Using convolution neural networks to learn enhanced fiber orientation distribution models from commercially available diffusion magnetic resonance imaging

Oeslle Lucena\textsuperscript{a,}, Sjoerd B. Vos\textsuperscript{b,c}, Vejay Vakharia\textsuperscript{e}, John Duncan\textsuperscript{e,f}, Keyoumars Ashkan\textsuperscript{g}, Rachel Sparks\textsuperscript{a}, Sebastien Ourselin\textsuperscript{a}

\textsuperscript{a}School of Biomedical Engineering and Imaging Sciences, King’s College London, UK
\textsuperscript{b}Centre for Medical Image Computing, Department of Computer Sciences, University College London, London, UK
\textsuperscript{c}Neuroradiological Academic Unit, University College London Queen Square Institute of Neurology, University College London, London, UK
\textsuperscript{d}Epilepsy Society MRI Unit, Chalfont St Peter, UK
\textsuperscript{e}Department of Clinical and Experimental Epilepsy, University College London, UK
\textsuperscript{f}National Hospital for Neurology and Neurosurgery, Queen Square, UK
\textsuperscript{g}King’s College Hospital Foundation Trust, UK

Abstract

Accurate local fiber orientation distribution (FOD) modeling based on diffusion magnetic resonance imaging (dMRI) capable of resolving complex fiber configurations benefit from specific acquisition protocols that impose a high number of gradient directions (b-vecs), a high maximum b-value (b-vals) and multiple b-values (multi-shell). However, acquisition time is limited in a clinical setting and commercial scanners may not provide robust state-of-the-art dMRI sequences. Therefore, dMRI is often acquired as single-shell (SS) (single b-value). Here, we learn improved FODs for commercially acquired dMRI. We evaluate the use of 3D convolutional neural networks (CNNs) to regress multi-shell FOS representations from single-shell representations, using the spherical harmonics basis obtained from constrained spherical deconvolution (CSD) to model FODs. We use U-Net and High-ResNet 3D CNN architectures and data from the publicly available Human Connectome Dataset and a dataset acquired at National Hospital For Neurology and Neurosurgery Queen Square. We evaluate how well the CNN models can resolve local fiber orientation 1) when training and testing on datasets with same dMRI acquisition protocol; 2) when testing on dataset

\textsuperscript{*}Corresponding author
Email address: oeslle.lucena@kcl.ac.uk (Oeslle Lucena)
with a different dMRI acquisition protocol than used training the CNN models; and 3) when testing on datasets where a fewer number of dMRI gradient directions than used training the CNN models. Our approach may enable robust CSD model estimation on dMRI acquisition protocols which are single shell and with a few gradient directions, reducing acquisition times, and thus, facilitating translation to time-limited clinical environments.

Keywords: Diffusion Weighted Image, Deep Learning, Constrained spherical deconvolution, Tractography, CSD

1. Introduction

Diffusion-weighted MRI (dMRI) captures water molecule diffusion and can reveal underlying organizational aspects relating to different tissue components. In the brain, this is of particular interest to investigate the organization of the white matter (WM), which is composed of bundles of neuronal axons that impose a direction to water molecules diffusion Tanner (1979); Shapey et al. (2019), resulting in anisotropic diffusion with the preferred direction of diffusion being along the axon fiber Basser et al. (1994); Jeurissen et al. (2014).

From acquired dMRI signals, it is possible to non-invasively estimate WM tissue microstructure information, such as axon diameter Assaf et al. (2008), and compute local fiber orientation distributions (FOD or fODF) at the voxel level Alexander et al. (2002). FODs are used for fiber tractography Berman (2009); Jeurissen et al. (2019) which has an important role in presurgical planning Winston et al. (2014); Essayed et al. (2017); Mancini et al. (2019); ODonnell et al. (2017) and connectome analyses Setsompop et al. (2013). A common method to estimate local fiber orientation is diffusion tensor imaging (DTI) Basser et al. (1994). However, DTI only models a single fiber population and it cannot resolve complex fiber configurations in the brain such as crossing fibers Alexander et al. (2002). To address this issue, more robust methods for representing FODs have been presented that can resolve fiber crossing based on spherical harmonics (SH) Canales-Rodríguez et al. (2019); Dell’Acqua and Tournier (2018) or other approaches such as Q-ball imaging Tuch (2004), PAS-MRI Jansons and Alexander (2003) and the ball-and-sticks model Behrens et al. (2007).

Constrained spherical deconvolution (CSD) is one method based on SH capable of modeling complex fiber configurations Tournier et al. (2007); Jeurissen et al. (2014); Dhollander and Connelly (2016). Single-shell single-
tissue CSD (S-CSD) models each voxel as a single compartment with one corresponding FOD, irrespective of any underlying tissue components Tournier et al. (2007). However, distorted FODs are found when multiple tissues are present in a voxel due to the partial volume effect (PVE). Multi-shell multi-tissue CSD (M-CSD) extends S-CSD by modeling one anisotropic compartment (corresponding to WM) and two isotropic compartments (corresponding to GM and CSF). M-CSD provides a more reliable and accurate estimation of the WM FOD Jeurissen et al. (2014). M-CSD makes use of different attenuation levels across multiple b-values (or shells) to separate the voxel into three components. Nevertheless, a limitation with this approach is that it requires multi-shell (MS) dMRI which results in longer acquisition times compared to DTI and SS dMRI.

Alternatively Dhollander and Connelly (2016), single-shell 2-tissue (2TS-CSD) and single-shell 3-tissue CSD (SS3T-CSD) can overcome PVE by modeling isotropic compartments similar to M-CSD using the $b=0 \text{ s/mm}^2$ as a second shell. Within the additional shell, a multi-tissue signal profile is computed by 2TS-CSD and one isotropic (either GM or CSF) compartment is computed. SS3T-CSD uses an iterative approach to fit a CSD model for the three tissues compartments. Compared to M-CSD, 2TS-CSD can only model one isotropic tissue. While SS3T-CSD is more robust than S-CSD or 2TS-CSD, it is constrained to a high SH order fit ($l_{max} = 8$) requiring an acquisition with a high number of gradient acquisitions. Therefore, improving FOD modeling for commercially available dMRI acquisitions is an active topic of research.

Usually, accurate FOD models that are able to resolve complex fiber configurations require specific dMRI acquisition protocols with a high number of gradient directions (b-vecs), a high maximum b-value and/or multiple b-values Neher et al. (2017); Daducci et al. (2013); Descoteaux (1999); Vos et al. (2016). A higher number of gradient directions and high b-values both improve FOD angular resolution, enabling the model to better distinguish between complex fiber configurations, such as fiber crossings Jones et al. (2013); Tournier et al. (2013). Multiple b-values allows for multiple compartment modeling, correcting for PVE Jeurissen et al. (2014). Additionally, local fiber reconstruction is more accurate for images with a high signal-to-noise ratio Tournier et al. (2013). All of these constraints for dMRI acquisition impose longer acquisition times. Despite these advantages, clinical uptake has been limited due to the longer acquisition times and the need of expert staff to set up the acquisition parameters Ordóñez-Rubiano et al. (2019).

Deep learning (DL) has been successfully implemented for a variety
of medical imaging tasks Litjens et al. (2017). DL methods learn an underlying mathematical representation that can non-linearly map data from one representation to another representation, for instance mapping from raw image intensity to a predicted class or another intensity space LeCun et al. (2015). DL has been successfully applied to dMRI through image quality transfer (IQT) to improve its spatial resolution Alexander et al. (2017); Tanno et al. (2017), fit neurite orientation dispersion and density imaging (NODDI) Zhang et al. (2012) and spherical meaning technique (SMT) Kaden et al. (2016) models from the q-space Alexander et al. (2017); Golkov et al. (2016), and improve local fiber orientation model fitting Koppers et al. (2016); Nath et al. (2019).

In Alexander et al. (2017), IQT was proposed with linear regression and random forest models to (a) infer high-resolution dMRI patches from a lower spatial resolution and (b) to learn a mapping between parameters of different models. dMRI parameter mapping was evaluated to go from a low order DTI model to a higher order model, NODDI and SMT were both evaluated. Further work around this idea has evaluated IQT using a CNN patch-based regression to infer higher resolution patches and to quantify uncertainty for the regression Tanno et al. (2017). Fitting between models has been also proposed to compute diffusion kurtosis imaging Lu et al. (2006) and NODDI from shorter q-space MS dMRI signal intensities using a multilayer perceptron network (MLP) Golkov et al. (2016).

In Koppers et al. (2016), a MLP was trained to infer SH coefficients across different dMRI shells. Here, a MLP network used SH model coefficients from the same order calculated from one shell or combination of shells to infer the SH coefficients for a different shell with the restriction that all shells have the same number of dMRI gradient directions. One limitation of their approach is that the mapping only uses voxel information and does not take into account neighborhood information, which may help provide important spatial context. Another limitation is that this approach did not evaluated for generalization on data of different acquisition protocols than the training dataset.

In Nath et al. (2019), a neural network composed of regular hidden and residual layers (ResDNN) was trained for mapping S-CSD to FODs computed from histology. For the training dataset data was acquired from ex-vivo histology images from macaques using 3D structure tensor analysis Schilling et al. (2016). Angular correlation coefficient (ACC) was used to evaluate FOD accuracy. ResDNN outperformed FOD methods from single-shell dMRI such as S-CSD and Q-ball imaging when evaluating the ACC between the dMRI FOD and the FOD derived from the histology imaging.
ResDNN was also evaluated for reproducibility on 12 paired in-vivo dMRI obtained from the human connectome project (HCP). For each patient, two dMRI scans were acquired and ACC was computed between FODs obtained from ResDNN and between FODs obtained from the S-CSD. The major disadvantage of this work is that the model is trained on Macaque imagery and then transferred to human imaging. Nonetheless, as there is no ground truth or histology FOD estimates for the human imaging was against not clear ResDNN improves the FOD calculation. Furthermore, the baseline comparison used for the human imaging was S-CSD which is suboptimal as several more robust approaches exist - i.e. M-CSD.

In this work, we aim to compute a more accurate and reliable FOD using data that are still the most common in clinical settings: single-shell dMRI acquisitions. To achieve this aim, we present a framework to train a CNN to learn how to regress M-CSD model coefficients from 2TS-CSD model coefficients using a patch-based approach. We evaluate two different 3D CNN architectures and used two datasets. To evaluate our models, an extensive evaluation is performed as follows: (1) training and testing on a dataset with the same dMRI acquisition protocol, (2) testing on a dataset with different dMRI acquisition protocol than the training dataset, and (3) testing on dMRI that have fewer gradient directions than in the training dataset where test data can have a) the same or b) different dMRI acquisition protocol than the training dataset.

2. Methods

2.1. Pipeline overview

Our pipeline consists of the following steps. First, we preprocess all the MS dMRI data to correct for signal drift, geometric distortions and eddy-current induced distortions. We used data from the publicly available Human Connectome Dataset (HCP) Sotiropoulos et al. (2013) and an in-house dataset which we refer to as QS Dataset (see dataset details in Section 3.1). Secondly, we construct a paired dataset composed of SS dMRI and the original MS dMRI (Section 2.2). From this data, we compute the CSD models 2TS-CSD and M-CSD (Section 2.3) to SS dMRI and MS dMRI, respectively. Finally, we train a CNN model to regress the M-CSD model coefficients from 2TS-CSD model coefficients (Section 2.3) using the paired dataset.
2.2. Training dataset

From the preprocessed MS dMRI, a paired SS dMRI was constructed by selecting an appropriate shell from the MS dMRI based on the compromise of a minimum number of gradient directions for a given b-value to best characterize the angular frequency components of the dMRI signal Tournier et al. (2013).

Although \( l_{\text{max}} = 6 \) can provide an FOD with a higher angular contrast for high b-values \((b = 3000 \text{ s/mm}^2)\) Tournier et al. (2013), we aim to use a lowest feasible \( l_{\text{max}} \) order to ensure maximum applicability in clinical settings. Therefore, we used a \( l_{\text{max}} = 4 \), comprising 15 coefficients, as the order of our CSD modeling.

As a result, for the HCP dataset we constructed SS dMRI for all 3 b-values (1000, 2000, 3000 s/mm\(^2\)) where each shell has 90 directions. For the QS we constructed paired SS dMRI for 2 b-values (700, 2500 s/mm\(^2\)) with 32 and 64 directions for each shell respectively. For more details on the datasets see Section 3.1.

2.3. CSD modeling

CSD models the FOD as SH components and applies a nonnegativity constraint as a soft regularizer using a linear least-squares fit Dell’Acqua and Tournier (2018). The original dMRI signal intensities is approximated by a convolution of the FOD model with a signal attenuation profile for a single fiber population, referred to as response function Dell’Acqua and Tournier (2018); Tournier et al. (2007).

After the CSD model fitting, we applied a multi-tissue informed log-domain intensity normalization Tournier et al. (2019) to the M-CSD and the 2TS-CSD model coefficients to correct for the effects of (residual) intensity inhomogeneities.

2.3.1. M-CSD Modeling

M-CSD is computed from using a least squares fit of the dMRI signal intensities following the equation Jeurissen et al. (2014):

\[
\begin{bmatrix}
\hat{x}_1 \\
\vdots \\
\hat{x}_n
\end{bmatrix} = \underset{x}{\arg\min} \left\| \begin{bmatrix}
C_{1,1} & \cdots & C_{1,n} \\
\vdots & \ddots & \vdots \\
C_{m,1} & \cdots & C_{m,n}
\end{bmatrix} \begin{bmatrix}
x_1 \\
\vdots \\
x_n
\end{bmatrix} - \begin{bmatrix}
d_1 \\
\vdots \\
d_m
\end{bmatrix} \right\|^2_2
\]

subject to

\[
\begin{bmatrix}
A_1 & 0 & 0 \\
0 & \ddots & 0 \\
0 & 0 & A_n
\end{bmatrix} \begin{bmatrix}
x_1 \\
\vdots \\
x_n
\end{bmatrix} \geq \mathbf{0}
\]
where $\mathbf{d}_i$ is the vector of dMRI signal intensities on the $i$-th shell, $\mathbf{x}_j$ is the unknown vector of coefficients of the FOD of tissue $j$, $C_{i,j}$ is the matrix relating the coefficients of 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 opposed on the linear fit, where $\mathbf{A}_j$ is the tissue specific matrix relating the coefficients of the FOD of tissue $j$ to their amplitudes, effectively imposing positivity on each FOD. To perform the optimization, we adopted the original CSD optimization algorithm available from MRtrix Tournier et al. (2019).

2.3.2. 2TS-CSD modeling

For the 2TS-CSD model Dhollander and Connelly (2016), a similar approach is used where the $b = 0 \text{s/mm}^2$ (b-zero) is used as a second shell. To ensure Equation 1 has a unique solution, we set $j = 2$, reducing the number of response functions 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 Dhollander and Connelly (2016).

2.4. CNN training

Two CNNs were chosen for model evaluation, a 3D High-Resolution Network (HighResNet) Li et al. (2017) and a 3D U-Net Çiçek et al. (2016). Figure 1 shows a graphical representation of the network architectures. For both models, a patch-based training was used where it is necessary to reduce the effective receptive field (ERF) for both networks to reflect the selected patch size (Section 2.5) to avoid distortions at the patch boundary voxels Luo et al. (2016).

2.4.1. HighResNet

The original HighResNet architecture comprises of 3 levels of dilated convolutions and 9 residual connections resulting in 0.81M trainable parameters. HighResNet was originally proposed as a compact network that could achieve large ERFs Luo et al. (2016) without requiring a downsample-upsample pathway to capture low and high level features Çiçek et al. (2016); Milletari et al. (2016). Dilated convolutions are used to produce accurate predictions and detailed probabilistic maps alongside object boundaries Li et al. (2017). The final HighResNet architecture in this paper has the number of layers modified to reduce the ERF. It comprises of 2 levels of dilated
convolutions and 4 residual connections resulting in 0.16M trainable parameters. A parametric rectified linear unit (PReLU) activation function was used in place of a ReLU as PReLU adaptively learns the rectifier parameters and this property has been shown to improve CNNs accuracy in other applications He et al. (2015).

2.4.2. U-Net

The original 3D U-Net is a U-shaped network that has downsample-upsample pathway composed of 14 convolutional layers Çiçek et al. (2016) resulting in 19.08M trainable parameters. To achieve a comparable ERF as the patch size used for training, we adapted this architecture such that the final U-Net used in this work is composed of 10 convolutional layers resulting in 3.93M 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)) to reduce the ERF.

2.4.3. Data augmentation

Classical techniques for on-the-fly augmentation includes axis flipping, scaling, and rotation have been successfully applied to DL training in small 3D medical imaging datasets Wasserthal et al. (2018); Li et al. (2017); Gibson et al. (2018). Nonetheless, directly applying these techniques to SH coefficients damages the integrity of FODs, since the transformations are not in the SH harmonics domain, i.e. rotating spatially an image will not rotate the SH components appropriately resulting in unrealistic FODs. Therefore, we focused on applying a 3D random rotation in the SH domain to augment
our dataset as in Nath et al. (2019). Rotations were applied to 2TS-CSD and M-CSD coefficients to preserve FOD structure and the relationships between neighboring CSD coefficients.

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 network CSD coefficients output, as measured by $\text{loss}(y, \hat{y}) = \frac{\|y - \hat{y}\|^2}{2}$ where $y$ is the ground truth (M-CSD) and $\hat{y}$ is the output coefficients inferred from the network. Each CNN is initialized using He uniform function He et al. (2015) and trained for 400 epochs, based on experimentally chosen convergence, with a weight decay of $1E-6$. Training started with a learning rate of $3E-2$, which was then reduced by $1/2$ every 50 epochs. The entire network was trained using patches sampled from the intracranial space. To achieved that, it was used a binary skull-stripped mask as a prior to provide an intensity-based likelihood for the patch sampling. The binary skull-stripped mask was computed using MRtrix skull-stripping Tournier et al. (2019).

For each iteration in an epoch, a subject from the training set is randomly selected. Subsequently, the data is augmented by applying a random FOD rotation in the range of $[-25, 25]$ degrees to the subject. As a next step, 40 patches of size $32 \times 32 \times 32 \times 15$ were randomly sampled from the intracranial space, where 15 is the number of 2TS-CSD coefficients. The number of patches were experimentally selected to achieve the lowest validation loss while being able to be loaded on the available graphics process unit memory. An epoch is finished when all subject data from the training set have been selected to optimize the loss function. For every epoch, a new set of random augmentations and patches are computed from each subject data.

3. Experimental Design

3.1. Datasets

In this work, we used two datasets to conduct our analysis: the publicly available HCP Sotiropoulos et al. (2013) and a dataset acquired at National Hospital For Neurology and Neurosurgery Queen Square (QS) Dataset. All dMRI from QS were corrected for signal drift, geometric distortions and eddy-current induced distortions as in Mancini et al. (2019). The data from the HCP dataset was corrected following the protocols described in Sotiropoulos et al. (2013) prior to download. More detailed information about each dataset is found below.
3.1.1. QS dataset

The QS dataset is composed of 50 volumetric MS dMRI scans acquired from patients with epilepsy who appeared "structurally normal" on a T1-weighted MRI (T1). Small lesions are still found but only focally in the GM and hence not distorting or affecting the WM FOD estimation. All patients underwent MRI as part of routine clinical procedures acquired on a 3T GE MR750 that included 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/mm\(^2\), respectively and single \( b = 0 \) s/mm\(^2\) with reverse phase-encoding.

3.1.2. HCP dataset

We used a subset of HCP data composed of 45 subjects Sotiropoulos et al. (2013). These MS data was acquired in a 3T scanner with the following parameters: 1.25 mm isotropic resolution with 90 gradient directions for each \( b = \{1000, 2000, 3000 \) s/mm\(^2\}\) and 18 images at \( b = 0 \) s/mm\(^2\).

3.2. Evaluation Metrics

To compare the accuracy of CSD models we compute the mean absolute error (MAE) \( \text{MAE}(y, \hat{y}) = \frac{|y - \hat{y}|}{n} \) and the angular correlation coefficient (ACC) Anderson (2005) between the output inferred from a trained network an the M-CSD coefficients for voxels in the WM. The WM binary mask was computed using geodesic information flows as a segmentation tool Cardoso et al. (2015). ACC is a similarity metric computed at voxel level between two different sets of SH coefficients \( u \) and \( v \) of the same order, where \( j \) is the SH order. ACC is computed as:

\[
\text{ACC}(u, v) = \frac{\sum_{j=1}^{\infty} \sum_{m=-j}^{j} u_{j,m} v_{j,m}^*}{\left[ \sum_{j=1}^{\infty} \sum_{m=-j}^{j} u_{j,m}^2 \right]^{0.5} \left[ \sum_{j=1}^{\infty} \sum_{m=-j}^{j} v_{j,m}^2 \right]^{0.5}} \quad (2)
\]

ACC has a scale in the interval \([-1, 1]\), where 1 implies a perfect linear correlation between two functions on a sphere, whereas a negative -1 would imply a negative correlation Schilling et al. (2018). The value of -1 will not be reach because of the antipodal symmetry of the even order SH coefficients.

3.3. Experiments

We assessed the performance of the CNN models in three ways: 1) when training and testing on datasets with same dMRI acquisition protocol; 2)
when testing on dataset with a different dMRI acquisition protocol than used when training the CNN models; and 3) testing on datasets where a fewer number dMRI gradient directions than used when training the CNN models. The details for each experiment are described below.

3.4. Experiment 1

We assessed how well CNN models were able to regress M-CSD model coefficients when using the same acquisition protocol for training and testing. Model performance is assessed against M-CSD, the ground truth, for both QS and HCP datasets (Section 2.2). In this experiment, a 5-fold cross-validation was used where 3 folds were used for training, 1 fold for validation, and 1 fold for testing.

3.5. Experiment 2

In this stage, we assessed the generalizability of the CNN models to regress dMRI obtained from a different acquisition protocol. Here, without further tuning, we used the models from Experiment 1 and tested on a dataset with a different dMRI acquisition protocol than used in training the CNN models. For example, a model trained on QS dataset, where the input is 2TS-CSD model coefficients obtained from $b=700 \text{ s/mm}^2$, is used to estimate M-CSD from the HCP dataset, where the input is 2TS-CSD model coefficients obtained from $b=2000 \text{ s/mm}^2$.

3.6. Experiment 3

In this stage, we assessed robustness of the CNN models under scenarios for SS dMRI gradients with a fewer gradient directions than used for training the CNN model. Models from Experiment 1, with no further tuning, are used to infer images with a fewer dMRI gradients directions than the original acquisition (see Section 3.1 for datasets acquisition details). Similar to Experiment 2, testing on dMRI data with a different acquisition protocol than used when training the CNN models was also performed. The test data had a fewer dMRI gradients directions than the original acquisition.

To construct an SS dMRI dataset with fewer gradient directions, we subsampled by a half the number gradient directions from the SS and MS dMRI (see Section 2.2 for training dataset). To achieve that, we first used the command `dirgen` from MRtrix that reorders a 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 Tournier et al. (2019). Then, we truncated by half the number of gradient directions for both $b=0 \text{ s/mm}^2$ and the shell selected to generate
the SS dMRI - b=700 or 2500 s/mm$^2$ for the QS dataset and b=1000 or 2000 or 3000 s/mm$^2$ for the HCP dataset. Finally, the last step consisted in computing 2TS-CSD for the subsampled SS dMRI.

3.7. Implementation

All experiments were performed on a workstation equipped with a CPU (Xeon® W-2123, 8 x 3.60 GHz; Intel) and 32 GB of memory and a GPU (GeForce Titan V; NVIDIA) with 12 GB of on-board memory. All code was implemented in python using PyTorch Paszke et al. (2019) for the networks training, NiftyNet Gibson et al. (2018) for data loading and sampling parts, and SHtools Wieczorek and Meschede (2018) for SH rotations as data augmentation. The code used for train the CNN models is available online$^1$.

4. Results

In this section, we present results from Experiments 1-3 that are detailed in sections 4.1- 4.3, respectively.

As a reference for Tables and Figures, a method with following acronym QS 700-HCP 2000 CNN U-Net means that the model was trained on the QS data where the input was the 2ST-CSD model coefficients for b=700 s/mm$^2$ and it was tested on HCP data 2TS-CSD model coefficients for b=2000 s/mm$^2$. A method called 2TS-CSD (QS 700) means that the baseline approach was computed on the QS data for b=700 s/mm$^2$.

4.1. Experiment 1

In this experiment, we evaluated how well the CNN models performed when testing on a dataset with the same dMRI acquisition as the training dataset compared to the baseline method 2TS-CSD. Table 1 reports the average MAE and ACC values, between the indicated CSD coefficients and the M-CSD coefficients. As shown in Table 1, the CSD coefficients from U-Net and HighResNet are the most similar to M-CSD while the baseline 2TS-CSD is the least similar. Pronounced ACC improvements are found on QS dataset (improvement 7% of ACC mean) compared to HCP dataset (improvement 3% of ACC mean) for the best methods.

Figure 2 shows the cumulative distribution function (CDF) of ACC values over all WM voxels. The CNN models have more skewed curves, demonstrating more voxels with a high level of agreement when compared to 2TS-CSD.

---

$^1$https://github.com/OeslleLucena/RegressionFOD
Figures 5 and Figures 6 show qualitative results of heatmaps for ACC over all WM voxels. As shown in Figures 5 and Figures 6, ACC heatmaps show a high similarity across the three tested methods in WM voxel far from boundaries, where PVE effects are most likely to be minimal, and low correlation in boundary voxels, which are most likely to suffer from PVE. As with the quantitative results, HCP dataset have distinctly fewer low ACC values compared to QS dataset.

Table 1: Overall MAE and ACC mean and median between M-CSD and CSD model coefficients for Experiment 1. The minimum MAE and maximum ACC mean and median appear in bold. The label Train indicates data used to train the CNNs and the Test indicates the dataset used for the inference. HR = HighResNet

| Train-Test       | Method                  | MAE mean (std) | ACC mean | ACC median |
|------------------|-------------------------|---------------|----------|------------|
|                  |                         | mean median   | mean     |            |
| QS 700-QS 700    | 2TS-CSD (QS 700)        | 0.427(0.042)  | 0.882(0.126) | 0.928      |
|                  | CNN HR                 | 0.210(0.059)  | 0.940(0.064) | 0.957      |
|                  | CNN U-Net              | **0.204(0.024)** | 0.939(0.092) | **0.973** |
| QS 2500-QS 2500  | 2TS-CSD (QS 2500)      | 0.326(0.035)  | 0.938(0.099) | 0.984      |
|                  | CNN HR                 | **0.138(0.016)** | **0.965(0.068)** | **0.988** |
|                  | CNN U-Net              | 0.140(0.015)  | 0.963(0.075) | **0.988** |
| HCP 1000-HCP 1000| 2TS-CSD (HCP 1000)     | 0.247(0.042)  | 0.964(0.058) | 0.991      |
|                  | CNN HR                 | 0.162(0.019)  | **0.982(0.029)** | **0.993** |
|                  | CNN U-Net              | **0.159(0.019)** | 0.982(0.029) | **0.993** |
| HCP 2000-HCP 2000| 2TS-CSD (HCP 2000)     | 0.223(0.045)  | 0.970(0.052) | 0.993      |
|                  | CNN HR                 | **0.137(0.020)** | **0.984(0.029)** | **0.994** |
|                  | CNN U-Net              | **0.147(0.019)** | 0.984(0.029) | **0.994** |
| HCP 3000-HCP 3000| 2TS-CSD (HCP 3000)     | 0.273(0.040)  | 0.961(0.066) | 0.987      |
|                  | CNN HR                 | **0.162(0.020)** | **0.980(0.034)** | **0.992** |
|                  | CNN U-Net              | **0.159(0.020)** | 0.980(0.033) | **0.992** |

Finally, Figures 7 and Figures 8 has a visual representation of FODs, represented by glyphs showing the direction and distribution of diffusion parameters per voxel. The glyphs show that both U-Net and HighResNet are better able to resolve multiple fiber populations, small rotations and scaling within select regions.

Due to similar performance on results for HCP for different b-values in Experiment 1 (Section 4.1), we conducted further analysis for Experiments 2 and 3 (Sections 4.2 and 4.3) for data acquired at b=2000 s/mm² only.
4.2. Experiment 2

In this experiment, we evaluated how well the CNN models performed when testing on dataset with a different dMRI acquisition protocol than used when training the CNN models compared to the baseline method 2TS-CSD. Table 2 reports the average MAE and ACC values. As shown in Table 2, both U-Net and HighResNet were quantitatively more similar in terms of ACC and MAE to M-CSD when compared to the baseline 2TS-CSD.

Figure 3 shows the CDF of ACC values over all WM voxels. Here, we see distinct improvements for the QS 700 data as input when compared to the other inputs (QS 2500, HCP 2000). Once again, the CNN models had the most skewed CDF curves showing higher correlation to M-CSD compared to 2TS-CSD.

Figures 5 and Figures 6 show qualitative results of heatmaps for ACC over all WM voxels. Although ACC heat maps show more errors in “pure” WM voxels on CNN outputs for tests on Experiment 2 when compared to Experiment 1, U-Net and HighResNet captured better finer details on fiber crossing area and boundary voxels compared to 2TS-CSD (Figures 5 and Figures 6).

Finally, Figures 7 and Figures 8 has a visual representation of FODs. As
Table 2: MAE and ACC mean and median between M-CSD and CSD model coefficients for Experiment 2. The minimum MAE and maximum ACC mean and median are in bold. The label Train indicates data used to train the CNNs and the Test indicates the dataset used for the inference. HR = HighResNet.

| Train-Test Method | Method | MAE mean (std) | ACC mean | median |
|-------------------|--------|----------------|----------|--------|
| -                 | 2TS-CSD (QS 700) | 0.427 (0.042) | 0.882 (0.126) | 0.928 |
| HCP 2000-QS 700   | CNN HR | 0.349 (0.052) | 0.927 (0.080) | 0.962 |
|                   | CNN U-Net | 0.345 (0.055) | 0.920 (0.093) | 0.956 |
| -                 | 2TS-CSD (QS 2500) | 0.326 (0.035) | 0.938 (0.099) | 0.984 |
| HCP 2000-QS 2500  | CNN HR | 0.196 (0.022) | 0.954 (0.072) | 0.980 |
|                   | CNN U-Net | 0.214 (0.019) | 0.949 (0.078) | 0.978 |
| -                 | 2TS-CSD (HCP 2000) | 0.223 (0.045) | 0.970 (0.052) | 0.993 |
| QS 700-HCP 2000   | CNN HR | 0.514 (0.048) | 0.970 (0.046) | 0.986 |
|                   | CNN U-Net | 0.372 (0.052) | 0.972 (0.039) | 0.985 |
| -                 | 2TS-CSD (QS 2500) | 0.185 (0.031) | 0.979 (0.036) | 0.992 |
| QS 2500-HCP 2000  | CNN HR | 0.180 (0.031) | 0.978 (0.039) | 0.993 |
|                   | CNN U-Net | 0.180 (0.031) | 0.978 (0.039) | 0.993 |

shown in Figure 7, tests on QS dataset result FOD glyphs show a higher qualitatively agreement towards single fiber populations compared to multiple fiber populations. Small FODs were output when CNN models trained on HCP dataset were tested on QS 700 data. However, as shown in Figure 8, when tested on HCP dataset, CNN trained models on QS 2500 dataset were capable to resolve fiber crossing. CNN models trained on the QS 700 dataset could not resolve fiber crossing properly but they were able to resolve FODs scaling and small rotations having a good similarity with M-CSD.

Figure 3: ACC Cumulative distribution functions (CDFs) for ACC for Experiment 2. The ideal case is a high peak centered in 1 corresponding to a perfect ACC match to M-CSD. HR = HighResNet.
4.3. Experiment 3

We evaluated how well the CNN models performed when testing on datasets where a fewer number dMRI gradient directions than used when training the CNN models compared to the baseline method 2TS-CSD. Table 3 reports the average MAE and ACC values. As shown in Table 3, the CNN models achieved higher ACC and lower MAE than 2TS-CSD and are more similar to M-CSD (skewer CDFs) (Figures 4).

Table 3: MAE and ACC mean, std and median between M-CSD and CSD model coefficients for Experiment 3. The minimum MAE and maximum ACC mean and median are in bold. The label Train indicates data used to train the CNNs and the Test indicates the dataset used for the inference. HR = HighResNet.

| Train-Test          | Method  | MAE      | ACC      |
|---------------------|---------|----------|----------|
|                     |         | mean(std)| mean(std)| median  |
| -                   | 2TS-CSD (QS 700) | 0.477(0.045) | 0.850(0.138) | 0.892    |
| QS 700-QS 700       | CNN HR  | 0.249(0.033) | 0.917(0.106) | 0.957    |
|                     | CNN U-Net | 0.243(0.030) | 0.902(0.110) | 0.943    |
| HCP 2000-QS 700     | CNN HR  | 0.385(0.058) | 0.902(0.110) | 0.943    |
|                     | CNN U-Net | 0.378(0.049) | 0.896(0.109) | 0.937    |
| -                   | 2TS-CSD (QS 2500) | 0.391(0.036) | 0.911(0.114) | 0.959    |
| QS 2500-QS 2500     | CNN HR  | 0.189(0.023) | 0.942(0.091) | 0.975    |
|                     | CNN U-Net | 0.193(0.022) | 0.940(0.094) | 0.974    |
| HCP 2000-QS 2500    | CNN HR  | 0.229(0.027) | 0.936(0.089) | 0.969    |
|                     | CNN U-Net | 0.242(0.021) | 0.935(0.091) | 0.970    |
| -                   | 2TS-CSD (HCP 2000) | 0.283(0.057) | 0.955(0.066) | 0.984    |
| HCP 2000-QS 2000    | CNN HR  | 0.175(0.032) | 0.977(0.037) | 0.990    |
|                     | CNN U-Net | 0.177(0.031) | 0.977(0.036) | 0.990    |
| QS 700-HCP 2000     | CNN HR  | 0.325(0.051) | 0.965(0.045) | 0.982    |
|                     | CNN U-Net | 0.369(0.052) | 0.966(0.045) | 0.981    |
| QS 2500-HCP 2000    | CNN HR  | 0.221(0.038) | 0.965(0.047) | 0.987    |
|                     | CNN U-Net | 0.220(0.039) | 0.968(0.049) | 0.987    |

Figures 5 and Figures 6 show qualitative results of heatmaps for ACC over all WM voxels. As shown in Figures 5 and Figures 6, when tested on same a dataset dMRI acquisition protocol with a fewer gradients than it was originally trained, ACC heat maps show higher correlation in “pure” WM voxels and lower correlation boundary voxels similar to Experiment 1. Similarly to Experiment 2, we found low ACC in “pure” WM voxels for CNN outputs tested on a dataset with dMRI acquisition protocol than when tested the on dataset with the same dMRI acquisition protocol used for training. However, our models were able to capture better finer details on fiber crossing areas and at boundary voxels when compared to 2TS-CSD.
Figure 4: ACC Cumulative distribution functions (CDFs) for ACC for from Experiment 3. The ideal case is a high peak centered in 1 corresponding to a perfect ACC match to M-CSD. HR = HighResNet.

Finally, Figures 7 and Figures 8 has a visual representation of FODs. We observe similar behavior to Experiment 2 - small FODs were output when CNN models trained on HCP datasets and tested on QS 700 data (Figures 7) and CNN trained models on QS dataset when tested on HCP dataset were capable to resolve fiber crossing (Figures 8).

5. Discussion

We evaluated CNN-based patch regression to estimate M-CSD model coefficients from 2TS-CSD model coefficients. Our results demonstrate quantitatively and qualitatively that our method can estimate M-CSD model coefficients from 2TS-CSD model coefficients on datasets with the same dMRI acquisition protocol as the training set (Experiment 1). The models are also able to generalize when applied to dMRI datasets acquired with a different protocol compared to the training dataset (Experiment 2). Finally, and most importantly, the method is robust to dMRI acquisition protocols with a fewer gradient directions than the training datasets used in this work (Experiment 3), indicating that these methods can be trained on high-quality data and used to improve FOD estimation in lower-quality dMRI data.

Overall, larger improvements, in terms of ACC, MAE and FOD estimation, were observed in a specialist-acquired clinical protocol (QS) compared to a high quality research protocol (HCP). The HCP dataset has a high spatial and angular resolution that allows 2TS-CSD to resolve complex fiber configurations and compute very accurate FODs compared to the QS dataset, in which individual shells may not have enough gradient directions to capture these subtle differences. Because of these factors the CNN regression may show greater improvements in CSD coefficients estimation.
when the 2TS-CSD coefficients are less accurate. Our approach may enable faster commercial dMRI acquisition with fewer gradient directions, thereby, reducing acquisition times and thus facilitating translation in time-limited clinical environments Ordóñez-Rubiano et al. (2019).

MS dMRI enables better modeling of FODs, especially with voxels where PVE may complicate estimating the model coefficients Jeurissen et al. (2014). Our work has the potential to estimate CSD model coefficients of similar quality to M-CSD coefficients from SS dMRI. One application of this method is to improve the analysis of retrospective data. Additionally, estimation of M-CSD coefficients from 2TS-CSD coefficients may benefit tractography Smith et al. (2012); Mancini et al. (2019).

In this work, we evaluated two common neural network architectures, U-Net Çiçek et al. (2016) and HighResNet Gibson et al. (2018). Both CNN models perform similarly throughout all experiments (Table 1–3). The aim of this work was not to find the best CNN to perform CSD coefficients regression but to show the capability of deep learning to facilitate enhanced FODs for commercially available dMRI acquisition protocols. One future avenue of research is to investigate how the different networks and trainable parameters influences the regression of CSD coefficients.

There are two key limitations in this work. First, we used datasets to train our CNN models using where the data was from the same scanner and had the same acquisition protocols. Although we demonstrate that our approaches are capable of generalizing across dRMI acquisition protocols (Experiments 2 and 3), further improvements in the regression model may be obtained by combining datasets across different sites for CNN models training. Secondly, we did not included a validation on subjects with pathologies that would distort WM tissue connectivity, such as tumors or edema. Although the QS dataset Mancini et al. (2019) was acquired from patients with epilepsy if any small lesions or abnormalities were present they were not big enough to distort normal anatomy.

6. Conclusions

In this work, we presented a 3D CNN to regress M-CSD model coefficients from 2T-CSD model coefficients. Two CNN model architectures, U-Net and HighResNet, were evaluated on their ability to resolve local fiber orientation distributions (FODs) 1) on the same dataset; 2) across different dMRI acquisition protocols and 3) on a dMRI with fewer gradient directions. Our approach may enable robust CSD model estimation on dMRI
acquisition protocols which are single shell and with a few gradient directions, resulting in a faster commercial dMRI acquisition. Future validation is required to demonstrate this approach generalization on training datasets acquired at multiple sites and on patients with brain pathologies that distort normal anatomy, such as tumors.

Data availability statement

HCP data is publicly available dataset\(^2\). We used the 45 subjects from WU-Minn Retest set. QS data which support the findings in this paper was 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.

Acknowledgments

This research was funded/supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy’s and St Thomas’ NHS Foundation Trust and King’s College London and/or the NIHR Clinical Research Facility. Oeslle Lucena is funded by EPSRC Research Council (EPSRC DTP EP/R513064/1). Sjoerd B. Vos is funded by the National Institute for Health Research University College London Hospitals

\(^2\)https://db.humanconnectome.org
Biomedical Research Centre (NIHR BRC UCLH/UCL High Impact Initiative BW.mn.BRC10269). We also thank NVIDIA for the Titan V GPU used in this work. The views expressed are those of the author(s) and not necessarily those of the NIHS, the NIHR or the Department of Health.

**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. **Conflict of interest.** The authors declare that they have no conflict of interest.
Figure 5: ACC heat maps overlaid on a T1 image for all experiments on one subject from QS datasets. HR = HighResNet. The ACC values are in JET colormap and shown in between $[-0.5, 1]$. Both U-Net and HighResNet have higher agreement to M-CSD when compared to the baseline 2TS-CSD.
Figure 6: ACC heat maps overlaid on a T1 for all experiments on one subject of HCP dataset. HR = HighResNet. The ACC values are in JET colormap and shown in between $[-0.5, 1]$. Both U-Net and HighResNet have higher agreement to M-CSD when compared to the baseline 2TS-CSD.
Figure 7: A glyph representation of the FOD for the regions indicated by the red (magnified by 3×) for one subject from the QS dataset. HR = HighResNet. The ACC values are in grayscale and shown in between [−0.5, 1]. Both U-Net and HighResNet show higher similarity to M-CSD when compared to the baseline 2TS-CSD.
Figure 8: A glyph representation of the FOD for the regions indicated by the red (magnified by 2×) for one subject from the HCP dataset. HR = HighResNet. The ACC values are in grayscale and shown in between $[-0.5, 1]$. Both U-Net and HighResNet show higher similarity to M-CSD when compared to the baseline 2TS-CSD.
References

J. Tanner, Self diffusion of water in frog muscle, Biophysical journal 28 (1979) 107–116.

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, Frontiers in neuroscience 13 (2019) 23.

P. J. Basser, J. Mattiello, D. LeBihan, Mr diffusion tensor spectroscopy and imaging, Biophysical journal 66 (1994) 259–267.

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.

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

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

J. Berman, Diffusion mr tractography as a tool for surgical planning, Magnetic resonance imaging clinics of North America 17 (2009) 205–214.

B. Jeurissen, M. Descoteaux, S. Mori, A. Leemans, Diffusion mri fiber tractography of the brain, NMR in Biomedicine 32 (2019) e3785.

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.

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.
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) 101883.

L. J. ODonnell, 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.

K. Setsompop, R. Kimmling, 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.

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.

F. Dell’Acqua, J.-D. Tournier, Modelling white matter with spherical deconvolution: How and why?, NMR in Biomedicine (2018) e3945.

D. S. Tuch, Q-ball imaging, Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine 52 (2004) 1358–1372.

K. M. Jansons, D. C. Alexander, Persistent angular structure: new insights from diffusion magnetic resonance imaging data, Inverse problems 19 (2003) 1031.

T. E. Behrens, H. J. Berg, S. Jbabdi, M. F. Rushworth, M. W. Woolrich, Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?, Neuroimage 34 (2007) 144–155.

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.

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, in: Proc ISMRM, volume 24, 2016, p. 3010.
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.

A. Daducci, E. J. Canales-Rodríguez, 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 transactions on medical imaging 33 (2013) 384–399.

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

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.

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.

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 in Biomedicine 26 (2013) 1775–1786.

E. G. Ordóñez-Rubiano, F. A. Valderrama-Arias, J. A. Forbes, J. M. Johnson, I. Younus, J. H. Marin-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.

G. Litjens, T. Kooi, B. E. Bejnordi, 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, Medical image analysis 42 (2017) 60–88.

Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (2015) 436–444.

D. C. Alexander, D. Zikic, A. Ghosh, R. Tanno, V. Wotschel, J. Zhang, E. Kaden, T. B. Dyrby, S. N. Sotiropoulos, H. Zhang, et al., Image quality transfer and applications in diffusion mri, NeuroImage 152 (2017) 283–298.
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.

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.

E. Kaden, F. Kruggel, D. C. Alexander, Quantitative mapping of the per-axon diffusion coefficients in brain white matter, Magnetic resonance in medicine 75 (2016) 1752–1763.

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 transactions on medical imaging 35 (2016) 1344–1351.

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.

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, Magnetic resonance imaging 62 (2019) 220–227.

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.

K. Schilling, V. Janve, Y. Gao, I. Stepniewska, B. A. Landman, A. W. Anderson, Comparison of 3d orientation distribution functions measured with confocal microscopy and diffusion mri, Neuroimage 129 (2016) 185–197.

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.

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) 551739.

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, in: International Conference on Information Processing in Medical Imaging, Springer, 2017, pp. 348–360.

Ö. Çiček, 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.

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.

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.

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.

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

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, Computer methods and programs in biomedicine 158 (2018) 113–122.

A. W. Anderson, Measurement of fiber orientation distributions using high angular resolution diffusion imaging, Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine 54 (2005) 1194–1206.
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 transactions on medical imaging 34 (2015) 1976–1988.

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.

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.

M. A. Wieczorek, M. Meschede, Shtools: Tools for working with spherical harmonics, Geochemistry, Geophysics, Geosystems 19 (2018) 2574–2592.

R. E. Smith, J.-D. Tournier, F. Calamante, A. Connelly, Anatomically-constrained tractography: improved diffusion mri streamlines tractography through effective use of anatomical information, Neuroimage 62 (2012) 1924–1938.