How Can Spherical CNNs Benefit ML-Based Diffusion MRI Parameter Estimation?

Tobias Goodwin-Allcock[1]([✉])[iD], Jason McEwen[2][iD], Robert Gray[3][iD],
Parashkev Nachev[3][iD], and Hui Zhang[1][iD]

[1] Department of Computer Science and Centre for Medical Image Computing,
University College London, London, UK
`tobias.goodwin-allcock@ucl.ac.uk`
[2] Kagenova Limited, Guildford, UK
[3] Department of Brain Repair and Rehabilitation, Institute of Neurology,
UCL, London, UK

Abstract. This paper demonstrates spherical convolutional neural networks (S-CNN) offer distinct advantages over conventional fully-connected networks (FCN) at estimating scalar parameters of tissue microstructure from diffusion MRI (dMRI). Such microstructure parameters are valuable for identifying pathology and quantifying its extent. However, current clinical practice commonly acquires dMRI data consisting of only 6 diffusion weighted images (DWIs), limiting the accuracy and precision of estimated microstructure indices. Machine learning (ML) has been proposed to address this challenge. However, existing ML-based methods are not robust to differing gradient schemes, nor are they rotation equivariant. Lack of robustness to differing gradient schemes requires a new network to be trained for each scheme, complicating the analysis of data from multiple sources. A possible consequence of the lack of rotational equivariance is that the training dataset must contain a diverse range of microstucture orientations. Here, we show spherical CNNs represent a compelling alternative that is robust to new gradient schemes as well as offering rotational equivariance. We show the latter can be leveraged to decrease the number of training datapoints required.

Keywords: Spherical CNNs · Machine learning · Diffusion MRI

1 Introduction

Diffusion MRI (dMRI) plays an important role in neuroscientific and clinical research because it can help infer tissue microstructure [3]. To infer tissue microstructure from dMRI, we use mathematical models to estimate parameters from dMRI data. At each voxel, the measurements made according to some

acquisition scheme - commonly consisting of gradient schemes coupled with their diffusion sensitising factor - are fitted to mathematical models, such as the diffusion tensor (DT) [2]. From these models, dMRI parameters can be derived to reveal the microstructure such as fractional anisotropy (FA), which characterises the anisotropy of the tissue. The computation of these parameters from dMRI data is known as dMRI parameter estimation, which has traditionally been achieved with model fitting. However, the fidelity of dMRI parameters estimated in this way is limited by relatively high noise in the data, requiring more measurements to be acquired than what are routinely made in the clinic [10].

As in many other fields, dMRI parameter estimation has recently been revolutionised by exploiting deep learning (DL), yielding greatly increased accuracy than conventional fitting when the acquisition scheme has a small number of samples [1,8]. However, current deep-learning methods, e.g. fully-connected networks (FCN), are ignorant of the acquisition scheme of a given acquisition, rendering these methods potentially not generalisable to new acquisition schemes. This complicates the application of a DL model to data acquired from multiple sources. Moreover, these methods do not exhibit rotational equivariance, a property that may help reduce the demand for training data.

There have been a number of attempts to capture the relationship between an acquisition scheme and the corresponding data [4,12]. However, they do not utilise the topological features of the associated gradient schemes. Gradient schemes for a given diffusion sensitising factor can be represented by points on the unit sphere. Therefore, spherical convolutional neural networks (S-CNN), recently proposed as an alternative to FCNs [7,13], provide a more natural solution to this problem. However, currently there exists no direct evidence of the theoretical benefits of S-CNNs, such as rotational equivariance and robustness to different gradient schemes. Here we aim to provide the very first empirical evidence of these advantages in the context of estimating rotation-invariant dMRI parameters.

The rest of the paper is described as follows: Sect. 2 how machine learning has been used to solve the dMRI parameter estimation problem and the theoretical beneficial properties of S-CNNs; Sect. 3 then goes on to explain how we empirically test these properties; Sect. 4 summarises the results and discusses future work.

2   ML Solutions to the dMRI Parameter Estimation Problem and the Theoretical Benefits of S-CNNs

Deep learning (DL) has been proposed as a solution to dMRI parameter estimation from small numbers of diffusion weighted images (DWI). This section provides (1) an example of the current machine learning standard for voxel-wise estimation (2) the theoretical limitations of this architecture (3) the architecture features that theoretically benefit S-CNNs. As an example, we show how the dMRI parameter FA is estimated from a common clinical diffusion MRI acquisition consisting of 6 DWIs.

Deep learning models, F, map the dMRI signals, $\mathbf{s}$, and their corresponding acquisition scheme - consisting of b-values, $\mathbf{B}$, and gradient scheme $\mathbf{G}$ - directly to dMRI parameters, denoted t.

$$t = F(\mathbf{s}, \mathbf{B}, \mathbf{G}; \theta) \tag{1}$$

This function is learnt by optimising the trainable parameters, θ, on training data. After training, the quality of the network estimation depends on many factors. The two factors explored in this work are the training data distribution and the choice of network architecture.

2.1 Fully-Connected Networks

The first and most common deep learning network architecture applied to dMRI data is the FCN [1,8]. Conventionally these have been implemented following:

$$t = F_{\text{FCN}}(\mathbf{s}; \theta_{FCN}), \tag{2}$$

where F_{FCN} is a fully-connected network with trainable parameters θ_{FCN}. The network's input consists of the dMRI signals. Absent from the equation is the acquisition scheme so the network is ignorant of the acquisition scheme. Estimation from a new set of DWIs is accurate only if the acquisition scheme for the new data is consistent with the acquisition scheme during training [12].

An FCN's architecture is not designed to be rotationally equivariant. A theoretical consequence of lacking rotational equivariance is that the training dataset may have to contain a diverse set of tissue microstructure orientations for FCNs to accurately estimate independent of fibre orientation.

2.2 Spherical CNNs

S-CNNs theoretically improve over FCNs - both in terms of robustness to the gradient scheme and robustness to the training data distribution - because of the difference in network architecture.

An S-CNN's architecture differs greatly to an FCN's but not in the way one may expect. In S-CNNs, the convolution isn't across multiple voxels, like traditional CNNs, but over the spherical image space. Therefore, S-CNNs are voxel wise networks just like FCNs. The spherical image is generated at each voxel from the dMRI signals, $\mathbf{s}$, along with their corresponding gradient scheme $\mathbf{G}$. We see that this architecture may naturally address the highlighted limitations of FCNs. Firstly, an S-CNN's input is informed of the gradient scheme as shown in the following equation:

$$t = F_{\text{S-CNN}}(\mathbf{s}, \mathbf{G}; \theta_{\text{S-CNN}}) \tag{3}$$

where $F_{\text{S-CNN}}$ is an S-CNN with trainable parameters $\theta_{\text{S-CNN}}$. We hypothesise this input will allow S-CNNs to be robust to a change in gradient scheme at inference time as long as the diffusion sensitising factors are the same.

Another benefit of S-CNN's is the rotationally equivariant architecture [6]. We hypothesise that this property will allow S-CNNs to extrapolate from a training dataset with a common primary fibre orientation and, during the inference stage, well estimate tissue with fibres oriented along any direction. As a result, S-CNNs do not require a diverse set of tissue microstructure orientations in the training dataset, reducing demands on the training dataset.

3   Experiments

Each claim made in this paper is evaluated with an individual experiment. The first experiment evaluates network robustness to differing gradient schemes; the second assesses network robustness to the distribution of the primary fibre orientations in the training set.
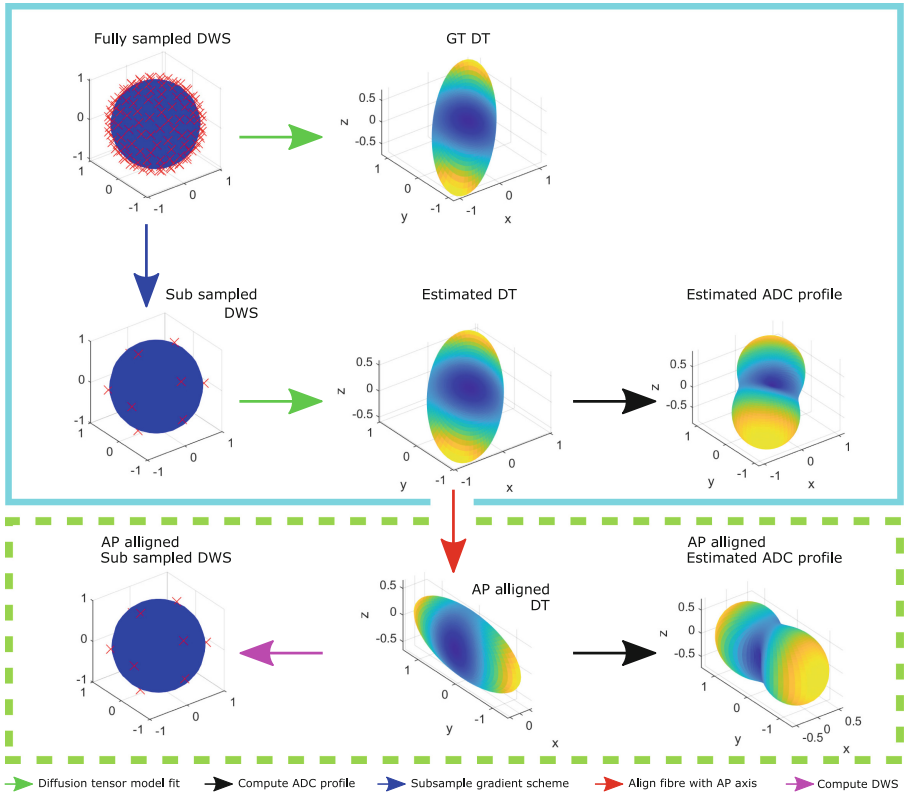
3.1   Experiment 1

Study Design. In this experiment we test if the networks are robust to new gradient schemes at inference time. In order to show this we propose an experiment where both networks are trained with a typical gradient scheme and then, in the inference phase, the trained networks are applied to data collected using a gradient scheme (1) the same as the training gradient scheme and another (2) different to the training scheme.

Network Architectures and Training Parameters. Both network architectures are voxel-wise networks. The S-CNN architecture we use, known as the hybrid spherical CNN architecture [5], was chosen as it has been shown to be highly rotationally equivariant whilst also being computationally efficient. The specific network parameters follow the spherical MNIST experiment and the input to this network is a densely sampled spherical signal, described later. The FCN network architecture implementation, used as the baseline, is consistent with the established implemented FCN techniques for dMRI parameter estimation [1,8]. This network input follows the standard practice consisting of 6 $b = 0$ normalised diffusion-weighted signals. The network hyperparameters are: three hidden layers with number of units $= [100,100,10]$ and ReLU activation function.

The training parameters are chosen in order for FCN to perform optimally and consistent between the networks so that any difference between trained networks is solely because of their architectures. To achieve this, the training parameters are consistent with the FCN literature; specifically the training regime used Adam optimiser for 50 epochs with learning rate set to 0.001, the batch size 32 and the loss metric MSE.

Generating Densely Sampled Spherical Signals. S-CNNs require densely sampled spherical signals as input. Densely sampled spherical signals are generated for each voxel by utilising the property of the 1-to-1 mapping between

six-directional dMRI signals and the 6 independent values of the diffusion tensor. Due to this property, six-directional dMRI signals, with all of their noise, are perfectly and uniquely described by a DT. From this DT a spherical function called the ADC profile may be derived and sampled to generate the input required for S-CNNs. This process is visually described in Fig. 1.

Fig. 1. This figure shows how the testing data is generated for the both experiments (blue box). The figure also shows how the training data is generated for experiment 1 (blue box) and experiment 2 (green box). In the blue box, from the 90 directional $b = 0$ normalised diffusion weighted signals (DWS) we estimate the ground truth diffusion tensor, and generate a 6-dir clinical scan by subsampling. The clinical scan's measurements are used as input to the FCN, but for compatibility with the S-CNN they must first be re-parameterised as an ADC profile. This is achieved by exploiting the 1-to-1 relationship described in Sect. 3.1. For the second experiment training data is required that has the primary fibre oriented along the AP axis. This is achieved by changing the orientation of the estimated DT and calculating the 6 DWS, for FCN, or calculating the ADC profile for S-CNN. (Color figure online)

Datasets. A dataset is required for training and testing the models. All of the deep learning models evaluated in this paper are supervised machine learning techniques, therefore, the training dataset must consist of a set of input values paired with ground truth output values. For the high-quality ground truth output, a dataset is required that contains a sufficiently large number of DWIs to provide accurate estimation of FA.

For this reason, we have chosen dMRI data from the Human Connectome Project (HCP) which includes 90 DWIs at $b = 1000 \, s/mm^2$. Ground-truth FA maps are computed from the complete set of DWIs. Subsets of six-directional DWIs are sampled from the 90 DWIs to be in maximal agreement with the chosen gradient schemes to mimic clinical scans. Although clinical scans are collected with a much lower spatial resolution than HCP data, for the purpose of this evaluation, HCP data can be argued to resemble typical clinical scans. As the networks we evaluate are voxel-wise, the main effect of a change in spatial resolution is a change in SNR. While HCP's increased resolution reduces the SNR, the state-of-the-art gradients of the scanner used to acquire HCP data increases the SNR. The resulting SNR is comparable with clinical data. Training was performed with one participant and, to show generalisation, data from 12 unseen participants was used for testing.

Evaluation. Both networks are trained using diffusion weighted signals corresponding to one gradient scheme, the Skare scheme [14], and evaluated with signals corresponding to both the original Skare scheme as well as a new scheme, Jones [9]. Quantitative measurements of the estimation error are calculated with root mean square error (RMSE) over specific FA ranges over the whole image. Statistical significance between distributions is quantified with paired t-tests. Qualitative assessments are made using maps of estimates and errors relative to the ground truth.

3.2 Experiment 2

Study Design. In this experiment we test the network's robustness to the distribution of the primary fibre orientations in the training set by testing the networks on a set of microstructure configurations whose primary fibre orientation lies both inside and outside of the training dataset's distribution. We achieve this by restricting the primary fibre orientation in the training dataset to align with the anterior-posterior axis and testing on microstructure oriented in all directions. We hypothesise that networks robust to the distribution of primary fibre orientation in the training set will estimate the FA equally well independent of the direction of the primary fibre orientation.

Restricting Primary Fibre Orientation. The training distribution of the primary fibre orientations is restricted to the anterior-posterior axis by adapting the dense ADC sampling algorithm. In the dense ADC sampling algorithm, after a noisy DT's estimation, the noisy DT's shape and size are extracted by

eigendecomposition. Next, the primary fibre orientation, otherwise known as the principal eigenvector, is set to the anterior-posterior axis whilst the secondary eigenvector is set to the superior-inferior axis; see visual description in the light green dashed box in Fig. 1. This new AP-aligned diffusion tensor is used to generate the input required for FCN, by computing the 6 directional $b = 0$ normalised diffusion-weighted signals from the DT forward model, and the input required for S-CNN, by densely sampling the ADC profile.

Evaluation. Network architectures, training scheme, training dataset and evaluation metrics are the same as experiment 1. Only one gradient scheme is required for this experiment so the Skare scheme is chosen for training and testing. Training data undergoes primary fibre orientation restriction, whilst testing data is unrestricted. To show further benefits of rotational equivariance we test to see if this property allows S-CNNs to estimate with high fidelity when starved of training data points. For this, an S-CNN is trained with only 10 % of the total training datapoints. The distribution of estimation error over the primary fibre orientation is evaluated.

4 Results and Discussion

4.1 For Experiment 1

Qualitative and quantitative results of experiment 1 are shown in Figs. 2 and 3 respectively.

Figure 2 shows an example slice of a GT FA map from a test subject along with the model fitting, FCN and S-CNN estimations and error maps when the gradient scheme is the same or different between training and testing. When the training and testing schemes are the same FCN's performance is consistent with the literature, estimating FA faithfully. When the gradient schemes are different the FCN estimates poorly. Estimation is especially poor in areas of high FA, such as the corpus callosum. S-CNN's estimations are similar to the GT regardless of the gradient scheme. This shows S-CNN's robustness to differing gradient schemes.

Figure 3 reinforces the qualitative observations with quantitative measures showing boxplots of the mean RMSE over the 12 subjects. We see for anisotropic signals (FA > 0.4) the conventional FCN performance is significantly worse (p $< 8e{-}12$) when applied to a new gradient scheme than on the training scheme, whereas, the S-CNN models estimation fidelity does not decrease when applied to a new scheme. In both figures, on differing gradient schemes the FCN model is shown to be more inaccurate in regions of high FA. This loss of accuracy in high FA regions may be caused by the signal attenuation in these regions being greatly dependant on the direction of measurement. Therefore, for low FA voxels the DWS from any two gradient schemes are similar. Whereas, for areas of high FA the DWS will be vastly different and therefore the training distribution will be different to the testing distribution.
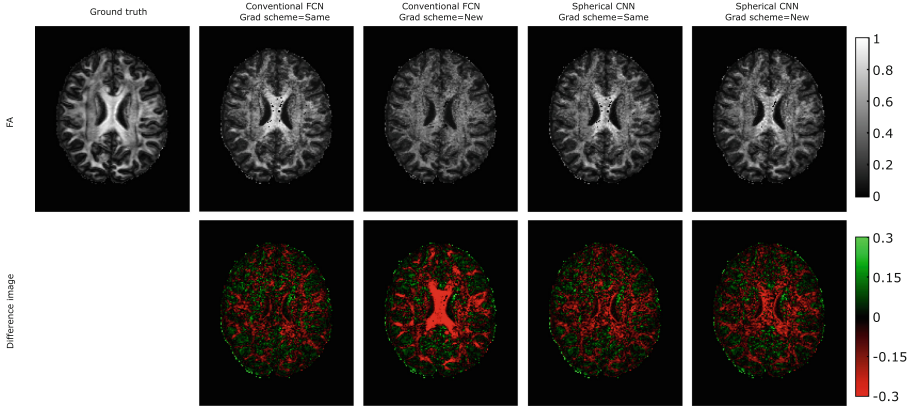
Fig. 2. Results of experiment 1 are visualised on an subject unseen during training. Conventional fully-connected networks are compared against the S-CNN model at estimating FA with either the same gradient scheme as used in training or a new gradient scheme. The FCN shows a drop in performance when the gradient scheme is different between training and testing. The S-CNN doesn't have this issue, therefore, it is robust to differing gradient schemes.

Gradient scheme robustness may also be achieved with FCN by first encoding the signals in terms of their corresponding spherical harmonic representation [11]. This approach is essentially equivalent to S-CNN but does not offer the property of rotational equivariance, the effect of which is demonstrated in the next subsection.

4.2   For Experiment 2

The results of experiment 2 are shown in Figs. 3, 4 and 5. Figure 4 qualitatively shows the effect of estimating the full brain volume using networks trained only on tissue microstructure aligned with the anterior-posterior axis. The FCN model consistently underestimates FA in regions where the underlying tissue microstructure does not align with anterior-posterior direction (e.g. the corpus callosum which consists of left-right white matter tracts). In contrast, the S-CNN models accurately estimate FA independently of the primary fibre direction, and the error is far less structured than the FCN's.

This is mirrored in the quantitative measurements over the 12 testing subjects, shown in Fig. 3. The FCN model only well estimates isotropic signals (FA < 0.2); however the performance difference between different methods is relatively small, with the difference between the mean RMSE for different methods no larger than 0.04. When the signal becomes anisotropic, the FCN model estimates the signal significantly worse than the S-CNN models, with the difference in the mean RMSE for different methods no smaller than 0.08. As the signals get more anisotropic the FCN model performs even worse until in the top FA bracket

$(0.8 \leq \mathrm{FA} < 1)$ the mean RMSE between the FCN and the S-CNN models is 0.5, 10 times the difference in error of the isotropic signals.

Figure 5 shows the distribution of the absolute error over the full range of primary fibre orientations for anisotropic signals. The FCN model well estimates tissue microstructure aligned with the anterior-posterior axis, seen during training. However the error quickly grows as the primary fibre orientation deviates from this axis. This adverse feature is not exhibited by the S-CNN model as the estimation error is low and independent of training dataset distribution of the primary fibre orientation. The lack of rotational equivariance in FCNs hinders estimation performance when generalising to microstructure with primary fibre orientation not sampled in the training distribution. This has potential for orien-
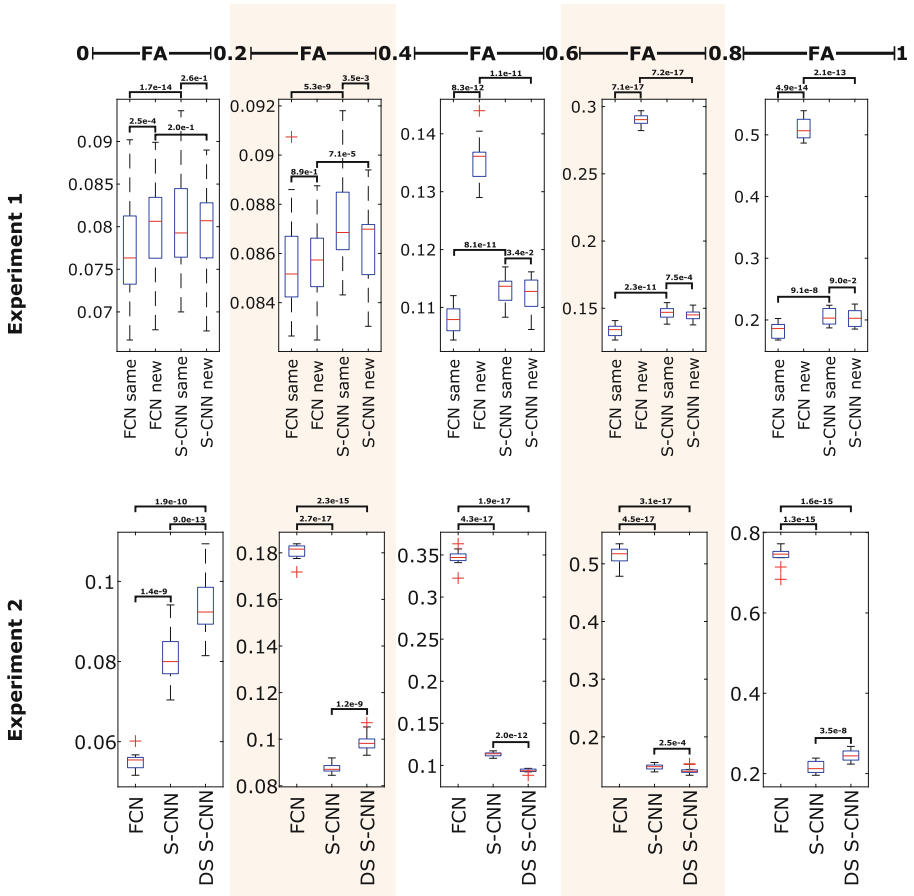
Fig. 3. Boxplots of the RMSE over the 12 testing subjects for both experiments, with p-values between distributions from paired t-test shown. The testing data is uniformly split into 5 GT FA ranges. For both experiments the problems with FCN become apparent when applied to anisotropic signal.
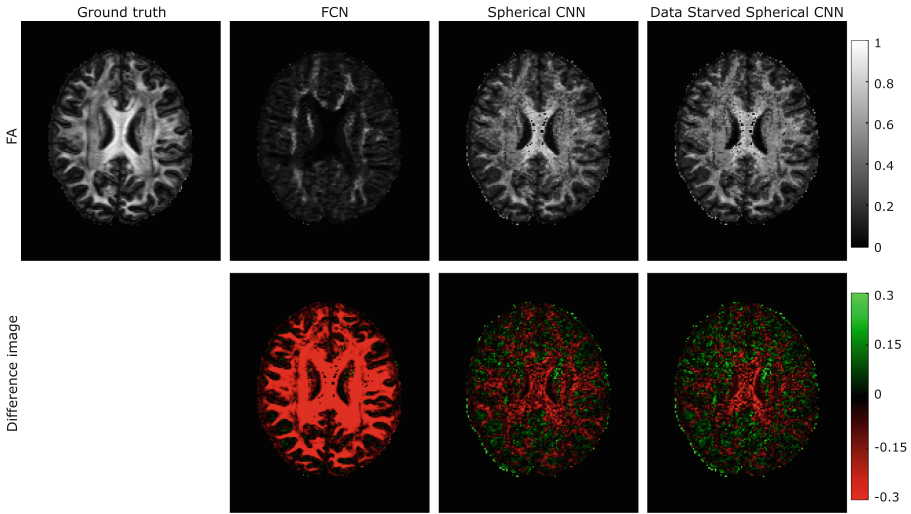
Fig. 4. Comparison between the FCN model and the S-CNN model at estimating the FA when trained only with microstrucure oriented along the anterior-posterior axis. These are also compared against a training data starved S-CNN. We see that spherical CNN outperforms FCN as the noise is greatly reduced and less structured. Data starved S-CNN performs similarly well to the S-CNN trained using the full dataset.
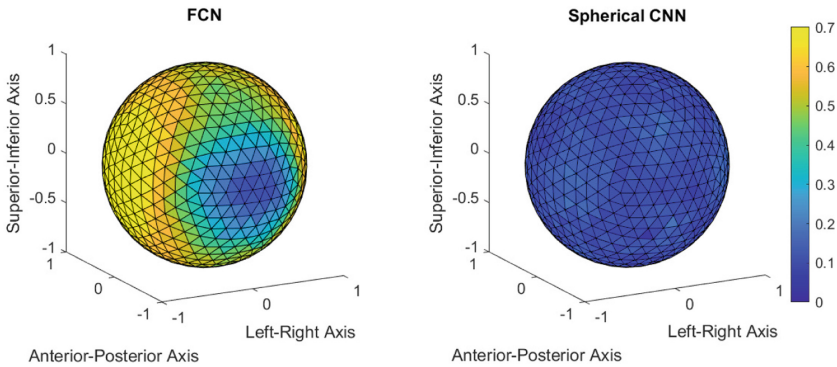
Fig. 5. The distribution of the estimation error over the sphere is compared between the fully-connected network and spherical CNN. To colour each tile we find the voxels whose GT primary fibre orientation lies within the tile's surface and compute the mean RMSE. For this evaluation only voxels with FA > 0.6 were used. We see that FCN's error is only low along the AP axis and high in all other directions. S-CNN's error is low over the entire surface of the sphere.

tation bias in the training dataset to lead to poor estimation for under-sampled directions.

In Figs. 3 and 4 another advantage of rotational equivariance is shown. Performance of the S-CNN network is not greatly changed when only a tenth of the training dataset is used. This is shown in the maps of FA estimated by the data starved S-CNN model in Fig. 4 and the RMSE over the 12 testing subjects shown in Fig. 3. The robustness to a lack of training data is due to S-CNNs being rotationally equivariant. Therefore, during training only a diverse distribution of the microstructure shape is required and not their orientations as well. This property of S-CNNs is a real benefit as it reduces the number of training datapoints required for good estimation at inference stage and, when simulating training data, allows for denser sampling of the shape parameters as SO(3) need not be sampled.

5   Conclusion

In this work we explore the advantages of S-CNNs for dMRI parameter estimation over conventional FCNs. Representing diffusion weighted signals as a spherical image is here demonstrated to gain robustness to the gradient scheme absent from conventional FCNs, at no cost to fidelity. This removes the need to retrain a new network for every gradient scheme, a feature especially beneficial when combining data from multiple sites. S-CNN is shown to be superior to FCN methods additionally because of its rotational equivariance property. This enables the network to encode information about the pattern of the signal irrespective of primary fibre orientation. This eliminates the need to sample diffusion primary fibre orientations, reducing the number of samples needed to cover the full parameter space.

References

1. Aliotta, E., Nourzadeh, H., Sanders, J., Muller, D., Ennis, D.B.: Highly accelerated, model-free diffusion tensor MRI reconstruction using neural networks. Med. Phys. **46**(4), 1581–1591 (2019). https://doi.org/10.1002/mp.13400
2. Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. J. Magn. Resonan. Ser. B **103**(3), 247–254 (1994). https://doi.org/10.1006/jmrb.1994.1037
3. Bodini, B., Ciccarelli, O.: Diffusion MRI in Neurological Disorders. Diffusion MRI: From Quantitative Measurement to In vivo Neuroanatomy, 2nd edn, pp. 241–255 (2014). https://doi.org/10.1016/B978-0-12-396460-1.00011-1
4. Chen, G., et al.: Estimating tissue microstructure with undersampled diffusion data via graph convolutional neural networks. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12267, pp. 280–290. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59728-3_28
5. Cobb, O.J., et al.: Efficient generalized spherical CNNs. In: ICLR 2021 (2021). https://arxiv.org/abs/2010.11661

6. Cohen, T.S., Geiger, M., Köhler, J., Welling, M.: Spherical CNNs. In: ICLR 2018, January 2018. https://arxiv.org/abs/1801.10130

7. Elaldi, A., Dey, N., Kim, H., Gerig, G.: Equivariant spherical deconvolution: learning sparse orientation distribution functions from spherical data. In: Feragen, A., Sommer, S., Schnabel, J., Nielsen, M. (eds.) IPMI 2021. LNCS, vol. 12729, pp. 267–278. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-78191-0_21

8. Golkov, V., et al.: q-Space deep learning: twelve-fold shorter and model-free diffusion MRI scans. IEEE Trans. Med. Imaging **35**(5), 1344–1351 (2016). https://doi.org/10.1109/TMI.2016.2551324

9. Jones, D.K., Horsfield, M.A., Simmons, A.: Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magn. Reson. Med. **42**(3), 515–525 (1999). https://doi.org/10.1002/(SICI)1522-2594(199909)42:3⟨515::AID-MRM14⟩3.0.CO;2-Q

10. Jones, D.K.: The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study. Magn. Reson. Med. **51**(4), 807–815 (2004). https://doi.org/10.1002/mrm.20033

11. Lin, Z., et al.: Fast learning of fiber orientation distribution function for MR tractography using convolutional neural network. Med. Phys. **46**(7), 3101–3116, mp.13555 (2019). https://doi.org/10.1002/mp.13555

12. Park, J., et al.: DIFFnet: diffusion parameter mapping network generalized for input diffusion gradient schemes and b-values. IEEE Trans. Med. Imaging **41**(2), 491–499 (2022). https://doi.org/10.1109/TMI.2021.3116298

13. Sedlar, S., Alimi, A., Papadopoulo, T., Deriche, R., Deslauriers-Gauthier, S.: A spherical convolutional neural network for white matter structure imaging via dMRI. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12903, pp. 529–539. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87199-4_50

14. Skare, S., Hedehus, M., Moseley, M.E., Li, T.Q.: Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. J. Magn. Resonan. **147**(2), 340–352 (2000). https://doi.org/10.1006/jmre.2000.2209