Contents lists available at ScienceDirect

Computers in Biology and Medicine

Enhancing the estimation of fiber orientation distributions using convolutional neural networks

Oeslle Lucena [a,*], Sjoerd B. Vos [b,c], Vejay Vakharia [d], John Duncan [d,e], Keyoumars Ashkan [f], Rachel Sparks [a], Sebastien Ourselin [a]

[a] School of Biomedical Engineering and Imaging Sciences, King's College London, UK
[b] Centre for Medical Image Computing, Department of Computer Sciences, University College London, London, UK
[c] Neuroradiological Academic Unit, University College London Queen Square Institute of Neurology, University College London, London, UK
[d] Department of Clinical and Experimental Epilepsy, University College London, UK
[e] National Hospital for Neurology and Neurosurgery, Queen Square, UK
[f] King's College Hospital Foundation Trust, UK

ARTICLE INFO

ABSTRACT

Local fiber orientation distributions (FODs) can be computed from diffusion magnetic resonance imaging (dMRI). The accuracy and ability of FODs to resolve complex fiber configurations benefits from acquisition protocols that sample a high number of gradient directions, a high maximum b-value, and multiple b-values. However, acquisition time and scanners that follow these standards are limited in clinical settings, often resulting in dMRI acquired at a single shell (single b-value). In this work, we learn improved FODs from clinically acquired dMRI. We evaluate patch-based 3D convolutional neural networks (CNNs) on their ability to regress multi-shell FODs from single-shell FODs, using constrained spherical deconvolution (CSD). We evaluate U-Net and High-Resolution Network (HighResNet) 3D CNN architectures on data from the Human Connectome Project and an in-house dataset. We evaluate how well each CNN can resolve FODs 1) when training and testing on datasets with the same dMRI acquisition protocol; 2) when testing on a dataset with a different dMRI acquisition protocol than used to train the CNN; and 3) when testing on a dataset with a fewer number of gradient directions than used to train the CNN. This work is a step towards more accurate FOD estimation in time- and resource-limited clinical environments.

1. Introduction

Diffusion magnetic resonance imaging (dMRI) can be used to investigate the organization of white matter (WM). WM is composed of neuronal axon fiber bundles that impose a preferential direction to the diffusion of water molecules [1,2], resulting in anisotropic diffusion along the axon fiber bundles [3,4].

WM tissue microstructural organization, such as axon diameter [5] and local fiber orientation distribution (FOD or fODF) [6], can be estimated from dMRI acquisitions. FODs can be used to perform fiber tractography [7,8], which has an important role in presurgical planning [9–12] and connectome analyses [13]. A common method to estimate local fiber orientation is diffusion tensor imaging (DTI) [3]. However, DTI only models single fiber populations and cannot resolve complex fiber configurations such as fiber crossings [6]. More robust methods for

representing FODs have been proposed based on spherical deconvolution [14–16] or other approaches that estimate diffusion orientation distribution functions from q-space [16,17].

Constrained spherical deconvolution (CSD) estimates a more complex FOD from dMRI signals [4,18]. Single-shell single-tissue CSD (S-CSD) models each voxel as a single compartment with one corresponding FOD, irrespective of the underlying tissue components [18]. However, S-CSD FODs suffer from partial volume effects (PVEs) when multiple tissues are present. Multi-shell multi-tissue CSD (M-CSD) reduce PVEs by modeling one anisotropic compartment, corresponding to WM, and two isotropic compartments, corresponding to grey matter (GM) and cerebrospinal fluid (CSF) [4]. M-CSD relies on decay behaviors of distinct tissue types across multiple b-values (or shells) to separate each voxel into three components. However, M-CSD is limited in clinical settings since it requires dMRI acquisitions with multiple shells (MS)

which has longer acquisition times compared to DTI or single-shell (SS) dMRI.

Single-shell 2-tissue CSD (2TS-CSD) and single-shell 3-tissue CSD (SS3T-CSD) [19] attempt to resolve PVEs by modeling isotropic compartments, similar to M-CSD, using the b = 0 s/mm^2 image as a second shell. With the additional shell, a multi-tissue signal profile is computed by 2TS-CSD assuming one anisotropic component (WM) and one isotropic component (either GM or CSF). SS3T-CSD uses an iterative approach to fit a CSD model for the three tissues compartments by first fitting one pair of components (WM and GM) and then fitting another pair of components (WM and CSF) [19].

FODs that are able to resolve complex fiber configurations typically require specific dMRI acquisition protocols with a high number of gradient directions (b-vecs), a high maximum b-value and/or multiple b-values [20–23]. A higher number of gradient directions and higher b-values improve FOD angular resolution, enabling FODs to better resolve complex fiber configurations, such as fiber crossings [24,25]. However, local fiber reconstruction is more accurate for images with a high signal-to-noise ratio (SNR), which can be achieved from low b-value shells [25]. Despite these advantages, clinical uptake is limited due to longer acquisition times and the need for expert staff to set up the acquisition protocols [26]. Therefore, improving FOD modeling for clinically available dMRI acquisitions is an active topic of research.

Deep learning (DL) has been successfully used for a variety of medical imaging tasks [27] to learn an underlying mathematical representation that non-linearly maps data between representations [28]. DL regression has been presented to improve dMRI spatial resolution [29], to map model coefficients from q-space [30] to either neurite orientation dispersion and density imaging (NODDI) [31] or diffusion kurtosis imaging (DKI) models [32], to output high quality CSD coefficients from CSD coefficients computed on dMRI with fewer gradient directions [33], to fit SH coefficients between different shells [34], and to map from S-CSD coefficients to FODs computed from histology [35].

In Tanno et al. [29]; a CNN patch-based regression was presented to infer higher resolution patches and quantify uncertainty. In a different work, Golkov et al. [30] computed model coefficients (DKI or NODDI) from q-space signal intensities using a multilayer perceptron network (MLP). In Koppers et al. [34]; a MLP was trained using SH coefficients calculated from one or a combination of shells to infer the SH coefficients of the same order for a different shell. In Nath et al. [35]; a neural network composed of regular hidden and residual layers (ResDNN) was trained to map S-CSD coefficients to FODs computed from macaque histology. In Lin et al. [33]; 3D CNNs were used to estimate M-CSD coefficients of a good quality MS dMRI acquisition from M-CSD coefficients computed from dMRI with fewer gradient directions.

In this work, we aim to compute a more accurate and reliable FOD from SS dMRI, the most common dMRI acquisition in clinical applications. Our goal is to demonstrate that 3D CNNs, U-Net [36] and High-Resolution Network (HighResNet) [37], can regress M-CSD coefficients from 2TS-CSD coefficients. We present an application of CNNs that has not been proposed elsewhere in the literature and is non-obvious from other proposed regression tasks for dMRI which have focused on regression between shells using MLPs [34] or regression from few gradient directions to many gradient directions for MS acquisitions [33]. We present an extensive evaluation involving: (1) training and testing on datasets with the same dMRI acquisition protocol, (2) testing on datasets with different dMRI acquisition protocols than the training dataset, and (3) regressing from dMRI with fewer gradient directions than the training dataset.

2. Methodology

2.1. Pipeline overview

Our pipeline consists of the following steps. We construct a paired dataset composed of one SS dMRI and one MS dMRI (Section 2.2) from

the Human Connectome Project (HCP) [38] and an in-house dataset which we refer to as QS dataset (see Section 3.1 for a detailed description of the datasets). For each dataset, we compute two CSD models, 2TS-CSD and M-CSD (Section 2.3) from the SS dMRI and MS dMRI, respectively. Then, we train a CNN to regress the M-CSD coefficients from the 2TS-CSD coefficients (Section 2.4).

2.2. Training dataset

From a MS dMRI a paired SS dMRI dataset is constructed by selecting one shell from the MS dMRI. In this work, we select a shell that has at least a minimum number of gradient directions for a given b-value that can characterize the angular frequency components of the dMRI signal [25]. For the HCP dataset, we construct a SS dMRI for each of the three b-values (1000, 2000, 3000 s/mm^2) where each shell has 90 directions. For the QS dataset, we construct a SS dMRI for two b-values (700, 2500 s/mm^2) with 32 and 64 directions, respectively.

2.3. CSD modeling

CSD models FODs from the dMRI signal intensities. A dMRI signal intensity is approximated by convolving CSD coefficients with a signal attenuation profile, also known as a response function [18]. The CSD coefficients are computed by minimizing differences between the dMRI signal intensities and signal intensities approximated from the CSD coefficients using iterative linear least-squares. A soft regularizer is also applied to enforce a non-negativity constraint on the CSD coefficients.

Although $l_{max} = 8$ provides FODs with a high angular contrast for high b-values (b > 1000 s/mm^2) [25], in this work, we focused on $l_{max} = 4$, comprising 15 coefficients, to establish proof of concept.

After computing CSD coefficients, we applied multi-tissue informed log-domain intensity normalization [39] to correct intensity inhomogeneities.

2.3.1. M-CSD

M-CSD coefficients are computed following the equation [4]:

$$\begin{bmatrix} \widehat{x}_1 \\ \vdots \\ \widehat{x}_n \end{bmatrix} = \arg\min_x \left\| \begin{bmatrix} C_{1,1} & \cdots & C_{1,n} \\ \vdots & \ddots & \vdots \\ C_{m,1} & \cdots & C_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix} \right\|_2^2$$

$$\text{subject to} \begin{bmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \geq 0 \quad (1)$$

where d_i is the vector of dMRI signal intensities in the i-th shell, x_j is an unknown vector of coefficients for the FOD of tissue j, $C_{i,j}$ is the matrix relating the coefficients for the FOD of tissue j to the dMRI signal intensities measured on the i-th shell in q-space by spherical convolution. An additional constraint is imposed on the coefficients where A_j is a tissue specific matrix relating the coefficients of the FOD for tissue j to their signal amplitudes, effectively imposing non-negativity on all coefficients. To perform the optimization, we use the algorithm available within MRtrix [40].

2.3.2. 2TS-CSD

For the 2TS-CSD, a similar approach to the M-CSD is used where the b = 0 s/mm^2 (b-zero) image is used as a second shell. To ensure Equation (1) has a unique solution, we set $j = 2$, reducing the number of tissue components to two. We model one compartment as isotropic corresponding to CSF, and the other compartment as anisotropic, corresponding to WM. The CSF is selected as the isotropic compartment as it leads to a more accurate fit of the FOD compared to using GM as the isotropic compartment [19].

2.4. CNN training

We evaluate two 3D CNN architectures, a 3D High-Resolution Network (HighResNet) [41] and a 3D U-Net [36], on their ability to regress M-CSD coefficients from 2TS-CSD coefficients. Fig. 1 shows a graphical representation of the network architectures. For both networks, patch-based training is used. Patch-based training requires a reduction in the effective receptive field (ERF) to reflect the selected patch size (Section 2.5) and avoid distortions at the patch boundary [42]. Details of how the ERF is reduced for each network are described below.

2.4.1. HighResNet

The original HighResNet architecture is comprised of three levels of dilated convolutions and nine residual connections resulting in 0.81 M trainable parameters. HighResNet was originally proposed as a compact network that could achieve large ERFs [42] without requiring a downsample-upsample pathway to capture low and high level features [36,43]. Dilated convolutions are used to produce accurate predictions and detailed probabilistic maps near object boundaries [41]. In this paper, we modified the HighResNet architecture by reducing the number of layers to achieve the desired ERF. The final HighResNet architecture comprises two levels of dilated convolutions and four residual connections resulting in 0.16 M trainable parameters. A parametric rectified linear unit (PReLU) activation function was used in place of a ReLU. PReLU adaptively learns the rectifier parameters and has been shown to improve CNNs performance in other applications [44].

2.4.2. U-Net

The original 3D U-Net is a "U"-shaped network that has a downsample-upsample pathway composed of 14 convolutional layers [36] resulting in 19.08 M trainable parameters. We adapted the U-Net architecture to achieve the desired ERF by reducing the network to ten convolutional layers resulting in 3.93 M trainable parameters. We removed one encoder block (2 × (conv. + batch norm. + PReLu) + max pooling) and one decoder block (concat. + up-sampling + 2 × (conv. + batch norm. + PReLu)).

2.4.3. Data augmentation

Classic techniques for on-the-fly augmentation including axis flipping, scaling, and rotation have been successfully applied to CNN training for small 3D medical imaging datasets [37,41,45]. However, applying these techniques as implemented in traditional medical image processing tools is not appropriate for CSD coefficients as they are in the SH domain and not the 3D spatial domain. Therefore, we apply 3D random rotations in the SH domain.

2.5. Training setup

Each network is trained with an RMSprop optimizer to minimize the L_2 loss between the M-CSD coefficients and the CSD coefficients regressed by the CNN, measured by $loss(y, \widehat{y}) = \frac{\|y - \widehat{y}\|_2^2}{2}$ where y are the ground truth coefficients (M-CSD) and $\widehat{y}$ are the coefficients regressed by the network. The initial parameters for each CNN are set using the He uniform initialization [44]. Each CNN is trained for 400 epochs, with a weight decay of $1E − 6$. Training starts with a learning rate of $3E − 2$, which is then reduced by 1/2 every 50 epochs. Training weight decay, initial learning rate, and number of epochs were experimentally chosen based on observed convergence. Patches are uniformly sampled within the binary mask corresponding to the intracranial space computed using the skull-stripped algorithm of MRtrix [40].

For each iteration in an epoch, a subject from the training set is randomly selected. Subsequently, CSD coefficients are augmented by applying a random FOD rotation in the range of [−25, 25] degrees. From this augmented data, 40 patches of size $32 \times 32 \times 32 \times 15$ were randomly sampled from within the intracranial space, where 15 is the number of 2TS-CSD coefficients. The number of patches were experimentally selected to achieve low validation loss while being able to be loaded on the available graphics processing unit (GPU) memory. An epoch finishes when all subject from the training set have been selected once. For every epoch, a new set of random rotations and patches are computed for each subject.

3. Experimental design

3.1. Datasets

We use two datasets to conduct our analysis: the publicly available HCP [38] and a dataset acquired at the National Hospital for Neurology and Neurosurgery, Queen Square (QS). The details of each dataset and the applied preprocessing are described below.

3.1.1. QS dataset

The QS dataset is comprised of 50 volumetric MS dMRI scans acquired from patients with epilepsy who appeared "structurally normal" on a T1-weighted MRI (T1). Small lesions may be present focally in the GM but should not distort or affect the WM FOD computation. All patients underwent MRI scanning as part of routine clinical care. Images are acquired on a 3T GE MR750 including a T1 sequence (MPRAGE) and a MS dMRI sequence with 2 mm isotropic resolution and the gradient directions 11, 8, 32, and 64 at b = 0, 300, 700, and 2500 s/mm², respectively and one b = 0 s/mm² with reverse phase-encoding. The MS dMRI is corrected for signal drift, geometric distortions and eddy-

Fig. 1. HighResNet architecture [41] and U-Net architecture [36] used in this work.

current induced distortions as in Mancini et al. [11].

3.1.2. HCP dataset

We use a subset of the HCP dataset comprised of 45 subjects [38]. Images are acquired on a 3T scanner with the following parameters: 1.25 mm isotropic resolution with 90 gradient directions for each b = 1000, 2000, and 3000 s/mm^2 and 18 images at b = 0 s/mm^2. Images in the HCP dataset are corrected following the protocols described in Sotiropoulos et al. [38] prior to download.

3.2. Evaluation metrics

To evaluate the CSD model accuracy, we compute mean absolute error (MAE) and the angular correlation coefficient (ACC) [46] for all voxels in the WM (Section 2.3). WM voxels are identified using geodesic information flows (GIF) to segment the WM as a binary mask [47]. The "ground truth" CSD is the M-CSD computed using dMRI with all gradient directions and shells.

MAE measures how well CSD coefficients match with lower values indicating high similarity to the ground truth CSD coefficients. MAE is computed as $MAE(y, \widehat{y}) = \frac{|y - \widehat{y}|}{n}$ where y is the ground truth M-CSD coefficients and $\widehat{y}$ are either the CSD coefficients inferred from a trained network or 2TS-CSD coefficients.

ACC is computed between two different sets of SH coefficients u, the M-CSD coefficients, and v, from either the CSD coefficients inferred from a trained network or 2TS-CSD coefficients, both models have the SH order j, where

$$ACC(u, v) = \frac{\sum\limits_{j=1}^{\infty} \sum\limits_{m=-j}^{j} u_{j,m} v_{j,m}^*}{\left[\sum\limits_{j=1}^{\infty} \sum\limits_{m=-j}^{j} u_{j,m}^2 \right]^{0.5} \left[\sum\limits_{j=1}^{\infty} \sum\limits_{m=-j}^{j} v_{j,m}^2 \right]^{0.5} + \alpha} \qquad (2)$$

ACC are in the range [−1, 1], where 1 implies a perfect linear correlation between two functions on a sphere, and −1 implies a negative correlation [48]. Note, in this work the value −1 is not achievable due to the non-negativity constraint imposed on the CSD coefficients.

3.3. Experiments

We assess CNN performance in the following scenarios: 1) training and testing on datasets with the same dMRI acquisition protocol (*Experiment 1 – Intra-Scanner Acquisition Performance*); 2) testing on datasets with different dMRI acquisition protocols than the training dataset (*Experiment 2 – Inter-Scanner Acquisition Performance*); and 3) testing on datasets with fewer dMRI gradient directions than the training dataset (*Experiment 3 – Downsampled Imaging Performance*). Additionally, we assess performance variability within specific brain regions. The details of each experiment are described below. For all experiments, the "ground truth" CSD is the M-CSD computed using the complete set of gradient directions and shells (Section 2.2).

3.3.1. Experiment 1 – Intra-Scanner Acquisition Performance

We assess how well CNNs were able to regress M-CSD coefficients when using dMRI acquired on the same scanner and with the same protocol for training and testing. In this experiment, 5-fold cross-validation is conducted where 3 folds are used for training, 1 fold for validation, and 1 fold for testing.

3.3.2. Experiment 2 – Inter-Scanner Acquisition Performance

We assess the generalizability of CNNs to regress M-CSD coefficients from dMRI acquired from different acquisition protocols and a different scanner. We use CNNs trained in Experiment 1 – Intra-Scanner Acquisition Performance, without additional model tuning, to test on a dataset with different dMRI acquisition protocol than used during training. For

example, a CNN trained on QS dataset using 2TS-CSD coefficients derived from the b = 700 s/mm^2 shell is evaluated on the HCP dataset using 2TS-CSD coefficients derived from the b = 2000 s/mm^2 shell.

3.3.3. Experiment 3 – Downsampled Imaging Performance

We assess the robustness of CNNs when acquisitions have fewer gradient directions than the training dataset. Using the same test datasets as in Experiment 1 – Intra-Scanner Acquisition Performance, we subsampled the number of gradient directions in the dMRI by 25%, 50%, and 75%. CNNs trained from Experiment 1 – Intra-Scanner Acquisition Performance, with no further tuning, were used to regress M-CSD coefficients from 2TS-CSD coefficients computed from the subsampled dMRI. CNNs were tested with both 2TS-CSD coefficients computed from the same dMRI acquisition protocol (as in Experiment 1 – Intra-Scanner Acquisition Performance) and dMRI with different acquisition protocols (as in Experiment 2 – Inter-Scanner Acquisition Performance).

Each subsampled dMRI was created as follows. For the original dMRI dataset, we first reorder the set of gradient directions such that if a scan is terminated prematurely, at any point, the acquired gradient directions will still be close to optimally distributed on the half-sphere [40]. Then, we truncated the number of gradient directions for both b = 0 s/mm^2 and the selected shell to generate the SS dMRI with the desired reduction in gradient directions. Finally, the 2TS-CSD is computed for the subsampled SS dMRI as described in Section 2.3.

3.3.4. Variability within brain regions

We assess the performance of the CNNs to regress M-CSD coefficients within the specific brain regions: frontal, occipital, parietal, and temporal lobes and the corpus callosum (CC). All brain regions were determined from a segmentation generated using [47] with WM labels derived from the Hammers Atlas [49].

3.4. Implementation

All experiments were performed on a workstation equipped with an Intel CPU (Xeon® W-2123, 8 × 3.60 GHz; Intel), 32 GB of memory and a NVIDIA GPU (GeForce Titan V) with 12 GB of on-board memory. All code was implemented in Python 3.6. PyTorch 1.6.0 [50] was used for network training. NiftyNet 0.4.0 [37] was used for data loading and sampling. Data augmentation was performed using SHtools 4.5 [51]. The code used to train the CNNs is available online.[1]

4. Results

For all Tables and Figures, the acronym QS 700-HCP 2000 CNN U-Net indicates that CSD coefficients are from a CNN trained on the QS dataset using the 2TS-CSD coefficients from b = 700 s/mm^2 shell and tested using as input 2TS-CSD coefficients from the HCP dataset from the b = 2000 s/mm^2 shell. The method 2TS-CSD (QS 700) indicates the baseline approach computed from the QS dataset computed for the b = 700 s/mm^2 shell. The "ground truth" is always the M-CSD coefficients computed using the complete set of gradient directions and shells.

4.1. Experiment 1 – Intra-Scanner Acquisition Performance

Table 1 reports the mean, standard deviation, and median MAE and ACC between the indicated CSD coefficients and the M-CSD coefficients when training and testing on data with the same dMRI acquisition protocol. CSD coefficients from U-Net are the most similar to M-CSD coefficients while the baseline 2TS-CSD coefficients are the least similar. Pronounced ACC improvements are found in the QS dataset (9% for the mean ACC) compared to the HCP dataset (5% for the mean ACC) for the best performing CNNs.

---

[1] https://github.com/OeslleLucena/RegressionFOD.

Table 1
Mean, standard deviation (std), and median MAE and ACC measured between M-CSD and CSD coefficients for Experiment 1. The best value, minimum value for MAE and maximum value for ACC, are indicated by bold text. Train-Test indicates the dataset used for training CNNs and inference, respectively. HR indicates HighResNet.

Train-Test	Method	MAE	ACC	
		mean(std)	mean(std)	median
–	*2TS-CSD* (QS 700)	0.972 (0.079)	0.851 (0.133)	0.882
QS 700-QS 700	*CNN* HR	0.450 (0.048)	**0.928 (0.088)**	0.959
	CNN U-Net	**0.440 (0.038)**	0.927 (0.097)	**0.960**
–	*2TS-CSD* (QS 2500)	0.826 (0.070)	0.917 (0.107)	0.966
QS 2500-QS 2500	*CNN* HR	**0.307 (0.031)**	**0.958 (0.069)**	**0.981**
	CNN U-Net	0.311 (0.030)	0.954 (0.077)	**0.981**
–	*2TS-CSD* (HCP 1000)	0.815 (0.130)	0.895 (0.117)	0.939
HCP 1000-HCP 1000	*CNN* HR	0.303 (0.038)	**0.961 (0.057)**	**0.983**
	CNN U-Net	**0.298 (0.042)**	0.960 (0.063)	**0.983**
–	*2TS-CSD* (HCP 2000)	0.742 (0.123)	0.909 (0.107)	0.953
HCP 2000-HCP 2000	*CNN* HR	0.284 (0.042)	**0.961 (0.062)**	**0.984**
	CNN U-Net	**0.283 (0.038)**	0.960 (0.066)	**0.984**
–	*2TS-CSD* (HCP 3000)	0.737 (0.109)	0.899 (0.112)	0.941
HCP 3000-HCP 3000	*CNN* HR	0.310 (0.046)	0.951 (0.077)	**0.980**
	CNN U-Net	**0.303 (0.041)**	**0.952 (0.075)**	**0.980**

Fig. 2 shows cumulative distribution functions (CDFs) for ACC. CSD coefficients regressed from CNNs have curves more skewed torward high ACC, indicating more voxels with a high level of agreement to the M-CSD coefficients, compared to 2TS-CSD coefficients.

Figs. 4 and 5 show qualitative heatmaps of the ACC for the WM

voxels. ACC heatmaps have a high similarity across the three tested methods in WM voxels far from boundaries, where PVEs are minimal, and low correlation in voxels near WM boundaries, which are most likely to have PVEs.

Figs. 6 and 7 present a glyph representation of FODs computed from the CSD coefficients, showing the fiber direction and distribution per voxel. The glyphs qualitatively show that CSD coefficients regressed from either U-Net or HighResNet are better able to resolve multiple fiber populations, small rotations and scaling within the selected regions than the 2TS-CSD coefficients.

Additional experiments where we trained the CNNs using $l_{max} = 8$, comprising 45 coefficients, are found in the Supplementary Material Section 1. Similar results were found for $l_{max} = 8$ as for $l_{max} = 4$. Finally, in Section 3 of Supplementary Material, we qualitatively evaluated the ability of the regressed CSD coefficients to perform tractography on QS 2500 and HCP 2000, generating four WM language-related tracts using ROI-based probabilistic tractography [11]. Tractography performed using CSD coefficients regressed from CNNs results in tracts more similar to tracts generated from M-CSD coefficients and with fewer spurious streamlines than tracts generated from 2TS-CSD coefficients (Supplementary Figures 2-3).

4.2. Experiment 2 – Inter-Scanner Acquisition Performance

We evaluate how well the CNNs perform when testing on datasets with different dMRI acquisition protocols than used during training. As models trained with the HCP dataset for different b-values had similar performance in Experiment 1 – Intra-Scanner Acquisition Performance (Section 4.1), we select one model (HCP 2000) for evaluation in this experiment.

Table 2 reports the mean, standard deviation, and median MAE and ACC between the indicated CSD coefficients and the M-CSD coefficients. U-Net and HighResNet were quantitatively more similar to M-CSD coefficients than the 2TS-CSD coefficients.

CSD coefficients regressed from CNNs have greater similarity to M-CSD coefficients compared to the 2TS-CSD coefficients indicated by the CDFs skewing toward higher ACC values as shown in Fig. 3.

Figs. 4 and 5 show ACC heatmaps for the WM voxels. ACC heatmaps show higher errors in WM voxels far from the boundary in the CSD
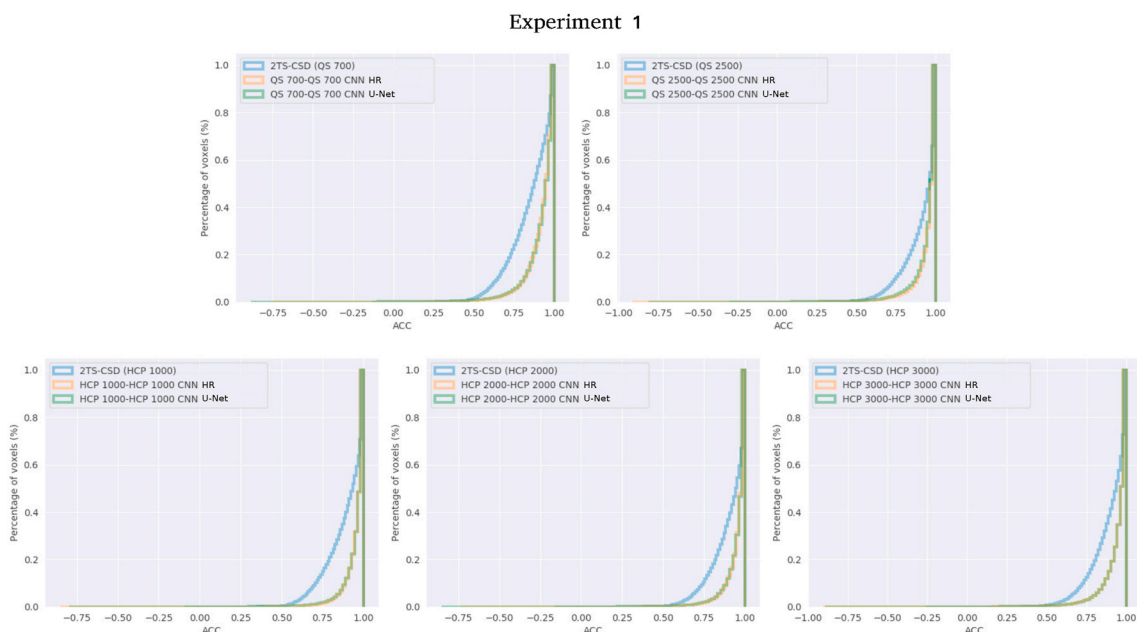
Experiment 1

Fig. 2. ACC cumulative distribution functions (CDFs) for Experiment 1. The ideal ACC CDF would give a single peak at one. Note HighResNet (HR) and U-Net have similar CDFs, causing overlay of the green and orange lines.
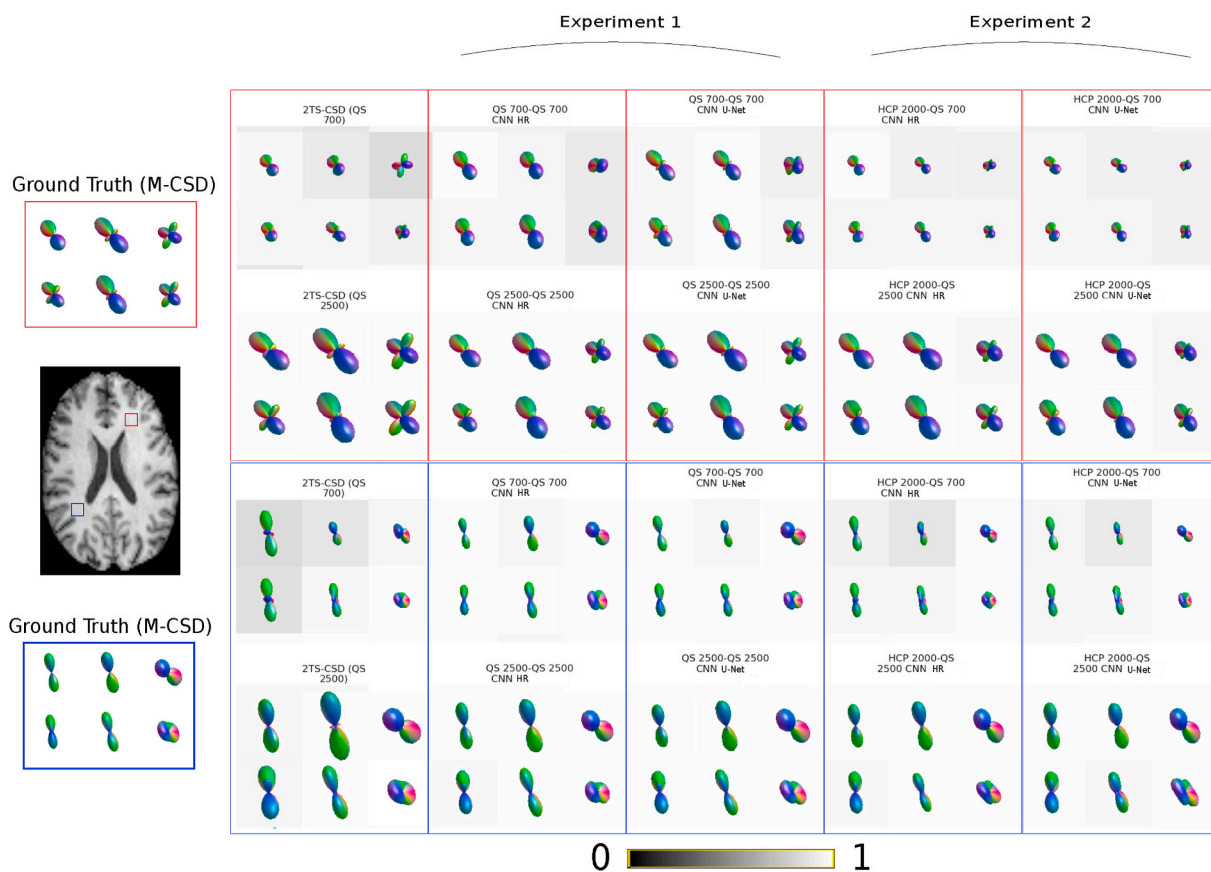
Experiment 2

Fig. 3. ACC Cumulative distribution functions (CDFs) for Experiment 2. The ideal ACC CDF would give a single peak at one. HR indicates HighResNet.

Fig. 4. ACC heatmaps overlaid on a T1 for one subject from the QS dataset. ACC is displayed in the JET colormap scaled between [0, 1]. HR indicates HighResNet.


Fig. 5. ACC heatmaps overlaid on a T1 for one subject from the HCP dataset. ACC is displayed in the JET colormap scaled between [0, 1]. HR indicates HighResNet.

Fig. 6. A glyph representation of the FODs for the regions indicated by the red and blue boxes (magnified by 3 ×) for one subject from the QS dataset. ACC is displayed in grayscale scaled between [0, 1]. HR indicates HighResNet.
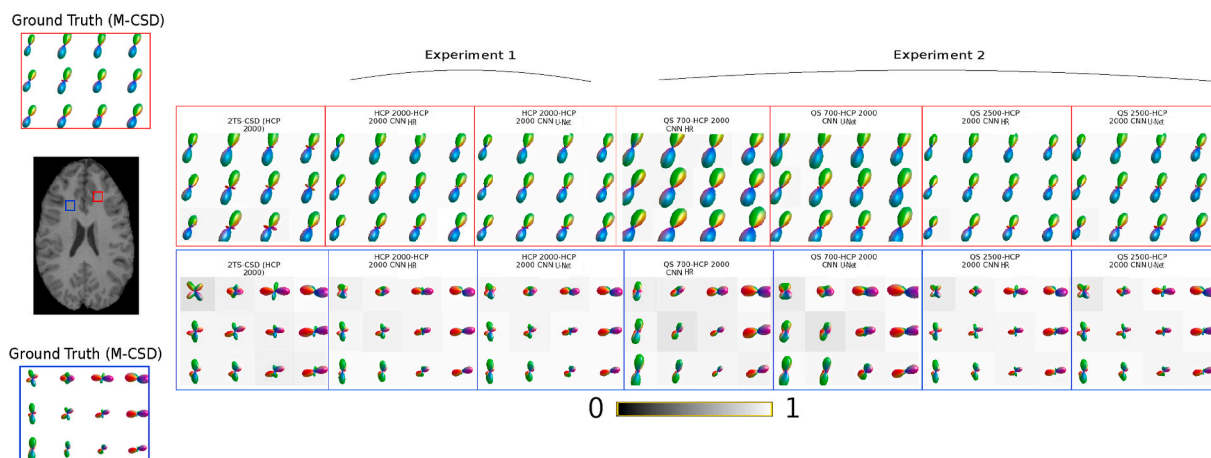
Fig. 7. A glyph representation of the FODs for the regions indicated by the red and blue boxes (magnified by 1.5 ×) for one subject from the HCP dataset. ACC is displayed in grayscale scaled between [0, 1]. HR indicates HighResNet.

coefficients regressed from CNNs for Experiment 2 – Inter-Scanner Acquisition Performance compared to Experiment 1 – Intra-Scanner Acquisition Performance. However, CSD coefficients regressed from CNNs better capture finer details in regions containing fiber crossings and voxels near the WM boundaries compared to the 2TS-CSD coefficients.

Figs. 6 and 7 show a visual representation of the FODs as glyphs. FODs for the QS dataset show a higher qualitative agreement in single fiber populations compared to multiple fiber populations. FOD amplitudes were observed to be smaller than expected for the QS 700 dataset

when CSD coefficients were regressed with CNNs training on the HCP dataset. This may occur due to a poor angular contrast in the low b-value (b = 700 s/mm^2 32 directions) of the testing data, which the CNNs cannot properly regress.

As shown in Fig. 7, when evaluating on the HCP dataset, CNNs trained on QS 2500 were capable of resolving fiber crossings while CNNs trained on QS 700 did not capture these finer details. However, both CNNs resolved FOD scaling and small rotations.

Table 2

Mean, standard deviation (std), and median MAE and ACC measured between M-CSD and CSD coefficients for Experiment 2. The best value, minimum value for MAE and maximum value for ACC, are indicated by bold text. Train-Test indicates the dataset used for training CNNs and inference, respectively. HR indicates HighResNet.

Train-Test	Method	MAE	ACC	
		mean(std)	mean(std)	median
–	*2TS-CSD* (QS 700)	0.972 (0.079)	0.851 (0.133)	0.882
HCP 2000-QS 700	*CNN* HR	**0.702 (0.069)**	**0.910 (0.096)**	**0.943**
	CNN U-Net	0.702 (0.075)	0.902 (0.099)	0.935
–	*2TS-CSD* (QS 2500)	0.826 (0.070)	0.917 (0.107)	0.966
HCP 2000-QS 2500	*CNN* HR	**0.428 (0.039)**	**0.945 (0.075)**	**0.971**
	CNN U-Net	0.469 (0.034)	0.938 (0.081)	0.967
–	*2TS-CSD* (HCP 2000)	0.742 (0.123)	0.909 (0.107)	0.953
QS 700-HCP 2000	*CNN* HR	0.556 (0.075)	0.943 (0.070)	0.971
	CNN U-Net	0.684 (0.081)	**0.946 (0.073)**	**0.973**
QS 2500-HCP 2000	*CNN* HR	**0.441 (0.079)**	0.943 (0.074)	**0.973**
	CNN U-Net	0.453 (0.088)	0.939 (0.089)	**0.973**

4.3. Experiment 3 – Downsampled Imaging Performance

We evaluated how well CNNs regressed CSD coefficients from 2TS-CSD computed from dMRI with subsampled gradient directions. For this experiment, we selected the best CNNs from Experiment 1 – Intra-Scanner Acquisition Performance for each dataset, QS 2500 and HCP 2000.

Fig. 8 shows box plots of ACC for different amounts of gradient direction subsampling. CSD coefficients regressed from CNNs have better performance compared to 2TS-CSD coefficients for all levels of subsampling. This demonstrates robustness in our method to appropriately regress coefficients regardless of the number of gradient directions.

We qualitatively evaluated our method on its ability to generate accurate and reliable fiber bundles for QS 2500 and HCP 2000 when dMRI with 50% of the gradient directions to regress CSD coefficients was used as input. Tractography computed from CSD coefficients regressed from CNNs resulted in tracts more similar to tracts generated from M-CSD coefficients and with fewer spurious streamlines than tracts generated from 2TS-CSD coefficients as shown in Supplemental Material Figs. 2 and 3.
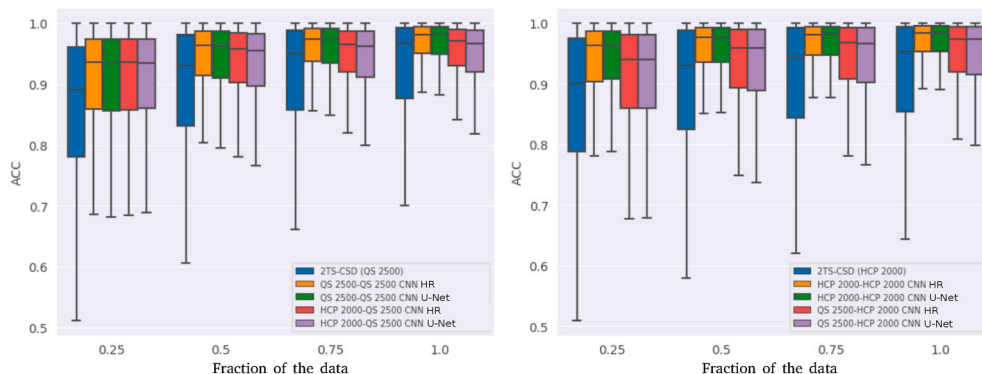
4.4. Variability within brain regions

We evaluate how well CNNs regressed CSD coefficients within specific brain regions to identify if there are spatial dependencies (Section 3.3.4). We selected the best CNNs from Experiment 1 – Intra-Scanner Acquisition Performance for each dataset, QS 2500 and HCP 2000. Fig. 9 shows ACC for the CC and frontal, occipital, parietal, and temporal lobes.

When regressing CSD coefficients for the HCP dataset, ACC between the regressed coefficients and M-CSD coefficients was higher compared to 2TS-CSD coefficients using CNNs trained on QS 2500. When regressing CSD coefficients for the QS dataset, CNNs trained on HCP 2000 had good performance in specific regions (occipital and temporal lobes) but relatively poor performance in other regions (CC, frontal and parietal lobes). However, differences in ACC for the CNN trained on HCP 2000 and the 2TS-CSD coefficients are relatively small and may be the result of random variation within models. One factor related to performance is the 2TS-CSD coefficients had relatively good performance (>0.9 for all regions), therefore, regression of CSD coefficients using CNNs had very little room to improve. The performance of CNNs to regress CSD coefficients on subsampled SS dMRI showed higher agreement with the M-CSD than the 2TS-CSD coefficients for all brain regions.
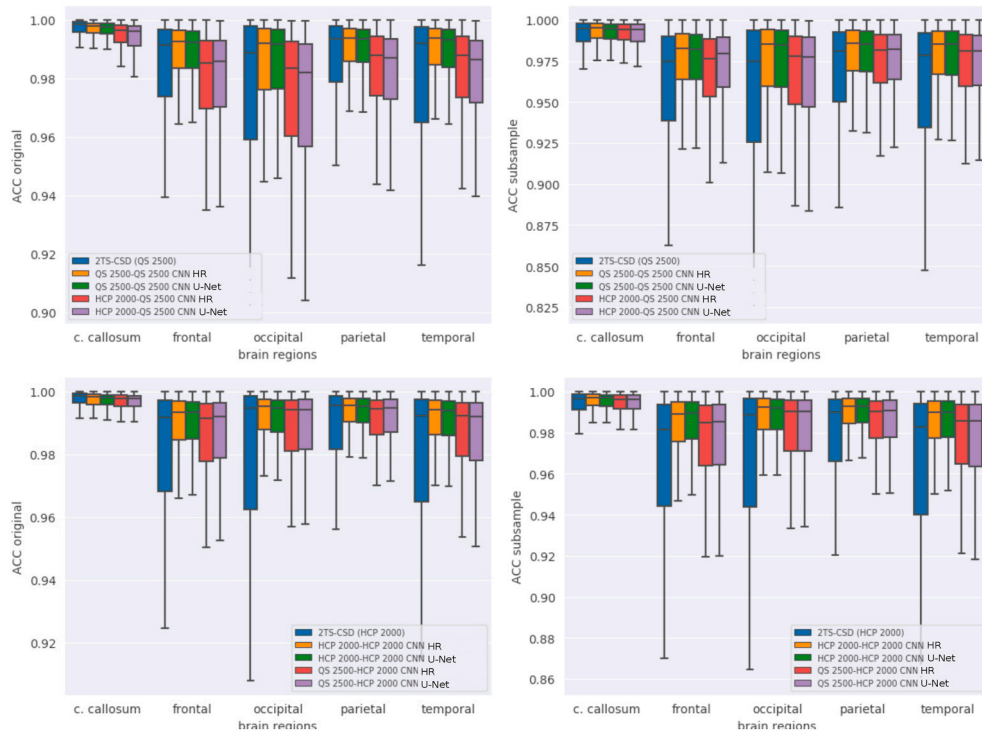
We evaluated our method on its ability to mitigate PVEs. We show a boxplot (Supplementary Figure 1) for the ACC in voxels at the WM/GM interface, the region most likely to contain PVEs. We demonstrate that CSD coefficients regressed from CNNs are more similar to M-CSD coefficients than 2TS-CSD coefficients at the WM/GM interface.

5. Discussion

We evaluated a patch-based CNN to regress M-CSD coefficients from 2TS-CSD coefficients. We demonstrated quantitatively (Tables 1 and 2 and Figs. 8 and 9) and qualitatively (Figs. 2–7) that CNNs can accurately regress CSD coefficients using data that are common in clinical settings - SS dMRI. CNNs can regress M-CSD coefficients from data with the same dMRI acquisition protocol as the training set (Experiment 1 – Intra-Scanner Acquisition Performance); generalize well to dMRI data acquired with different protocols than the training dataset (Experiment 2 – Inter-Scanner Acquisition Performance); and are robust to dMRI with fewer gradient directions than the training dataset (Experiment 3 – Downsampled Imaging Performance).

We evaluated two common neural network architectures, U-Net [36] and HighResNet [37]. Both architectures perform similarly throughout all experiments. The aim of this work was not to find the best CNN to perform CSD coefficient regression but to show the capability of DL to enhance CSD coefficients for clinically available dMRI acquisition protocols. Overall larger improvements, in terms of MAE and ACC, were observed in clinical protocols (QS) compared to research protocols (HCP). The HCP dataset has high spatial and angular resolution, which

Fig. 8. ACC for gradient directions subsampling. CNNs trained on QS 2500 (left) and on HCP 2000 (right). HR indicates HighResNet.

Fig. 9. ACC for specific brain regions tested on QS 2500 (Top) and on HCP 2000 (Bottom). ACC for specific brain regions when testing with dMRI containing all gradient directions (Left) and 50% of the gradient directions (Right). HR indicates HighResNet.

allows the 2TS-CSD to resolve complex fiber configurations and compute very accurate FODs while for the QS dataset individual shells are unable to properly capture these differences.

In this work, we investigated CNN-based regression methods for l_{max} = 4, comprising 15 coefficients, to establish proof of concept. However, we evaluated the ability to regress M-CSD coefficients for l_{max} = 8 for the best performing CNNs (QS 2500, HCP 2000) to assess a real-world setting. We found similar performance between both l_{max} orders (Supplementary Material Section 1).

Similar works have proposed DL-based approaches for regressing FODs. Koppers et al. [34] trained a MLP using SH coefficients calculated from one shell or combination of shells to infer SH coefficients of the same order for a different shell. Koppers et al. [34] regress SH coefficients between shells. In our work, we use CNNs to regress M-CSD coefficients from 2TS-CSD coefficients to go from a SS to MS representation.

Nath et al. [35] trained a neural network composed of regular hidden and residual layers (ResDNN) to regress from S-CSD coefficients to FODs computed from histology obtained from a macaque model. FODs computed from histology have higher resolution than FODs computed from dMRI. Due to a lack of ground truth in human data, the model was validated measuring reproducibility in a scan-rescan dataset comprising 12 subjects. ACC was computed between two FODs obtained from two different scanning sessions. In this work, we are interested in regressing M-CSD coefficients from 2TS-CSD coefficients, recognizing that even M-CSD coefficients have limitations for both angular and spatial resolution of WM fibers [4]. Hence, we measured ACC between the regressed coefficients and the "ground truth" M-CSD coefficients. Direct comparisons between Nath et al. [35] and our work are not possible, as ACC between scan-rescan FODs should be higher than ACC between CSD coefficients of varying quality.

Lin et al. [33] used a 3D CNN to regress M-CSD coefficients from M-CSD coefficients computed on dMRI with downsampled gradient directions (similar to Experiment 3 – Downsampled Imaging Performance). Lin et al. [33] evaluated regression of M-CSD coefficients from S-CSD coefficients for one subject in the HCP dataset. For this subject,

ACC ranged from 0.95 to 0.96 in WM regions and 0.87 to 0.89 in partial volume regions when dMRI was downsampled to 45 gradient directions at b = 2000 s/mm^2 shell (50% of the original gradient directions). In comparison, our method had median ACC ranged from 0.92 to 0.96 in 45 patients when the S-CSD model was computed on 50% of the gradient directions at b = 2000 s/mm^2 shell. Hence, our CNNs have similar performance to Lin et al. [33].

Our approach may enable faster commercial dMRI acquisition with fewer gradient directions, thereby reducing acquisition times and facilitating translation in time-limited clinical environments [26]. Our work has the potential to estimate CSD coefficients from SS dMRI with similar quality to M-CSD coefficients. This method could be applied to improve the analysis of retrospective data where changing the acquisition is not possible.

There are two key limitations in this work. First, we used datasets to train our CNNs from the same scanner with the same acquisition protocol. Although we demonstrate that our approach is capable of generalizing across dMRI acquisition protocols (Experiment 2 – Inter-Scanner Acquisition Performance and Experiment 3 – Downsampled Imaging Performance), further improvements in the CNNs may be obtained by combining datasets during training. Secondly, we did not validate on subjects with pathologies that would distort WM tissue connectivity, such as tumors. Although the QS dataset [11] was acquired from patients with epilepsy, if any small lesions or abnormalities were present they were not big enough to distort normal WM anatomy. One future avenue of research is to evaluate this approach on data with pathologies that distort normal anatomy and validate it in a clinical setting, e.g. presurgical planning.

6. Conclusions

In this work, we presented a 3D patch-based convolutional neural network (CNN) to regress multi-shell, multi-tissue constrained spherical deconvolution (M-CSD) coefficients from single-shell 2-tissue CSD coefficients (2TS-CSD). Two CNN architectures, U-Net and HighResNet, were evaluated on their ability to regress CSD coefficients 1) on the same

dataset; 2) across different diffusion MRI (dMRI) acquisition protocols, and 3) on dMRI with fewer gradient directions than the training dataset. Our approach enables robust multi-tissue CSD FODs on SS dMRI acquisition protocols with few gradient directions, allowing faster dMRI acquisition in clinical settings. Further validation is required to demonstrate this approach generalizes to datasets acquired at multiple sites and on patients with brain pathologies that distort normal anatomy, such as tumors.

Data availability statement

HCP data is a publicly available dataset.[2] We used the 45 subjects in WU-Minn Retest set. **QS data** were taken from patients examined as part of routine clinical care. Informed consent was obtained from all individual participants included in the study in accordance with the ethical standards of the institutional and/or national research committee. Consent to publish this data was not obtained and will not be made public to protect patient privacy.

CRediT authorship contribution statement

Oeslle Lucena: Conceptualization, Methodology, Software, Data curation, Validation, Formal analysis, Investigation, Writing - original draft, Visualization. **Rachel Sparks:** Conceptualization, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Sjoerd B. Vos:** Resources, Writing - review & editing, Supervision. **Vejay Vakharia:** Resources, Writing - review & editing, Supervision. **Seb Ourselin:** Conceptualization, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Keyoumars Ashkan:** Conceptualization, Writing - review & editing, Supervision, Project administration, Funding acquisition. **John Duncan:** Resources, Writing - review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104643.

Ethical approval

All data were evaluated retrospectively. All studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

---

[2] https://db.humanconnectome.org.

References

[1] J. Tanner, Self diffusion of water in frog muscle, Biophys. J. 28 (1979) 107–116.

[2] J. Shapey, S. Vos, T. Vercauteren, R. Bradford, S.R. Saeed, S. Bisdas, S. Ourselin, Clinical applications for diffusion mri and tractography of cranial nerves within the posterior fossa: a systematic review, Front. Neurosci. 13 (2019) 23.

[3] P.J. Basser, J. Mattiello, D. LeBihan, Mr diffusion tensor spectroscopy and imaging, Biophys. J. 66 (1994) 259–267.

[4] B. Jeurissen, J.-D. Tournier, T. Dhollander, A. Connelly, J. Sijbers, Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion mri data, Neuroimage 103 (2014) 411–426.

[5] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, P.J. Basser, Axcaliber: a method for measuring axon diameter distribution from diffusion mri, Magn. Reson. Med.: An Official Journal of the International Society for Magnetic Resonance in Medicine 59 (2008) 1347–1354.

[6] D. Alexander, G. Barker, S. Arridge, Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data, Magn. Reson. Med.: An Official Journal of the International Society for Magnetic Resonance in Medicine 48 (2002) 331–340.

[7] J. Berman, Diffusion mr tractography as a tool for surgical planning, Magn. Reson. Imag. Clin. N. Am. 17 (2009) 205–214.

[8] B. Jeurissen, M. Descoteaux, S. Mori, A. Leemans, Diffusion mri fiber tractography of the brain, NMR Biomed. 32 (2019) e3785.

[9] G.P. Winston, P. Daga, M.J. White, C. Micallef, A. Miserocchi, L. Mancini, M. Modat, J. Stretton, M.K. Sidhu, M.R. Symms, et al., Preventing visual field deficits from neurosurgery, Neurology 83 (2014) 604–611.

[10] W.I. Essayed, F. Zhang, P. Unadkat, G.R. Cosgrove, A.J. Golby, L.J. O'Donnell, White matter tractography for neurosurgical planning: a topography-based review of the current state of the art, Neuroimage: Clinical 15 (2017) 659–672.

[11] M. Mancini, S.B. Vos, V.N. Vakharia, A.G. O'Keeffe, K. Trimmel, F. Barkhof, C. Dorfer, S. Soman, G.P. Winston, C. Wu, et al., Automated Fiber Tract Reconstruction for Surgery Planning: Extensive Validation in Language-Related White Matter Tracts, NeuroImage: Clinical, 2019, p. 101883.

[12] L.J. O'Donnell, Y. Suter, L. Rigolo, P. Kahali, F. Zhang, I. Norton, A. Albi, O. Olubiyi, A. Meola, W.I. Essayed, et al., Automated white matter fiber tract identification in patients with brain tumors, Neuroimage: Clinical 13 (2017) 138–153.

[13] K. Setsompop, R. Kimmlingen, E. Eberlein, T. Witzel, J. Cohen-Adad, J.A. McNab, B. Keil, M.D. Tisdall, P. Hoecht, P. Dietz, et al., Pushing the limits of in vivo diffusion mri for the human connectome project, Neuroimage 80 (2013) 220–233.

[14] J.-D. Tournier, F. Calamante, D.G. Gadian, A. Connelly, Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution, Neuroimage 23 (2004) 1176–1185.

[15] E.J. Canales-Rodríguez, J.H. Legarreta, M. Pizzolato, G. Rensonnet, G. Girard, J. Rafael-Patino, M. Barakovic, D. Romascano, Y. Alemán-Gómez, J. Radua, et al., Sparse wars: a survey and comparative study of spherical deconvolution algorithms for diffusion mri, Neuroimage 184 (2019) 140–160.

[16] F. Dell'Acqua, J.-D. Tournier, Modelling white matter with spherical deconvolution: how and why? NMR Biomed. (2018), e3945.

[17] C.-F. Westin, H. Knutsson, O. Pasternak, F. Szczepankiewicz, E. Özarslan, D. van Westen, C. Mattisson, M. Bogren, L.J. O'Donnell, M. Kubicki, et al., Q-space trajectory imaging for multidimensional diffusion mri of the human brain, Neuroimage 135 (2016) 345–362.

[18] J.-D. Tournier, F. Calamante, A. Connelly, Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution, Neuroimage 35 (2007) 1459–1472.

[19] T. Dhollander, A. Connelly, A novel iterative approach to reap the benefits of multi-tissue csd from just single-shell (+ b= 0) diffusion mri data, Proc. ISMRM 24 (2016) 3010.

[20] P.F. Neher, M.-A. Cote, J.-C. Houde, M. Descoteaux, K.H. Maier-Hein, Fiber tractography using machine learning, Neuroimage 158 (2017) 417–429.

[21] A. Daducci, E.J. Canales-Rodrı, M. Descoteaux, E. Garyfallidis, Y. Gur, Y.-C. Lin, M. Mani, S. Merlet, M. Paquette, A. Ramirez-Manzanares, et al., Quantitative comparison of reconstruction methods for intra-voxel fiber recovery from diffusion mri, IEEE Trans. Med. Imag. 33 (2013) 384–399.

[22] M. Descoteaux, High angular resolution diffusion imaging (hardi), Wiley Encyclopedia of Electrical and Electronics Engineering (1999) 1–25.

[23] S.B. Vos, M. Aksoy, Z. Han, S.J. Holdsworth, J. Maclaren, M.A. Viergever, A. Leemans, R. Bammer, Trade-off between angular and spatial resolutions in in vivo fiber tractography, Neuroimage 129 (2016) 117–132.

[24] D.K. Jones, T.R. Knösche, R. Turner, White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion mri, Neuroimage 73 (2013) 239–254.

[25] J.-D. Tournier, F. Calamante, A. Connelly, Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging, NMR Biomed. 26 (2013) 1775–1786.

[26] E.G. Ordóñez-Rubiano, F.A. Valderrama-Arias, J.A. Forbes, J.M. Johnson, I. Younus, J.H. Marín-Muñoz, M. Sánchez-Montaño, D.A. Angulo, H.A. Cifuentes-Lobelo, W. Cortes-Lozano, et al., Identification of preoperative language tracts for intrinsic frontotemporal diseases: a pilot reconstruction algorithm in a middle-income country, World neurosurgery 125 (2019) e729–e742.

[27] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.

[28] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.

[29] R. Tanno, D.E. Worrall, A. Ghosh, E. Kaden, S.N. Sotiropoulos, A. Criminisi, D. C. Alexander, Bayesian image quality transfer with cnns: exploring uncertainty in

dmri super-resolution, in: M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, S. Duchesne (Eds.), Medical Image Computing and Computer Assisted Intervention – MICCAI 2017, Springer International Publishing, Cham, 2017, pp. 611–619.

[30] V. Golkov, A. Dosovitskiy, J.I. Sperl, M.I. Menzel, M. Czisch, P. Sämann, T. Brox, D. Cremers, Q-space deep learning: twelve-fold shorter and model-free diffusion mri scans, IEEE Trans. Med. Imag. 35 (2016) 1344–1351.

[31] H. Zhang, T. Schneider, C.A. Wheeler-Kingshott, D.C. Alexander, Noddi: practical in vivo neurite orientation dispersion and density imaging of the human brain, Neuroimage 61 (2012) 1000–1016.

[32] H. Lu, J.H. Jensen, A. Ramani, J.A. Helpern, Three-dimensional characterization of non-Gaussian water diffusion in humans using diffusion kurtosis imaging, NMR in Biomedicine, An International Journal Devoted to the Development and Application of Magnetic Resonance In vivo 19 (2006) 236–247.

[33] Z. Lin, T. Gong, K. Wang, Z. Li, H. He, Q. Tong, F. Yu, J. Zhong, Fast learning of fiber orientation distribution function for mr tractography using convolutional neural network, Med. Phys. 46 (2019) 3101–3116.

[34] S. Koppers, C. Haarburger, D. Merhof, Diffusion mri signal augmentation: from single shell to multi shell with deep learning, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 61–70.

[35] V. Nath, K.G. Schilling, P. Parvathaneni, C.B. Hansen, A.E. Hainline, Y. Huo, J. A. Blaber, I. Lyu, V. Janve, Y. Gao, et al., Deep learning reveals untapped information for local white-matter fiber reconstruction in diffusion-weighted mri, Magn. Reson. Imag. 62 (2019) 220–227.

[36] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3d u-net: learning dense volumetric segmentation from sparse annotation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 424–432.

[37] E. Gibson, W. Li, C. Sudre, L. Fidon, D.I. Shakir, G. Wang, Z. Eaton-Rosen, R. Gray, T. Doel, Y. Hu, et al., Niftynet: a deep-learning platform for medical imaging, Comput. Methods Progr. Biomed. 158 (2018) 113–122.

[38] S.N. Sotiropoulos, S. Jbabdi, J. Xu, J.L. Andersson, S. Moeller, E.J. Auerbach, M. F. Glasser, M. Hernandez, G. Sapiro, M. Jenkinson, et al., Advances in diffusion MRI acquisition and processing in the human connectome project, Neuroimage 80 (2013) 125–143.

[39] D. Raffelt, T. Dhollander, J.-D. Tournier, R. Tabbara, R.E. Smith, E. Pierre, A. Connelly, Bias field correction and intensity normalisation for quantitative analysis of apparent fibre density, Proc. ISMRM 25 (2017) 3541.

[40] J.-D. Tournier, R.E. Smith, D.A. Raffelt, R. Tabbara, T. Dhollander, M. Pietsch, D. Christiaens, B. Jeurissen, C.-H. Yeh, A. Connelly, Mrtrix3: A Fast, Flexible and Open Software Framework for Medical Image Processing and Visualisation, bioRxiv, 2019, p. 551739.

[41] W. Li, G. Wang, L. Fidon, S. Ourselin, M.J. Cardoso, T. Vercauteren, On the compactness, efficiency, and representation of 3d convolutional networks: brain parcellation as a pretext task. International Conference on Information Processing in Medical Imaging, Springer, 2017, pp. 348–360.

[42] W. Luo, Y. Li, R. Urtasun, R. Zemel, Understanding the effective receptive field in deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2016, pp. 4898–4906.

[43] F. Milletari, N. Navab, S.-A. Ahmadi, V-net, Fully convolutional neural networks for volumetric medical image segmentation, in: 2016 Fourth International Conference on 3D Vision (3DV), IEEE, 2016, pp. 565–571.

[44] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: surpassing human-level performance on imagenet classification, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1026–1034.

[45] J. Wasserthal, P. Neher, K.H. Maier-Hein, Tractseg-fast and accurate white matter tract segmentation, Neuroimage 183 (2018) 239–253.

[46] A.W. Anderson, Measurement of fiber orientation distributions using high angular resolution diffusion imaging, Magn. Reson. Med.: An Official Journal of the International Society for Magnetic Resonance in Medicine 54 (2005) 1194–1206.

[47] M.J. Cardoso, M. Modat, R. Wolz, A. Melbourne, D. Cash, D. Rueckert, S. Ourselin, Geodesic information flows: spatially-variant graphs and their application to segmentation and fusion, IEEE Trans. Med. Imag. 34 (2015) 1976–1988.

[48] K.G. Schilling, V. Janve, Y. Gao, I. Stepniewska, B.A. Landman, A.W. Anderson, Histological validation of diffusion mri fiber orientation distributions and dispersion, Neuroimage 165 (2018) 200–221.

[49] A. Hammers, R. Allom, M.J. Koepp, S.L. Free, R. Myers, L. Lemieux, T.N. Mitchell, D.J. Brooks, J.S. Duncan, Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe, Hum. Brain Mapp. 19 (2003) 224–247.

[50] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., Pytorch: an imperative style, high-performance deep learning library, in: Advances in Neural Information Processing Systems, 2019, pp. 8024–8035.

[51] M.A. Wieczorek, M. Meschede, Shtools: tools for working with spherical harmonics, G-cubed 19 (2018) 2574–2592.