher consistency between measurements was detected for the automatic segmentation, with between measurements correlations of volume = 0.92/0.65, MD = 0.94/0.75 for the automatic versus manual segmentation. **Conclusions:** The TractSeg method enables automatic CST segmentation in patients with brain pathology. Superior measurements consistency was detected for the automatic in comparison to manual fiber segmentation, which indicates an advantage when using this method for clinical and longitudinal studies.

Keywords

DTI, deep learning, segmentation, corticospinal tract, MRI, consistency measurements

Received: July 15, 2022; Revised: August 24, 2022; Accepted: September 7, 2022.

Introduction

Preoperative white matter (WM) fibers visualization and quantitative assessment of tissue organization in patients with brain pathology can assist in the selection of surgical approach and setting maximum boundaries for the extent of lesion resection while preserving function.[1–3]

Magnetic resonance imaging (MRI) is the method of choice for the assessment and monitoring of patients with brain lesions. MRI enables superior soft tissue contrast over other imaging modalities, allowing for better visualization of subtle tumor cell infiltration or disrupted parenchymal architecture, and enables extensive characterization of tissue structure and function.[4] In addition to structural imaging, diffusion tensor imaging (DTI) can be used for WM fibers visualization and quantitative assessment.[5] Using DTI tissue diffusivity can be detected by employing magnetic field

gradients in different directions (≥ 6) which are sensitized to diffusion in a particular direction. Next, voxel-based diffusion tensor can be calculated from the acquired images by solving the Stejskal-Tanner equation, allowing for extract

[1] Sagol Brain Institute, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel
[2] Department of Mathematics, Bar Ilan University, Ramat Gan, Israel
[3] Sackler Faculty of Medicine and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel
[4] Gonda Brain Research Center, Bar-Ilan University, Ramat Gan, Israel

*Equal contribution.

Corresponding Author:
Artzi Moran, Sagol Brain Institute, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel; Sackler Faculty of Medicine and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel, 6 Weizmann St. Tel-Aviv 64239 Israel.
Email: artzimy@gmail.com

of average diffusivity, fractional anisotropy, and fiber-tracking (Tensor) maps.[6,7]

Currently, reconstruction of the WM tracts based on DTI heavily relies on manual user intervention. The common method for fiber segmentation is streamlines tractography.[8] Streamlines tractography refers to fibers reconstruction by connecting a contiguous set of 3D points in the DTI space.[9] This method requires the setting of seed points (regions of interest —ROI) from which to initiate streamlines, followed by manual/semi-manual corrections of the reconstructed fibers based on prior anatomical knowledge, setting threshold criteria, etc. This dependency in user intervention is a major limitation in integrating DTI segmentation in routine clinical application, as it is time-consuming, and is heavily prone to inter- and intra-rater repeatability errors.[10]

In recent years, several methods have been proposed for automatic fibers segmentation. Poulin et al proposed the use of a Feed-Forward Neural Network, and a Recurrent Neural Network for fibers segmentation, demonstrating high performers on the ISMRM 2015 segmentation challenge,[11] enabling to recover of more than 50% of the spatial coverage.[12] Lam et al[13] proposed the TRAFIC tool, and trained the arcuate fasciculus frontotemporal bundle on the right and left, reaching ~52% and 70% of accuracy, respectively. Gupta et al[14] proposed FiberNET, a robust approach using convolutional neural networks (CNNs) to learn the shape features of the fiber bundles, which was then exploited to cluster WM fibers into bundles. They improved this approach and the proposed workflow aims to remove false positive fibers from a fiber bundle segmented using ROI-based segmentation tools.[15] Wassermann et al[16] proposed a TractQuerier approach that enables to automatically label WM anatomy across subjects by the construction of a dictionary that describes a query language on anatomy definition of the WM tracts. However, the implementations of these methods were on small cohorts and required a long runtime, because they require complex and computationally intensive processing pipelines.

Recently, TractSeg was proposed as a novel, fully automatic approach for direct WM tract segmentation.[17] TractSeg is based on a fully CNN (FCNN) that directly segments WM tracts in fields of fiber orientation distribution function (fODF) peaks. The advantages of this approach are that it provides full and precise segmentations, is simple to set up, short runtime (20.1 min compared to 1704 min of TractQuerier on the same data), and does not require intensive processing pipelines. The algorithm uses a semi-automatic approach to segment 72 anatomically well-described tracts in a cohort of 105 healthy subjects selected from the Human Connectome Project (HCP).[18] Only one recent study has demonstrated the application of the TracSeg model on a small sample of 28 patients with brain tumors.[17]

Although showing promising results for the task of automatic fiber segmentation, the implementation of TractSeg for pre-operative fiber reconstruction in patients with brain lesions is not straightforward. The TractSeg model was trained on healthy subjects and was based on a relatively homogeneous dataset in terms of the study population and MRI protocol. This does not represent the clinical scenario, where there are substantial brain morphological changes due to cerebral pathology, as well as high heterogeneity in terms of data acquisition parameters and data quality. In addition, the TractSeg model was trained based on a permissive definition of fiber space (larger fiber distribution) rather than on a more conservative one which is necessary for the identification of the critical area to preserve.

The aim of this study was to implement and optimize the TractSeg tool in patients with brain pathologies, using an extremely large clinical cohort of 649 DTI scans, focusing on the corticospinal tract (CST) segmentation, and evaluating consistency between measurements, in healthy subjects based on interrater reliability and repeatability assessment.

Materials and Methods

Subjects

A retrospective study was performed. A total of 649 MRI scans from 571 subjects, in our institute, between the years 2015 and 2020 were included. A total of 625 scans were obtained from 559 patients with brain pathologies, who were candidates for brain surgery. Of them, Glioblastoma (n = 254), Oligodendroglioma (n = 112), Astrocytoma (n = 107), Brain Metastasis (n = 31), Ependymoma (n = 2), Cavernoma (n = 10), Epilepsy (n = 30), Anaplastic pleomorphic xanthoastrocytoma (n = 1), Arteriovenous malformation (n = 1), Lymphoma (n = 2), Hemangioma (n = 1), Meningioma (n = 1), and 73 uncertain (patients in whom no biopsy was performed/the biopsy response was uncertain). A total of 24 scans were obtained from 12 healthy controls, scanned twice, with scan intervals < 2 weeks. The healthy subjects were used for between measurements consistency assessment. Inclusion criteria were: age 18 to 100 years, MRI scans including anatomical and diffusion tensor images. Exclusion criteria were: substantial imaging artifacts/low-quality MRI data as detected by visual assessment. The study protocol was approved by the Tel-Aviv Medical Center institutional review board (IRB)—IRB approval numbers 0200-10. No informed consent was required by the IRB for this retrospective study which utilized anonymous data. All procedures were carried out in accordance with relevant guidelines and regulations.

MRI Data

All scans included DTI images and high-resolution anatomical images (3D T1-weighted/3D FLAIR images). Scan parameters of the DTI data are given in Table 1.

Patients: MRI data was collected retrospectively from patients who were scanned for pre-surgical mapping, as part of their routine clinical assessment, performed with different vendors, MRI systems, and various acquisition parameters.

Table 1. MRI Acquisition Parameters.

Vendor	System/magnetic field (Tesla)	Sequence	TE	TR	Voxel dimension (mm³)	# Scans
Siemens	Prisma/3T	EPI	55-71	5300-11 400	6.13-17.58	320
	Skyra/3T	EPI	74-93	5000-8800	7.05-17.58	78
	Aera/1.5T	EPI	72-73	4900-6700	13.71-17.58	117
	Vida/3T	EPI	88	4200	8.86	1
	Avanto-fit/1.5T	EPI	71-75	5800-7600	9-13.71	133

Abbreviations: EPI, Echo Planar Imaging; TE/TR, Repetition Time/Echo Time.
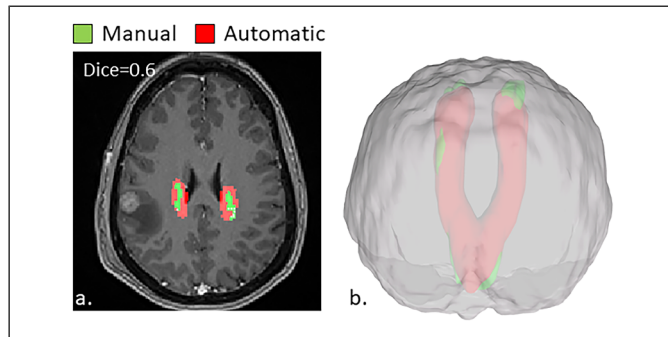
Figure 1. Corticospinal-tract (CST) segmentation results in a patient with a high grade brain tumor, superimposed on a diffusion image ($b = 0$). (a) Axial orientation; (b) 3D representation. Green = manual segmentation; red = automatic segmentation.

Healthy subjects: all scans were performed at the same MRI system and acquisition parameters.

Image Analysis

CST Segmentation. Data Annotation: Manual segmentation of the CST was performed by expert MRI technicians from the preoperative mapping unit, as part of the clinical service for preoperative evaluation in these patients. The CST was reconstructed using DSI Studio (http://dsi-studio.labsolver.org), a tractography software tool for diffusion MRI analysis. The CST was defined as they passed through 3 regions of interest, manually defined on color-coded maps and anatomical images at the level of the brain stem, and motor cortex. For the healthy controls (n = 12, a total of 24 scans), manual segmentation was performed twice by 2 raters (raters 1 and 2).

Image Preprocessing: Preprocessing of the DTI data was performed by FSL software (http://www.fmrib.ox.ac.uk/fsl/) and included skull stripping, eddy currents, and motion corrections. Functional anisotropy (FA) and mean diffusivity (MD) maps were calculated using the FSL FDT tool.[19] Next, in order to make it compatible with network requirements, the tract orientation peaks maps (TOMs) and the manual segmented CST fibers were resliced to $1.25 \times 1.25 \times 1.25$ image size, and were aligned image to MNI space (MNI_FA_template) via the extracted FA map.

Data Splitting: Patients (625 scans): The entire dataset was split in a stratified manner while ensuring that all images

Table 2. Between Measurements Consistency Assessment.

	Volume	FA	MD
Interrater reliability (manual method)	0.65	0.93	0.75
repeatability (automatic method)	0.92	0.92	0.94

Abbreviations: FA, fractional anisotropy; MD, mean diffusivity.

belonging to a given subject would be allocated to the same group. Data was split into 80% training and 20% testing datasets. The training dataset was further split into 80% training and 20% validation. *Healthy controls (24 scans):* The dataset of 12 subjects, scanned twice, was used only for between measurements consistency assessment.

TractSeg model training and evaluation were performed in a Pycharm environment (2019.3.2), using Python software (Python 3.7). The network was trained and tested on a single graphical processing unit (GPU), CUDA device, Nvidia RTX2080 Ti.

The TractSeg model was trained based on the manual segmentation of the CST mask and evaluated on the validation and testing dataset.

Evaluation of segmentation results: comparison between the manual and automatic segmentation results was performed on the patient's dataset based on the Dice score, the common measurement to calculate the overlap between segmentations[20]:

$$\text{DICE} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}$$

where TP = true positives, FP = false positives, and FN = false negatives that are coming between the manual and the automatic results.

TractSeg Performance in Patients With Lesions Adjacent to the CST: In order to assess TractSeg performance in challenging cases, we evaluated the model performance in cases in which lesions adjacent to the CST. For this aim, we realigned the MNI152 standard-space T1-weighted average structural image and the JHU DTI-based white-matter atlases[21,22] to the subject anatomy image space (3D T1-weighted or FLAIR image). Patients were annotated as having lesion adjacent to the CST based on visual inspection of the intersection between the lesion area detected based on the patient's anatomy image and the CST template detected based on the JHU atlas.
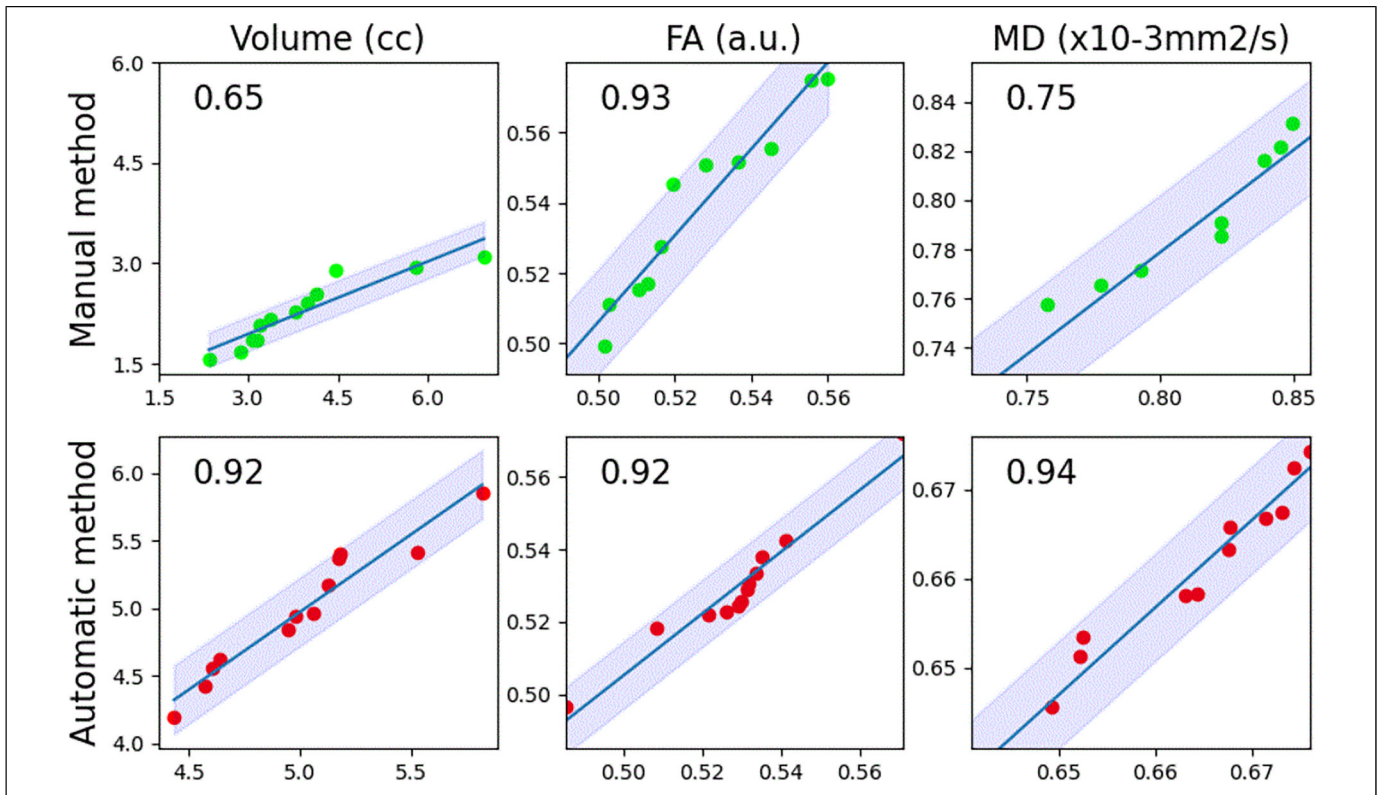
Figure 2. Consistency in repeated measurements. Correlations of volume, fractional anisotropy (FA), and mean diffusivity (MD) between the manual (top row, green dots) and the automatic (lower row, red dots) methods.
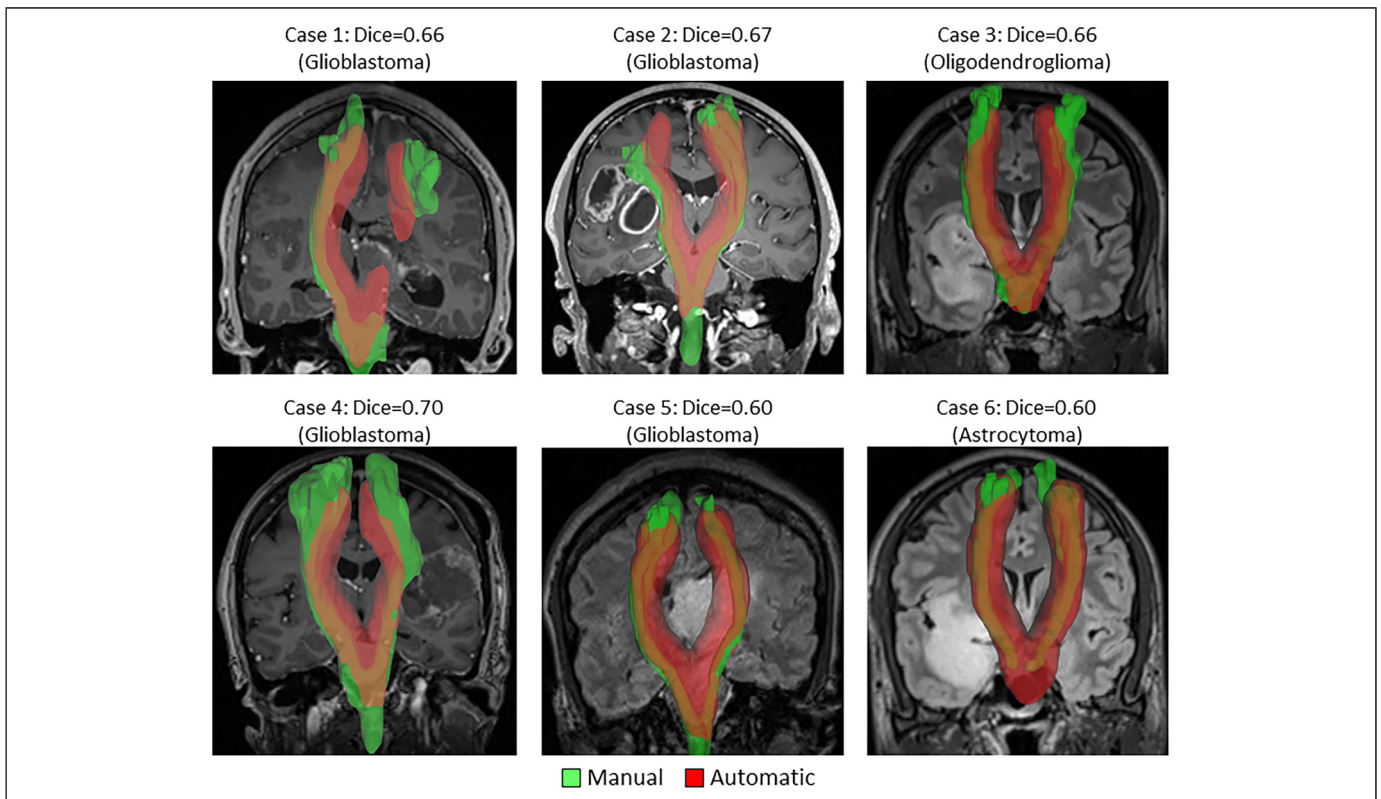
Figure 3. Corticospinal-tract (CST) segmentation results in 6 patients with lesions adjacent to the CST, showing the high similarity between the manual (green) and the automatic (red) segmentation results.
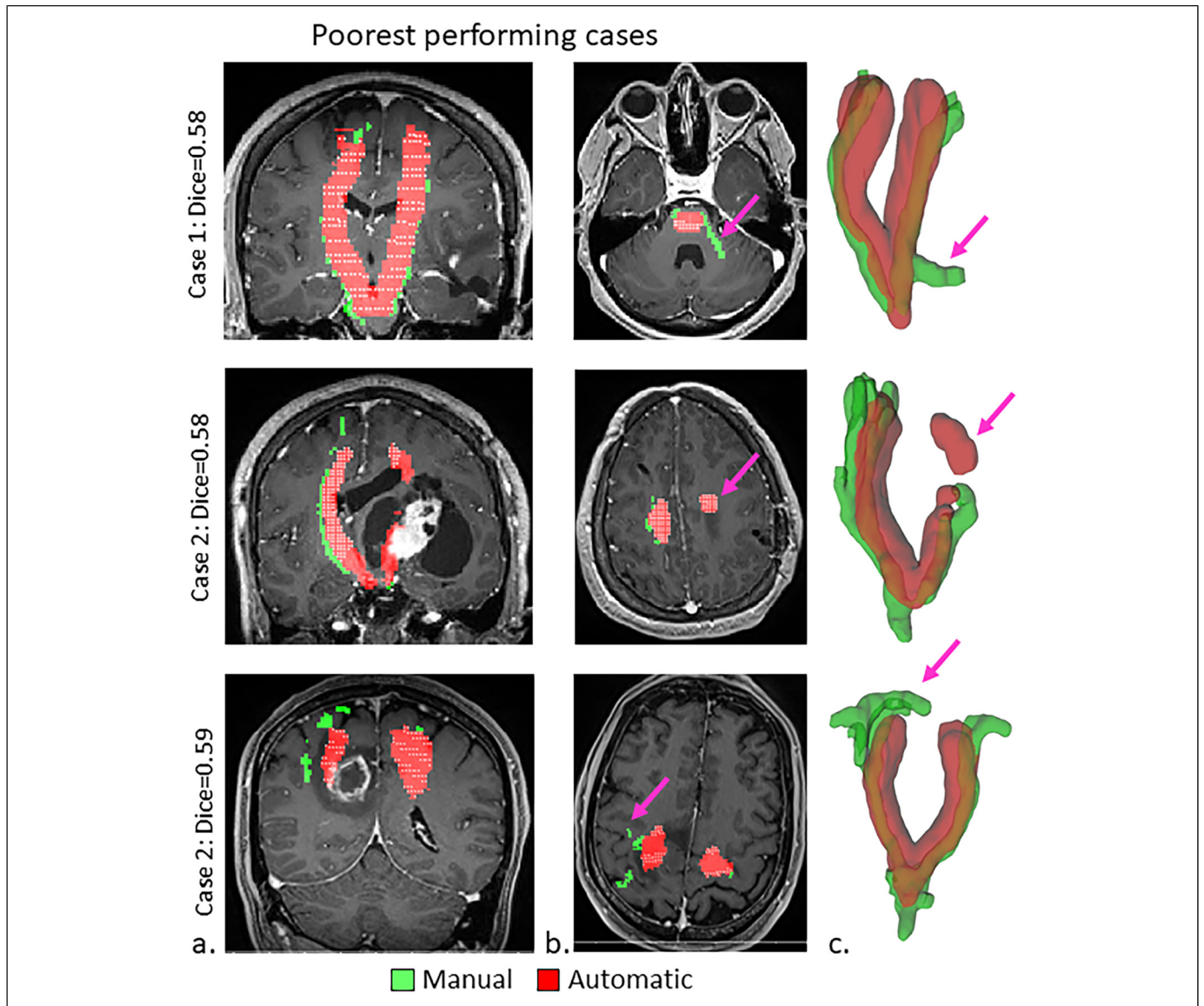
Figure 4. Examples for the poorest performing cases. (a) Axial orientation; (b) coronal orientation; (c) 3D representation. Green = manual segmentation; red = automatic segmentation. The purple arrow indicates areas of mismatch between the manual and automatic methods.

Between measurements consistency assessment: was evaluated using correlations of fibers volume, FA, and MD values, on the dataset of healthy subjects scanned twice, and based on:

1. *Interrater reliability* (*manual method*): was evaluated for the manual CST segmentation obtained by raters 1 and 2.
2. *Repeatability* (*automatic method*): was evaluated for automatic CST segmentation obtained in the repeated measurements (time points 1 and 2).

Results

Segmentation Results

Following optimization, network training was carried out with a batch size = 47 with a learning rate of 0.001, with a threshold of 0.4, with Dice score as the metric for model evaluation during training, and with a total of 50 epochs, while preserving the model which achieved the best level of accuracy during training.

For the automatic CST segmentation Dice scores of 0.63 ± 0.04 and 0.64 ± 0.04 were obtained for the training and testing datasets. A representative result of the automatic CST segmentation in a patient with high-grade glioma is given in Figure 1.

Between Measurements Consistency Assessment

Interrater reliability (manual method) and repeatability (automatic method) results are given in Table 2 and Figure 2. A significant higher between measurements correlation was obtained for the automatic versus the manual method for the volume and MD values.

TractSeg Performance in Patients With Lesions Adjacent to the CST

A total of 140 out of 225 patients (62%) were annotated as having lesions adjacent to the CST (65 from the validation and 75 from the test groups). For the automatic CST segmentation Dice scores of 0.641 ± 0.02 and 0.642 ± 0.047 were obtained for the training and testing subsets, similar to that obtained for the entire cohort. Figure 3 demonstrates the CST segmentation results in 6 patients with lesions adjacent to the CST, showing the high similarity between the manual and the automatic segmentation results. This finding demonstrated the robustness of the automated method even in cases with disordered brains.

Figure 4 shows the examples for the poorest performing cases, detected based on the lowest dice score and visual inspection. Purple arrows indicate the areas of mismatch between the manual and automatic methods. Cases $1 + 3$, demonstrate a mismatch due to the additions/expansion of the fiber in the manual segmentation. Yet those mismatches do not significantly affect the correct fiber identification and localization in both methods. Case 2, in which there is extensive pathology in the CST trajectory, demonstrated the superiority of the automatic method, which successfully enabled reconstruction of the upper segment of the CST in the affected hemisphere, at the level in the manual method failed to do.

Discussion

In this work, we demonstrated the usage of the TractSeg method for automatic CST segmentation in patients with brain pathologies. The model demonstrated robustness by coping with variability within the dataset in terms of brain deformation, expressed in a high variety of forms depending on the dimensions and type of brain pathology; as well as high variety in acquisition parameters, with the DTI data acquired with different vendors, systems and various acquisition parameters of MRI.

The TractSeg algorithm is a deep learning-based approach, which has become the state-of-the-art technique in the field of computer vision, leading to enhanced performance relative to conventional machine learning methods in various medical vision applications.[23] Deep learning approach was previously applied for WM tracts segmentation tasks based on diffusion MRI data, demonstrating the promising potential of this approach,[12,18] yet its application in a clinical setup is limited.

In order to evaluate the reliability of the automatic versus the manual segmentation method, we assessed consistency between repeated measurements based on the healthy control group. Substantially higher correlations between measurements were obtained for the CST segments by the automatic method indicating higher consistency for this approach. The superiority of the automatic versus the manual method in repeated measurements facilitates the use of this tool for longitudinal patient assessment (before and after therapy/surgery), and for clinical trials.

In this work, we focused on CST reconstruction as CST is the most common tract that is requested for preoperative reconstruction due to its importance in being responsible for motor control, and its spatial location which makes it more prone to be affected by brain pathology relative to other fibers. However, our results indicate that given a sufficient amount of manual labels, there is a feasibility to generate an automatic model for the reconstruction of other brain fibers in a clinical setup.

Study limitations include the absence of reliable ground truth manual labels, which makes it difficult to train and evaluate an accurate segmentation model. Another limitation is the absence of a sufficient amount of manual labels for other brain fibers, important for preoperative mapping as the optic radiation, longitudinal fasciculus, arcuate fasciculus, etc, which prevent their inclusion in the proposed automated segmentation model.

Conclusion

The TractSeg method was implemented for automatic segmentation in patients with brain pathologies, demonstrating superior consistency for the TractSeg-based CST segmentation. This finding indicates the advantage of using the automatic approach in comparison to the manual fiber segmentation approach, especially for longitudinal studies in patients, in a way that may yield more valid results while minimizing human time consumption.

ORCID iD

Moran Artzi https://orcid.org/0000-0003-4214-1875

References

1. Schult T, Hauser T-K, Klose U, Hurth H, Ehricke H-H. Fiber visualization for preoperative glioma assessment: Tractography versus local connectivity mapping. *PLoS One*. 2019;14(12):e0226153.
2. Ashmore J, Pemberton HG, Crum WD, Jarosz J, Barker GJ. Implementation of clinical tractography for pre-surgical planning of space occupying lesions: an investigation of common acquisition and post-processing methods compared to dissection studies. *PLoS One*. 2020;15(4):e0231440.

3. Essayed WI, Zhang F, Unadkat P, Cosgrove GR, Golby AJ, O'Donnell LJ. White matter tractography for neurosurgical planning: a topography-based review of the current state of the art. *NeuroImage: Clinical*. 2017;15(15):659-672.

4. Villanueva-Meyer JE, Mabray MC, Cha S. Current clinical brain tumor imaging. *Neurosurgery*. 2017;81(3):397-415.

5. Soares J, Marques P, Alves V, Sousa N. A Hitchhiker's guide to diffusion tensor imaging. *Front Neurosci*. 2013;7(12):31.

6. O'Donnell LJ, Westin C-F. An introduction to diffusion tensor image analysis. *Neurosurgery Clinics*. 2011;22(2):185-196.

7. Jellison BJ, Field AS, Medow J, Lazar M, Salamat MS, Alexander AL. Diffusion tensor imaging of cerebral white matter: a pictorial review of physics, fiber tract anatomy, and tumor imaging patterns. *American Journal of Neuroradiology*. 2004;25(3):356-369.

8. Jeurissen B, Descoteaux M, Mori S, Leemans A. Diffusion MRI fiber tractography of the brain. *NMR Biomed*. 2019;32(4): e3785.

9. Mukherjee P, Chung S, Berman J, Hess C, Henry R. Diffusion tensor MR imaging and fiber tractography: technical considerations. *Am J Neuroradiology*. 2008;29(5):843-852.

10. Rheault F, De Benedictis A, Daducci A, et al. Tractostorm: the what, why, and how of tractography dissection reproducibility. *Hum Brain Mapp*. 2020;41(7):1859-1874.

11. Shi D, Pan Z, Li X, Guo H, Zheng Q. Diffusion coefficient orientation distribution function for diffusion magnetic resonance imaging. *J Neurosci Methods*. 2021;348(15):108986.

12. Poulin P, Cote M-A, Houde J-C, et al. *Learn to track: deep learning for tractography*. Springer; 2017. 540-547.

13. Lam PDN, Belhomme G, Ferrall J, Patterson B, Styner M, Prieto JC. TRAFIC: fiber tract classification using deep learning. *Int Soc Optics Photonics*. 2018;10574:1057412.

14. Gupta V, Thomopoulos SI, Rashid FM, Thompson PM. *FiberNET: an ensemble deep learning framework for clustering white matter fibers*. Springer. 2017:548-555.

15. Gupta V, Thomopoulos SI, Corbin CK, Rashid F, Thompson PM. *FIBERNET 2.0: an automatic neural network based tool for clustering white matter fibers in the brain*. IEEE. 2018:708-711.

16. Wassermann D, Makris N, Rathi Y, et al. The white matter query language: a novel approach for describing human white matter anatomy. *Brain Struct Funct*. 2016;221(9):4705-4721.

17. Richards TJ, Anderson KL, Anderson JS. Fully automated segmentation of the corticospinal tract using the TractSeg algorithm in patients with brain tumors. *Clin Neurol Neurosurg*. 2021;210:107001.

18. Wasserthal J, Neher P, Maier-Hein KH. TractSeg-fast and accurate white matter tract segmentation. *NeuroImage*. 2018;183:239-253.

19. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. Fsl. *Neuroimage*. 2012;62(2):782-790.

20. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging*. 2015;15(1):1-28.

21. Neurovault. JHU DTI-based white-matter atlases. https://identifiers. org/neurovault.collection:264.

22. University m. MNI152 standard-space T1-weighted average structural template image. http://nist.mni.mcgill.ca/atlases/.

23. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik*. 2019;29(2):102-127.