Robust Estimation of the Microstructure of the Early Developing Brain Using Deep Learning

Hamza Kebiri[1,2,3(✉)], Ali Gholipour[3], Rizhong Lin[2,4], Lana Vasung[5], Davood Karimi[3], and Meritxell Bach Cuadra[1,2]

[1] CIBM Center for Biomedical Imaging, Lausanne, Switzerland
hamza.kebiri@unil.ch
[2] Department of Radiology, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland
[3] Computational Radiology Laboratory, Department of Radiology, Boston Children's Hospital and Harvard Medical School, Boston, MA, USA
[4] Signal Processing Laboratory 5 (LTS5), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[5] Department of Pediatrics, Boston Children's Hospital, and Harvard Medical School, Boston, MA, USA

Abstract. Diffusion Magnetic Resonance Imaging (dMRI) is a powerful non-invasive method for studying white matter tracts of the brain. However, accurate microstructure estimation with fiber orientation distribution (FOD) using existing computational methods requires a large number of diffusion measurements. In clinical settings, this is often not possible for neonates and fetuses because of increased acquisition times and subject movements. Therefore, methods that can estimate the FOD from reduced measurements are of high practical utility. Here, we exploited deep learning and trained a neural network to directly map dMRI data acquired with as low as six diffusion directions to FODs for neonates and fetuses. We trained the method using target FODs generated from densely-sampled multiple-shell data with the *multi-shell multi-tissue constrained spherical deconvolution* (MSMT-CSD). Detailed evaluations on independent newborns' test data show that our method achieved estimation accuracy levels on par with the state-of-the-art methods while reducing the number of required measurements by more than an order of magnitude. Qualitative assessments on two out-of-distribution clinical datasets of fetuses and newborns show the consistency of the estimated FODs and hence the cross-site generalizability of the method.

D. Karimi and M. B. Cuadra—Equal contribution.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43990-2_28.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14226, pp. 293–303, 2023.
https://doi.org/10.1007/978-3-031-43990-2_28

1 Introduction

Depiction of white matter fiber tracts is of paramount importance for brain characterization in health and disease. Diffusion-weighted magnetic resonance imaging (dMRI) is the method of choice to study axon bundles that connect different brain regions. Several models have been proposed to map the 4-dimensional diffusion signal to objects such as tensors or fiber orientation distribution functions (FODs) [23,32], which can be further processed to compute metrics such as tract orientation and apparent fiber density [16,26]. Model-based FODs are the mathematical frameworks of choice for microstructure estimation. In fact, FODs accurately describe the underlying microstructure by a deformed sphere in which different radii correspond to different intra-voxel fibers. Moreover, without FODs, tracking stops prematurely and favors shorter fiber tracts [8]. Standard FOD estimation methods [1,15,23,27,32] process voxels individually and thus do not exploit correlations between neighboring voxels. As a result, these methods demand dMRI measurements with multiple b-values and a high number of gradient directions to account for the response function of each tissue type [15].

These acquisitions require prolonged scans that are not affordable for newborn and fetal subjects because of the sensitivity of these cohorts and the increased risk of motion. Acquisitions have to be fast to freeze in-plane motion; yet data dropout rates are high in these cohorts because of motion artifacts. Reconstructing FODs in developing brains has been performed [4,5,7,30] using high-quality datasets and rich information including several gradient directions, higher and/or multiple b-values, and high signal-to-noise ratio (3 T magnetic field strength). Additionally, the datasets were acquired in a controlled and uniform research setting with healthy volunteers, which can hardly be reproduced in the clinical environment. Moreover, and in contrast to adult brains, anisotropy increases in white matter fibers of developing brains because of increased water volume and poor alignment of the fibers [6]. Gray matter on the other hand, during early gestational weeks, is highly anisotropic because of the complexity of the formation of cell bodies, glial cells, and the different neuronal structures [6]. This dynamic period for microstructure [2,9] makes FOD estimation a more challenging task. Adaptive learning-based methods can be leveraged to learn from high-quality datasets and exploit this knowledge in clinical routine acquisitions.

Deep learning models, first suggested in [12], promise to overcome the error accumulation of suboptimal processing steps that are characteristic of standard estimation techniques. This end-to-end learning paradigm has been then applied in dMRI for several purposes [13,18,22,24,25]. The authors in [24] have accurately predicted tensor maps with six diffusion measurements. In [22], a 2D convolutional neural network (CNN) was used to predict the orientation of the fibers in a classification approach whereas [25] deployed a 3D CNN to predict FODs using a small neighborhood of the diffusion signal. In [18], a feedforward neural network was used to predict the FODs and found that 44 directions can be sufficient. However, the network does not exploit neighboring voxels correlations. A more recent work [13] used a Transformer-CNN block to first map 200

to 60 directions and the latter to FODs. However, for uncooperative cohorts such as neonates or fetuses, this number of measurements is unrealistic to acquire.

To the best of our knowledge, no learning-based method to predict FODs has been reported for newborn and fetal brains. In this work, we demonstrate that a deep convolutional neural network with a large field of view (FOV) can accurately estimate FODs using only 6–12 diffusion-weighted measurements. Our contribution is three-fold. We first show that a deep learning method can achieve an accuracy level that is comparable with the agreement between the state-of-the-art methods, while drastically reducing the number of measurements, for developing brains. We then show a low agreement between state-of-the-art methods in terms of different metrics using data from a highly controlled setting, namely the developing Human Connectome Project (dHCP). This addresses the need to build reproducible and reliable pipelines for white matter characterization [29], particularly for the developing brain. Finally, we demonstrate the generalizability of our method on two clinical datasets of fetuses and newborns that were acquired in completely different settings than those used in the training data.

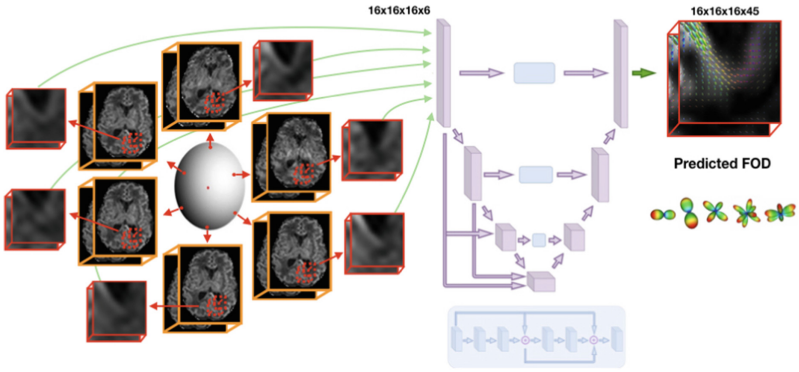

Fig. 1. The proposed framework to predict Fiber Orientation Distribution (FOD) functions in the spherical harmonics domain (SH-L_{max} order 8). The network takes 3D input patches from 6 diffusion measurements and outputs patches of SH coefficients. Network architecture details can be found in Figure S1 of Supplementary materials and code at (https://github.com/Medical-Image-Analysis-Laboratory/Perinatal_fODF_DL_estimation).

2 Methodology

2.1 Paradigm

The method is based on directly learning a mapping between the raw diffusion signal and the FOD in a supervised manner (Fig. 1). Our model inputs are respectively 6 single-shell (b = 1000 s/mm^2) measurements for the neonate network

(DL_n) and 12 single-shell (b = 400 s/mm^2) measurements for the fetal inference network (DL_f), trained on pre-term subjects (as in [17,19]). To be independent on gradient directions, projection of the signal onto spherical harmonics basis (SH) (SH-L_{max} order 2 and 4 respectively for the two networks) was performed to predict the FOD represented in the SH basis (SH-L_{max} order 8). To train the model, these target coefficients are estimated from 300 multi-shell measurements using MSMT-CSD [15]. These measurements are distributed over 3 shells of $\{400, 1000, 2600\}$ s/mm^2 with 64, 88 and 128 samples, respectively, and 20 b0 (b = 0 s/mm^2) images. The few measurements used as input to the model were based on the scheme in [31], whereby the gradient directions minimized the condition number of the diffusion tensor reconstruction matrix.

2.2 Data Processing

dHCP Newborns - We have selected two subsets from the developing Human Connectome Project (dHCP) dataset (1) 100 subjects (weeks: [32.1, 44.7], mean: 40, standard deviation: 2.4) and (2) 68 pre-term subjects (weeks: [29.3, 37.0], mean: 34.9, standard deviation: 1.8). The data was acquired with a 3T Philips Achieva scanner in a multi-shell scheme (b $\in \{0, 400, 1000, 2600\}$ s/mm^2) [14] and was denoised, motion and distortion corrected [3]. It has a final resolution of $1.17\times1.17\times1.5$ mm^3 in a field of view of $128\times128\times64$ voxels. We have upsampled the data to 1 mm isotropic resolution to account for network isotropic 3D patches. We have additionally normalized the input data by b0. A white matter mask was generated using the union of the *White Matter* and the *Brainstem* labels provided by the dHCP, and the voxels where Fractional Anisotropy (FA) was higher than 0.25. A resampling of the dHCP labels from T2-w resolution (0.5 mm^3 isotropic) to 1 mm^3 resolution was performed.

Clinical Newborns and Fetuses - Acquisitions of 8 neonates ([38.1, 39.4, 40.1, 40.4, 40.7, 40.9, 41.8, 42] weeks), were performed during natural sleep at 3T (Siemens Trio and Skyra). Five b0 images and 30 b = 1000 s/mm^2 were acquired. The TR-TE were 3700–104 ms and voxel size was 2 mm isotropic. Eight fetal subjects ([24, 25, 26.3, 26.6, 26.7, 26.9, 29.4, 38.7] gestational weeks, GW) were scanned using a 3T Siemens Skyra MRI scanner (TR = 3000–4000 ms, TE = 60 ms) with one b0 and 12 diffusion-sensitized images at b = 500 s/mm^2. All subjects were processed for noise [34] and bias field inhomogeneities [33]. Rigid registration to a T2 atlas [11] was performed and b-vectors were rotated accordingly for fetal data. The different volumes were upsampled to 1 mm^3 and normalized by b0. The studies were approved by the institutional review board committee.

2.3 Training

Two networks, DL_n and DL_f (see Subsect. 2.1 above), were trained using Adam optimizer [21] to minimize the ℓ_2 norm loss function between the predicted 45 SH coefficients and the ground truth FOD SH coefficients generated using the 300 directions and the 4 b-values ($\{0, 400, 1000, 2600\}$ s/mm^2), i.e.

$$minimize \sum_{i=1}^{45} \left\| FOD_i^{pred} - FOD_i^{GT} \right\|^2$$

We used 70% of the subjects for training, 15% for validation, and 15% for testing. We used the number of FOD peaks (extracted from Dipy [10]) to balance patch selection per batch. The central voxel of each patch was constrained to be in the generated white matter mask and to be 1 peak in $\frac{2}{3}$ of the batch and more than one peak in $\frac{1}{3}$ of the batch. This condition implicitly guarantees the non selection of empty patches. The patch size was empirically set to 16^3 voxels. The batch size was set to 27 for DL_n and 9 for DL_f, and the initial learning rate to 10^{-4} and was decreased by 0.9 whenever the validation loss did not improve after one epoch. The total number of training epochs was 10000 and a dropout rate of 0.1 was used in all layers to reduce overfitting and improve generalization. In DL_f, Gaussian noise injection (mean = 0, sigma = 0.025) was applied as well as small rotations (uniformly from $[-5°, +5°]$) to make the model robust to minor uncorrected movements due to small differences in FOV and fetal head motion.

2.4 Evaluation of dHCP Newborns

Comparison with State-of-the-Art Methods - In addition to comparing our network (DL_n) prediction with FODs estimated using MSMT-CSD of 300 directions (considered as ground truth, GT), we have assessed the agreement between two mutually exclusive subsets extracted from the ground truth (gold standards 1 and 2, respectively GS1 and GS2). Each subset contains 150 directions (b $\in \{0, 400, 1000, 2600\}$ s/mm^2) with respectively 10, 32, 44, and 64 measurements (half measurements of GT data). GS1 and GS2 subsets can be considered as independent high-quality scans, and differences in terms of subsequent metrics can be considered as an upper bound error for the different methods deployed. Furthermore, we have computed three state-of-the-art methods: (1) Constrained spherical deconvolution (CSD) [32] using the 128 gradient directions of the highest shell, i.e. b = 2600 s/mm^2 and 20 b0 images; (2) Constant Solid Angle ODF (Q-Ball) model [1] that we refer to as CSA and (3) the Sparse Fascicle Model (SFM) [27] model for which we have used the default regularization parameters. We also compared our method with the multilayer perceptron (MLP) in [18], which has been shown to outperform the method of [25].

Error Metrics - Quantitative validation was performed based on the number of peaks, the angular error and the apparent fiber density (AFD) [26]. The number of peaks was generated from the FOD predicted by the network and the ones estimated by the different methods (GT, GS1, GS2, CSD, CSA and SFM) using the same parameters (mean separation angle of 45°, a maximum number of 3 peaks and relative peak threshold of 0.5). The conservative choice of these parameters was guided by [28] which shows the limitations of current dMRI models at depicting multiple number of peaks and low angular crossing fibers. We have compared these models in terms of confusion matrices, and the *agreement rate* (AR). AR is defined for each number of peaks p as: $AR = \frac{A_p}{\Sigma D_p}$ where A_p is the percentage of voxels on which both methods agree on p number of peaks

and D_p the percentage of voxels where at least one of the two methods predicts p and the other p' where $p \neq p'$. For the voxels containing the same number of peaks, we have computed the angular error with respect to the GT, as well as between GS1 and GS2. For voxels with multiple fibers, we have first extracted corresponding peaks between the two methods by computing the minimum angle between all configurations (4 for 2 peaks and 9 for 3 peaks); we then removed these peaks and recursively apply the same algorithm. We have also compared AFD, that is defined as the FOD amplitude. AFD was extensively demonstrated as a biologically plausible measure that is not only sensitive to the fiber partial volume fraction but also to fiber density or membrane permeability [26]. Statistical validation using paired t-test corrected for multiple comparisons with Bonferroni method was performed between the errors of the different methods with respect to GT and the difference between GS1 and GS2.

2.5 Evaluation of Clinical Datasets

DL_f was tested on fetal volumes whereas DL_n was tested on the clinical newborn dataset. Due to the lack of ground truth for both clinical datasets, we qualitatively assess the network predictions with 12 and 6 measurements, as compared to CSD using all available measurements (SH-L_{max} order 4 and 8), respectively for fetuses and newborns.

3 Results

The networks consistently learned a mapping between the six/twelve diffusion measurements and the ground truth FOD constructed with 300 measurements across 4 b-values, as evaluated on the independent test data (Figs. 2 and 3).

3.1 In-Domain Quantitative Evaluation in Newborns dHCP

Number of Peaks - We first observe a low agreement (AR) between the two gold standard acquisitions (GS1 vs. GS2), that is more pronounced for multiple fibers voxels. For instance, 1-peaks AR is 80.4%, 30.3% for 2-peaks and 27.9% for 3-peaks. SFM achieves a relatively high 1-peaks agreement with the GT of 83% and the lowest with multiple fibers voxels (10% and 3.5% for 2- and 3-peaks, respectively). In contrast, CSD estimates a high number of multiple fibers (16.5% and 5.9% for 1- and 2-peaks respectively) and achieves the lowest 1-peaks AR with 11.7%. In fact, the latter is biased towards multiple peaks estimation with more than 90% of the voxels modeled as either two or three peaks. This might be explained by the high b-value (b = 2600 s/mm^2) containing high levels of noise. Our method, DL_n, achieves an agreement for 1-, 2- and 3-peaks of respectively 79%, 16% and 3% that is globally the closest to the agreement between the gold standards when compared to other methods. We believe that the relatively low agreement for multiple intravoxel fiber orientations is due to their incongruence across GT subjects, and hence the absence of a

Table 1. Mean angular error, agreement rate on number of peaks and Apparent Fiber Density (AFD) error between GT (MSMT-CSD) and the different methods. ΔGS refers to GS1 and GS2 agreements. The number of measurements (N_m) and the b-values used are also reported. All results were statistically significant compared to ΔGS ($p \leq 9e^{-10}$ for angular error, except SFM three fibers, and $p \leq 4.5e^{-3}$ for AFD error).

Method	b-values (s/mm^2)	N_m	Angular error (Agreement rate in %)			AFD error
			Single fibers	Two fibers	Three fibers	
DL_n	{0, 1000}	7	12.6° (78.4%)	24.2° (15.8%)	33.3° (3.8%)	0.27 (±0.03)
CSD	{0, 2600}	148	7.5° (11.7%)	16.5° (16.5%)	27.2° (5.9%)	1.31 (±0.23)
CSA	{0, 400, 1000, 2600}	300	47.0° (27.7%)	41.4° (14.8%)	36.1° (7.9%)	3.46 (±0.46)
SFM	{0, 400, 1000, 2600}	300	51.4° (83%)	40.7° (10%)	35.4° (3.5%)	0.80 (±0.55)
ΔGS	{0, 400, 1000, 2600}	150	13.8° (80.4%)	29.1° (30.3%)	35.4° (27.9%)	0.2 (±0.025)

consistent pattern to be learned by the neural network. In fact, this is supported by the modest agreement between the two gold standards (ΔGS), in which both the subjects and the number of measurements are the same, only the gradient directions vary and already result in a drop of 70% in multiple fibers depiction. It is worth noting that the agreement between the different methods (CSD vs. CSA, SFM vs. CSA, CSD vs. DL_n, etc.) was also low. The confusion matrices for ΔGS agreement and the different methods can be found in Table 1 and the comparison with [18] in Section 3 of Supplementary materials.

Angular Error - The agreement in terms of the number of peaks does not guarantee that the fibers follow the same orientation. Table 1 shows the angular error and the agreement rate (AR) in numbers of peaks for the different configurations. In GS1 and GS2, the angular difference increases almost linearly for one, two and three fibers. Our learning model achieves an error rate that is comparable (although statistically different, $p \leq 9e^{-10}$) to GS1 and GS2. SFM and CSA achieve a higher error rate for single and two fiber voxels, whereas CSD achieves the lowest. This is because of the low AR and hence the error is computed among a small subset of common voxels between the GT and CSD as shown in Table 1. It is worth mentioning that using 15 directions instead of 6 as input to the network did not improve the results; and in general, these angular errors are higher than those reported for adult data, such as the Human Connectome Project as in [13]. We hypothesize this can be due to immature and high variability of the developing brain anatomy.

Apparent Fiber Density - The last column in Table 1 shows the differences between AFD averaged over the 15 test subjects. Our model achieves the closest error rate of 0.27 (±0.03) to the GT compared with the gold standards difference of 0.2 (±0.025), in terms of mean and standard deviation. The other methods have an increased error rate compared to DL_n with factors of around 2.5, 4.5 and 9.5-fold for SFM, CSD and CSA respectively. Results were statistically significant ($p \leq 4.5e^{-3}$) compared to the agreement between the gold standard models.

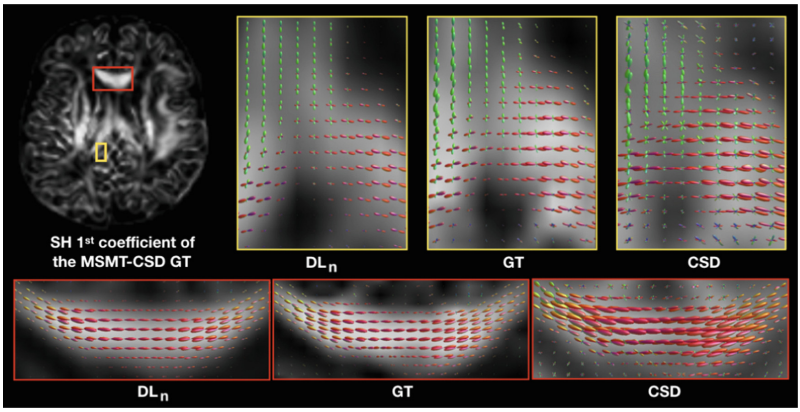

Fig. 2. Qualitative comparison between the deep learning method DL_n, the MSMT-CSD GT and CSD in two brain regions of a newborn dHCP subject.

3.2 Generalizability to Clinical Acquisitions (newborns and fetuses)

DL_f successfully generalized to fetal data as can be seen in Fig. 3 (right) for two subjects. Callossal fibers are clearly delineated on the top and bottom subjects. The radial coherence of cortical plate at early gestation [20] is also highlighted on the same panels. Similarly, DL_n generalized to the new newborn dataset (Fig. 3, left), despite differences in scanner and protocol. Both cortico-spinal tract and corpus callosum are shown in the bottom subject. As opposed to CSD that overestimated false positive crossing fibers, likely due to residual noise, the deep learning method trained on MSMT-CSD directly produced low amplitude FODs in isotropic or non-consistent regions. The results for six other subjects can be found in Figure S2 in Supplementary materials.

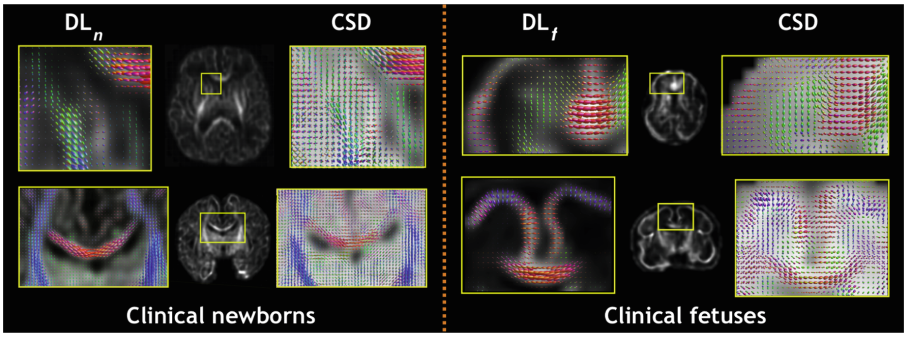

Fig. 3. The deep learning method compared to CSD in different brain regions for 2 newborn subjects (left) and 2 fetal subjects (right) of 25 (top) and 29.4 (bottom) weeks of gestation. FODs are superimposed to the first SH coefficient of the method used.

4 Conclusion

We have demonstrated how a deep neural network can successfully reconstruct high angular multi-shell FODs from a reduced number (6 to 12) of diffusion measurements. The substantially lower number of samples is compensated by learning from high-quality training data and by exploiting the spatial neighborhood information. The network was quantitatively evaluated on the dHCP dataset which was acquired in a highly controlled setting that cannot be reproduced in clinical settings. We showed that our method relying on six measurements can be leveraged to reconstruct plausible FODs of clinical newborn and fetal brains. We compared our model to commonly used methods such as CSD and MSMT-CSD between two gold standard datasets. The results exhibit low agreements between the different methods, particularly for multiple fiber orientations, despite using high angular multi-shell data. This highlights the need to build robust and reproducible methods for microstructure estimation in developing brains.

Acknowledgment. This work was supported by the Swiss National Science Foundation (project 205321-182602). We acknowledge the CIBM Center for Biomedical Imaging, a Swiss research center of excellence founded and supported by CHUV, UNIL, EPFL, UNIGE, HUG and the Leenaards and Jeantet Foundations. This research was also partly supported by the US National Institutes of Health (NIH) under awards R01NS106030 and R01EB032366; by the Office of the Director of the NIH under award S10OD0250111; and by NVIDIA Corporation; and utilized an NVIDIA RTX A6000 GPU.

References

1. Aganj, I., et al.: Reconstruction of the orientation distribution function in single- and multiple-shell q-ball imaging within constant solid angle. Magn. Reson. Med. **64**(2), 554–566 (2010)
2. Andescavage, N.N., et al.: Complex trajectories of brain development in the healthy human fetus. Cereb. Cortex **27**(11), 5274–5283 (2017)
3. Bastiani, M., et al.: Automated processing pipeline for neonatal diffusion MRI in the developing human connectome project. Neuroimage **185**, 750–763 (2019)
4. Chen, R., et al.: Deciphering the developmental order and microstructural patterns of early white matter pathways in a diffusion MRI based fetal brain atlas. Neuroimage **264**, 119700 (2022)
5. Christiaens, D., et al.: Multi-shell shard reconstruction from scattered slice diffusion MRI data in the neonatal brain. ISMRM (Paris) (2018)
6. Counsell, S.J., Arichi, T., Arulkumaran, S., Rutherford, M.A.: Fetal and neonatal neuroimaging. Handb. Clin. Neurol. **162**, 67–103 (2019)
7. Deprez, M., et al.: Higher order spherical harmonics reconstruction of fetal diffusion MRI with intensity correction. IEEE Trans. Med. Imaging **39**(4), 1104–1113 (2019)
8. Descoteaux, M.: High angular resolution diffusion MRI: from local estimation to segmentation and tractography. Ph.D. thesis, Univ. Nice Sophia Antipolis (2008)
9. Dubois, J., et al.: The early development of brain white matter: a review of imaging studies in fetuses, newborns and infants. Neuroscience **276**, 48–71 (2014)

10. Garyfallidis, E., et al.: Dipy, a library for the analysis of diffusion MRI data. Front. Neuroinform. **8**, 8 (2014)
11. Gholipour, A., et al.: A normative spatiotemporal MRI atlas of the fetal brain for automatic segmentation and analysis of early brain growth. Sci. Rep. **7**(1), 476 (2017)
12. Golkov, V., et al.: Q-space deep learning: twelve-fold shorter and model-free diffusion MRI scans. IEEE TMI **35**(5), 1344–1351 (2016)
13. Hosseini, S., et al.: CTtrack: a CNN+ transformer-based framework for fiber orientation estimation & tractography. Neurosci. Inform. **2**(4), 100099 (2022)
14. Hutter, J., et al.: Time-efficient and flexible design of optimized multishell HARDI diffusion. Magn. Reson. Med. **79**(3), 1276–1292 (2018)
15. Jeurissen, B., et al.: Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. Neuroimage **103**, 411–426 (2014)
16. Jeurissen, B., et al.: Diffusion MRI fiber tractography of the brain. NMR Biomed. **32**(4), e3785 (2019)
17. Karimi, D., et al.: Deep learning-based parameter estimation in fetal diffusion-weighted MRI. Neuroimage **243**, 118482 (2021)
18. Karimi, D., et al.: Learning to estimate the fiber orientation distribution function from diffusion-weighted MRI. Neuroimage **239**, 118316 (2021)
19. Kebiri, H., et al.: Slice estimation in diffusion MRI of neonatal and fetal brains in image and spherical harmonics domains using autoencoders. In: Cetin-Karayumak, S., et al. (eds.) CDMRI 2022. LNCS, vol. 13722, pp. 3–13. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-21206-2_1
20. Khan, S., et al.: Fetal brain growth portrayed by a spatiotemporal diffusion tensor MRI atlas computed from in utero images. Neuroimage **185**, 593–608 (2019)
21. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
22. Koppers, S., Merhof, D.: Direct estimation of fiber orientations using deep learning in diffusion imaging. In: Wang, L., Adeli, E., Wang, Q., Shi, Y., Suk, H.-I. (eds.) MLMI 2016. LNCS, vol. 10019, pp. 53–60. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47157-0_7
23. Le Bihan, D., et al.: Diffusion tensor imaging: concepts and applications. J. Magn. Reson. Imaging **13**(4), 534–546 (2001)
24. Li, H., et al.: SuperDTI: ultrafast DTI and fiber tractography with deep learning. Magn. Reson. Med. **86**(6), 3334–3347 (2021)
25. Lin, Z., et al.: Fast learning of fiber orientation distribution function for MR tractography using convolutional neural network. Med. Phys. **46**(7), 3101–3116 (2019)
26. Raffelt, D., et al.: Apparent fibre density: a novel measure for the analysis of diffusion-weighted magnetic resonance images. Neuroimage **59**(4), 3976–3994 (2012)
27. Rokem, A., et al.: Evaluating the accuracy of diffusion MRI models in white matter. PLoS ONE **10**(4), e0123272 (2015)
28. Schilling, K.G., et al.: Histological validation of diffusion MRI fiber orientation distributions and dispersion. Neuroimage **165**, 200–221 (2018)
29. Schilling, K.G., et al.: Prevalence of white matter pathways coming into a single white matter voxel orientation: the bottleneck issue in tractography. Hum. Brain Mapp. **43**(4), 1196–1213 (2022)
30. Shen, K., et al.: A spatio-temporal atlas of neonatal diffusion MRI based on kernel ridge regression. In: International Symposium on Biomedical Imaging (ISBI) (2017)
31. Skare, S., et al.: Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. J. Magn. Reson. **147**(2), 340–352 (2000)

32. Tournier, J.D., et al.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. Neuroimage **23**(3), 1176–1185 (2004)
33. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., et al.: N4ITK: improved N3 bias correction. IEEE Trans. Med. Imaging **29**(6), 1310–1320 (2010)
34. Veraart, J., Fieremans, E., Novikov, D.S.: Diffusion MRI noise mapping using random matrix theory. Magn. Reson. Med. **76**(5), 1582–1593 (2016)