LEARNING SPATIALLY-CONTINUOUS FIBER ORIENTATION FUNCTIONS

Tyler Spears[1] *P. Thomas Fletcher*[1]

[1] University of Virginia, Department of Electrical and Computer Engineering, Charlottesville, VA

ABSTRACT

Our understanding of the human connectome is fundamentally limited by the resolution of diffusion MR images. Reconstructing a connectome's constituent neural pathways with tractography requires following a continuous field of fiber directions. Typically, this field is found with simple trilinear interpolation in low-resolution, noisy diffusion MRIs. However, trilinear interpolation struggles following fine-scale changes in low-quality data. Recent deep learning methods in super-resolving diffusion MRIs have focused on upsampling to a fixed spatial grid, but this does not satisfy tractography's need for a continuous field. In this work, we propose FENRI, a novel method that learns spatially-continuous fiber orientation distribution functions for accurate tractography from low-resolution diffusion-weighted images. To quantify FENRI's capabilities in tractography, we also introduce an expanded simulated dataset built for evaluating learning-based tractography models. We demonstrate that FENRI accurately predicts fiber orientations from realistic low-quality data, and that FENRI tractography gives improved streamline reconstruction over the current use of simple interpolation. [1].

***Index Terms*—** Diffusion tensor imaging, Superresolution, Tractography, Implicit neural networks

1. INTRODUCTION

Mapping the human connectome relies upon a continuous and accurate representation of the underlying brain tissue. This is needed for tracing streamlines, resolving crossing fibers, and deciding when to terminate a tract. Often, tractography algorithms rely on simple trilinear interpolation to "fill out" a continuous field from discretely-sampled diffusion magnetic resonance images (dMRIs). If this interpolation could be improved, then tractography algorithms could produce more detailed and accurate human white matter (WM) fiber tracts.

In this work, we propose FENRI (**F**iber orientations from **E**xplicit **N**eural **R**epresentat**I**ons), a novel deep learning-based super-resolution model for estimating fiber orientation distribution functions (fODFs) continuously in space. We demonstrate FENRI's capabilities through the following experiments: 1) a quantitative evaluation of fODF reconstruction in Human Connectome Project (HCP) data, 2) a qualitative evaluation of tractography in HCP data, and 3) a quantitative measure of tractography performance on a new, expanded simulation dataset. As an image upsampler, FENRI outperforms more generic single-image super-resolution (SISR) methods on a variety of metrics. We also show how, as a tractography enhancement, FENRI's explicit representation sampling provides a powerful improvement over standard tractography methods.

Background. Reconstructing streamlines from diffusion-weighted images (DWIs) requires a model of neuron fiber directionality. One popular model is the general fODF represented by coefficients in the spherical harmonic (SH) orthonormal basis, estimated by constrained spherical deconvolution (CSD) [1]. Several deep learning models have recently been proposed to super-sample diffusion representations. For example, Qin et. al., 2021 used convolutional neural networks (CNNs), an efficient sub-pixel CNN (ESPCN) layer, and high-resolution T1w volumes to predict high-resolution diffusion model parameters [2, 3]. However, these previous works were limited to upsampling by an *integer upscaling factor*, e.g., $2\times$, which is not ideal for estimating continuous fields. The recently proposed implicit neural representation (INR) method, which learns continuous-valued representations in some Euclidean space, is one solution to this challenge [4]. INRs are most commonly applied to 3D rendering, but INR-like models have been used in SISR. For example, the Local Implicit Image Function where a low-resolution input image is encoded into a feature space and sampled continuously for upsampling [5]. To our knowledge, the only proposed model that utilizes INRs for super-resolving dMRIs is given in [6], which focused on uncertainty in continuous predictions rather than tractography. We place FENRI alongside these INR models, but note that FENRI does not model an *implicit* function, but an *explicit* function of SH coefficients.

2. METHODS

Our objective is to **predict fODFs at arbitrary, continuous spatial coordinates** given only low-resolution DWIs. We have a set of subject DWIs $\mathbf{S}$ with a continuous space $\mathbf{\Omega} \subset \mathbb{R}^3$. Now, $\mathbf{S}$ is sampled on a discrete, finite, rectilinear grid $\mathbf{P} \subset \mathbf{\Omega}$ at coordinates $p \in \mathbf{P}$. Given a *query coordinate* $q \in \mathbf{\Omega}$, we wish to find the vector of SH coefficients d_q that represents the fiber orientations at q. Thus, we construct a

[1] Source code and data can be found at `https://osf.io/dvnxw/`

function G_θ, with parameters θ, such that $G_\theta(\mathbf{S}, \mathbf{P}, q) = d_q$.

Encoder. The encoder compresses spatial and angular information found in the input DWIs, described as $E : (\mathbf{S}, \mathbf{P}) \rightarrow \mathbf{L}$, where $\mathbf{L}$ has feature vectors of length $c_\mathbf{L}$, and the spatial ordering of feature vectors in $\mathbf{L}$ is assumed to be equivalent to that of $\mathbf{P}$. We chose a 3D implementation of the Cascading Residual Network (CARN), a high-performing SISR model, as the core of our encoder [7]. We use ReLU activation throughout the encoder, and also add a kernel size 2 average-pooling layer and a single batch-norm layer at the decoder's output to improve performance and stability [8].

Continuous Decoder. The decoder predicts SH coefficients at any given query point $q \in \mathbf{\Omega}$ based on encoded features and the query coordinate itself, described by $D : (\mathbf{L}, \mathbf{P}, q) \rightarrow d_q$. The decoder at its core is a simple fully-connected network with n_D hidden layers [5] and the SiLU activation function. Every query point q requires 8 forward passes through the same decoder. The features in $\mathbf{L}$ (the "local ensemble" [5]) given to the decoder correspond to the $2\times2\times2$ spatial nearest-neighbors to q as determined by the input DWI grid $\mathbf{P}$. Each feature vector in this ensemble is given a forward pass through the decoder, and all 8 outputs are trilinearly weighted as the final step in the model prediction.

Here we describe a single forward pass of the decoder for a query q and sub-grid point $\mathbf{L}_i \in \mathbf{L}$, the i'th input vector local to q, and its input coordinate p_i. The first layer takes a concatenation of $\mathbf{L}_i, p_i, q$, and the Fourier positional encoding of $p_i - q$. We normalize $p_i - q$ to $[0, 1)$ (denoted $\mathrm{N}_\mathbf{P}$) and use the encoding map $\mathcal{F}$ (with the number of frequencies m) proposed as the "positional encoding" method in Tancik et. al., 2020 [9]. The full input size is $c_\mathbf{L} + 6 + 6m$, and the output is an SH coefficient vector.

3. EXPERIMENTS

Comparison Models. We compare FENRI to a variety of upsampling and tractography methods. The baseline method, which we label as "Trilin-DWI," is a trilinear upsampling of the noisy and low-resolution DWIs into the target spatial resolution. SH coefficients are then estimated from these upsampled DWIs with multi-shell, multi-tissue CSD (MSMT-CSD) [1]. We also evaluated a network that used the more common SISR ESPCN layer, which we named Fixed-Net (as in "fixed-size upsampling network"). Fixed-Net utilized a similarly parameterized CARN encoder to that used by FENRI, allowing Fixed-Net to also serve as a rough FENRI ablation model. Fixed-Net's encoder (by way of its ESPCN layer) upsampled low-resolution DWIs by $2\times$ the input spatial resolution, then used trilinear interpolation to resample its latent space into the target spatial resolution. The resampled latent space is then passed through a smaller CARN-style network to refine the trilinear resampling, producing an upsampled SH coefficient volume. For tractography, FENRI continuously samples its latent space $\mathbf{L}$ for SH coefficients at every tracking iteration, while all other models used trilinear sampling on predicted SH coefficients. The streamline tracking method was a deterministic gradient ascent-based tractography algorithm implemented for use on a graphics processing unit (GPU). We chose tracking parameters to closely match the defaults for the "SD Stream" tractography in MRtrix3 [10].

Experiment 1: HCP ODF Reconstruction. We tested all models on voxel-wise fODF reconstruction of the HCP Young Adult dataset [11]. All HCP DWIs were preprocessed with the standard HCP pipeline and normalized by b0 intensities [12, 10]. We degraded the DWIs to make them more clinically-realistic by: 1) angularly resampling gradient directions to the first 9 b0's, 45 b1000's and 45 b3000's of the HCP diffusion protocol, 2) downsampling from 1.25mm to 2.00mm, and 3) adding Rician-distributed noise to a 30 dB signal-to-noise ratio. Each model predicted a 1.25mm isotropic volume of even-degree SH coefficients with $l_{max} = 8$ from the degraded DWIs. Ground truth ODFs were estimated from 1.25mm DWIs with multi-shell multi-tissue CSD (MSMT-CSD) and normalization [10, 1, 13]. We evaluated models over three metrics: a weighted mean-squared error (WMSE) of the SH coefficients, the mean-squared Jensen-Shannon Distance (MSJSD) of the ODF, and the weighted average angular error (WAAE) of the ODF peaks. The WMSE is simple MSE with each degree l scaled to a standard normal distribution; this is also the loss function for FENRI and Fixed-Net. The MSJSD calculates the JS Distance between the predicted and target ODFs with density functions estimated over a discrete set of directions [14]. Finally, the WAAE compares ground truth ODF peaks to the nearest peaks in the predicted ODF [15]. The angular distance was minimized between (at most) the two largest target peaks and (at most) the three largest prediction peaks, with a minimum peak of 0.1. False positive and false negative peaks were penalized by $0.0073\pi/2$, the empirically-estimated normalized median peak amplitude times the maximum angular distance. FENRI and Fixed-Net were trained on 35 HCP subjects and tested on 29 subjects, repeated three times.

Experiment 2: HCP Qualitative Tractography. We compared FENRI-based tractography with trilinear interpolation tractography on a real human subject. No method exists for directly measuring a ground truth tractogram in real, *in vivo* HCP data, so our comparison was qualitative in nature. We input degraded DWIs (similar to Experiment 1) to both FENRI and Trilin-DWI and bilaterally reconstructed two tract bundles, chosen for their recognition in the literature - the cortico-spinal tract (CST) and the uncinate fasciculus (UF).

Experiment 3: ISMRM-sim Tractography. In the recently updated ISMRM-2015 tractography challenge dataset, expert-curated WM bundles formed the ground truth of a simulated DWI dataset [16]. We extend the original dataset from one to 15 brain-like simulacrums to allow for train-test splitting. We used symmetric image normalization registration (SyN) [17] to match the original dataset's simulated T1w to

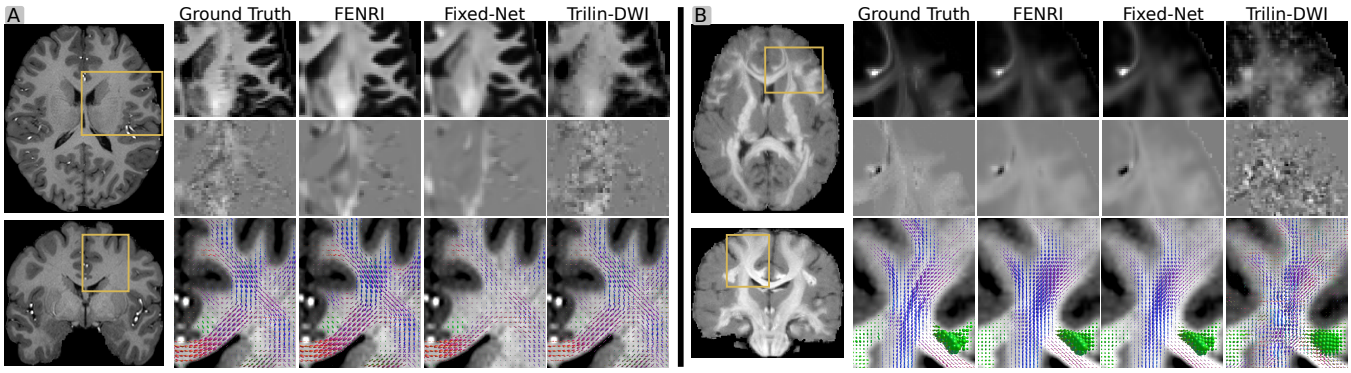

Fig. 1. Predicted SH coefficients on real HCP data and synthetic ISMRM-sim data. (A) Prediction examples on an HCP subject, located at the yellow selection box on the subject's T1w image (left). Rows 1 and 2 are axial slices of SH indices of degree 0 order 0, and degree 6 order -1, respectively. Row 3 contains coronal slices of each methods's ODF lobe plot, with lobes scaled for visual clarity. (B) Prediction examples on a synthetic ISMRM-sim subject. Image parameters match those in (A).

15 real HCP T1ws via the MNI-152 template, and warped the challenge's subjects WM bundles to each of the 15 target subjects. An example warped T1w image is in Figure 1B. This allowed the dataset to maintain biological variability alongside the expert-curated ground truth bundles. Then, each target subject's warped bundles were used to create simulated 0.9 mm DWIs, with 4 b0's, 90 b1000's, and 90 b3000's (gradient directions match the standard HCP sequence) via Fiberfox simulation [11, 18]. These data are normalized and degraded similarly to the DWIs in Experiment 1; we name this dataset "ISMRM-sim." Thus, FENRI and Fixed-Net must learn to upsample 2.0 mm degraded DWIs to 0.9 mm SH coefficients. Examples of ISMRM-sim SH coefficients are shown in Figure 1B. We aimed to measure how accurately each model could reconstruct the bundles that defined the ISMRM-sim DWIs. Tractography was seeded at the midpoint of each ground truth streamline in the following bundles: brainstem projection system (BPS), corpus callosum temporal (CC-t), corpus callosum u-shaped (CC-u), cingulum bundle (Cing), optic radiation (OR), inferior longitudinal fasciculus (ILF), superior longitudinal fasciculus (SLF), and uncinate fasciculus (UF). Performance was evaluated with the challenge's tractogram rating script on each separate bundle, giving the same metrics used in the challenge: overlap voxel ratio (OL), overreach voxel ratio normalized by bundle volume (OR), and segmentation f1 (Dice) score. Networks were trained on 9 ISMRM-sim subjects, and all methods were tested on 5 subjects.

4. RESULTS & DISCUSSION

Experiment 1: HCP ODF Reconstruction. The quantitative results for predicting HCP SH coefficients are given in Figure 2, and example predictions are illustrated in Figure 1A. Across all three metrics (WMSE, MSJSD, and WAAE), FENRI outperformed both Trilinear-DWI and Fixed-Net, while also showing equal or lower variance between test set

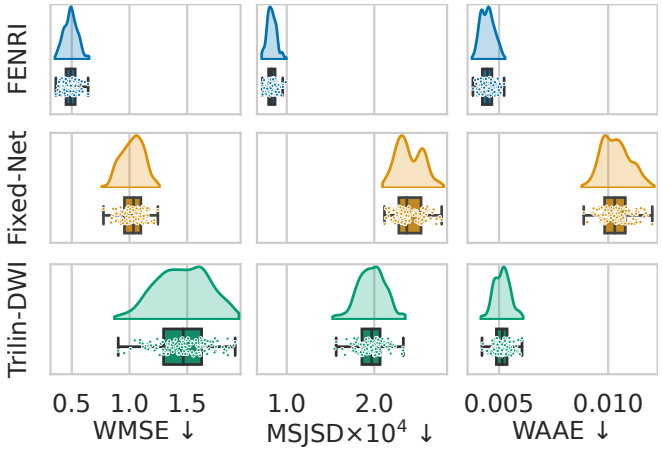

Fig. 2. Results of predicting HCP SH coefficients from low-resolution DWIs in white matter voxels. Each point is a score for a single HCP subject; the distribution plots summarize the point data. Arrows indicate the direction of better performance. WMSE: weighted mean squared error; MSJSD: mean-squared Jensen-Shannon distance; WAAE: weighted average angular error.

subjects. As shown in Figure 1A, FENRI gives high-quality predictions that preserve high frequency details, both in the spatial and angular sense. Against Fixed-Net, FENRI better reconstructs the high degree SH coefficients and produces an overall more accurate ODF, despite sharing an encoder architecture with Fixed-Net. While Trilin-DWI is impressive as a baseline, FENRI still gives quantitatively better predictions. One shortcoming of Trilin-DWI can be seen at the gray matter-WM boundary in Figure 1A, row 3, where Trilin-DWI cannot maintain edges over sharp turns in the gyri; FENRI, however, better preserves these boundaries. Additionally, Fixed-Net's performance is surprising compared to Trilin-

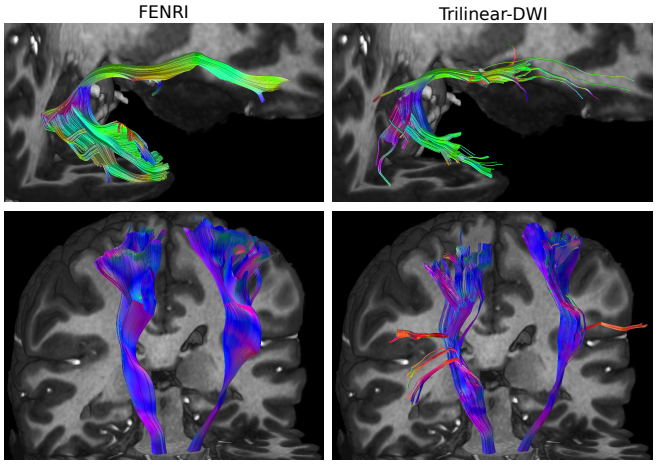

Fig. 3. Qualitative results of both trilinear and FENRI tractography on an HCP subject between FENRI and Trilin-DWI. Row 1 shows an example unilateral uncinate fasciculus, and row 2 shows the left and right cortico-spinal tracts.

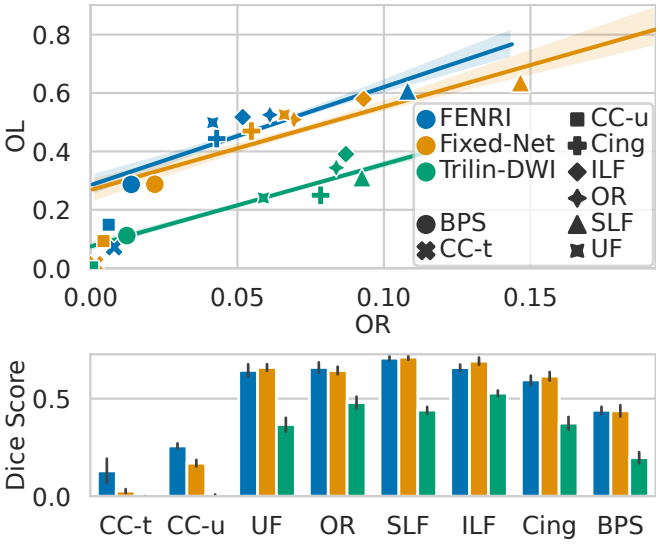

Fig. 4. Quantitative tractography results on ISMRM-sim. *(Top)* overreach (OR) vs. overlap (OL) of each tractography method over selected bundles. Points represent method and bundle score averaged over all subjects. Regression lines are fit to each method's scores, shaded on the 95% confidence interval. *(Bottom)* Dice scores for each method and bundle. Colors match the top row, and error bars represent the 95% confidence interval. See Section 3 for the full bundle names.

DWI. Fixed-Net outperformed Trilin-DWI on WMSE, the network's optimized loss function. However, Trilin-DWI outperforms Fixed-Net on both MSJSD and WAAE. We hypothesize that this is caused by Fixed-Net overfitting to the objective function, while struggling to reconstruct high-degree SH coefficients on the real HCP data. This can be seen in Figure 1A, where Fixed-Net's predicted degree 6 coefficient is relatively sparse. However, this weakness of Fixed-Net does not seem as limiting in the more homogeneous ISMRM-sim dataset, as shown in Figure 1B. Overall, these Fixed-Net results suggest that more common SISR methods may not be appropriate for predicting and upsampling ODFs on real data.

Experiment 2: HCP Qualitative Tractography. The qualitative results of both trilinear and FENRI tractography are shown in Figure 3. Tracking was poor when interpolating on Fixed-Net predictions of HCP data and were omitted for brevity. This poor tracking becomes clear when seeing the Fixed-Net predicted lobe plots shown in the final row of Figure 1A. When comparing FENRI tractography to trilinear DWI upsampling+tractography, FENRI generally produces smoother, more filled tracts when given the same seeds and tracking parameters. Row 1 in Figure 3 illustrates an example UF where FENRI-based tractography produces a smooth and dense tractogram when compared to Trilin-DWI. A similar result is found in reconstructed CST. Seed points were only given in the mid-Pons area of the brainstem, so all produced tracts must nearly traverse the entire brain longitudinally, maximizing the chance of integration errors while tracking.

Experiment 3: ISMRM-sim Tractography. Figure 4 shows the quantitative results of ISMRM-sim data tractography. Figure 1B illustrates that Trilin-DWI is challenged by the sparse, thin features of the simulation. These thin streamlines are obscured by downsampling and noise injection, so tractography becomes difficult without learned priors, such as those learned by FENRI and Fixed-Net. Comparing the learned networks, FENRI matches or exceeds Fixed-Net in most given metrics. As shown in Figure 4A, FENRI produces an equal or higher OL on all tested bundles, while having a generally lower OR, particularly in the SLF, ILF, and the Cingulum. In Figure 4B, FENRI typically matches or outperforms Fixed-Net's Dice score, especially on the CC bundles. However, Fixed-Net does outperform FENRI in the ILF, and matches FENRI in most other bundles. We hypothesize that the selected bundles are traversable by trilinear interpolation given a sufficiently high-resolution ODF. This warrants further analysis on other bundles, and gives opportunities for improving FENRI. Overall, these results suggest that using FENRI in tractography can produce more accurate tractograms, particularly in smaller WM bundles.

Discussion. We have proposed and evaluated FENRI, a first-of-its-kind explicit neural representation model built for tractography. We have shown that FENRI performs well in predicting and upsampling fODFs on a continuum of spatial resolutions when given degraded, clinical-quality DWIs. We have also shown that FENRI can be effectively used in tractography and holds great potential for improving a variety of tractography-focused methods. Finally, we have expanded the ISMRM 2015 challenge dataset to 15 HCP subjects and made these data publicly available for future works in quantitative evaluation of data-driven models in tractography.

Compliance with Ethical Standards This research was conducted using the publicly released HCP Young Adult human dataset. Ethical approval was not required.

Acknowledgments This work was partially supported by NSF Smart and Connected Health grant 2205417 and NIH NINDS grant 5UH3NS123308-02. Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

5. REFERENCES

[1] Ben Jeurissen, Jacques-Donald Tournier, Thijs Dhollander, Alan Connelly, and Jan Sijbers, "Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data," *NeuroImage*, vol. 103, pp. 411–426, Dec. 2014.

[2] Yu Qin, Yuxing Li, Zhizheng Zhuo, Zhiwen Liu, Yaou Liu, and Chuyang Ye, "Multimodal Super-Resolved q-Space Deep Learning," *Medical Image Analysis*, Apr. 2021.

[3] Wenzhe Shi, Jose Caballero, Ferenc Huszar, Johannes Totz, Andrew P. Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang, "Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network," in *CVPR 2016*, 2016.

[4] Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein, "Implicit Neural Representations with Periodic Activation Functions," June 2020.

[5] Yinbo Chen, Sifei Liu, and Xiaolong Wang, "Learning Continuous Image Representation with Local Implicit Image Function," Apr. 2021.

[6] William Consagra, Lipeng Ning, and Yogesh Rathi, "Neural Orientation Distribution Fields for Estimation and Uncertainty Quantification in Diffusion MRI," July 2023.

[7] Namhyuk Ahn, Byungkon Kang, and Kyung-Ah Sohn, "Fast, Accurate, and Lightweight Super-Resolution with Cascading Residual Network," in *ECCV 2018*, 2018.

[8] Yusuke Sugawara, Sayaka Shiota, and Hitoshi Kiya, "Super-Resolution Using Convolutional Neural Networks Without Any Checkerboard Artifacts," in *Proc. ICIP 2018*, Oct. 2018.

[9] Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng, "Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains," June 2020.

[10] J-Donald Tournier, Robert Smith, David Raffelt, Rami Tabbara, Thijs Dhollander, Maximilian Pietsch, Daan Christiaens, Ben Jeurissen, Chun-Hung Yeh, and Alan Connelly, "MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation," *NeuroImage*, vol. 202, Nov. 2019.

[11] David C. Van Essen, Stephen M. Smith, Deanna M. Barch, Timothy E. J. Behrens, Essa Yacoub, and Kamil Ugurbil, "The WU-Minn Human Connectome Project: An overview," *NeuroImage*, vol. 80, pp. 62–79, Oct. 2013.

[12] Matthew F. Glasser, Stamatios N. Sotiropoulos, J. Anthony Wilson, Timothy S. Coalson, Bruce Fischl, Jesper L. Andersson, Junqian Xu, Saad Jbabdi, Matthew Webster, Jonathan R. Polimeni, David C. Van Essen, and Mark Jenkinson, "The minimal preprocessing pipelines for the Human Connectome Project," *NeuroImage*, vol. 80, pp. 105–124, Oct. 2013.

[13] Thijs Dhollander, Rami Tabbara, Jonas Rosnarho-Tornstrand, Jacques-Donald Tournier, David Raffelt, and Alan Connelly, "Multi-tissue log-domain intensity and inhomogeneity normalisation for quantitative apparent fibre density," in *Proc. ISMRM*, 2021.

[14] Davood Karimi, Lana Vasung, Camilo Jaimes, Fedel Machado-Rivas, Simon K. Warfield, and Ali Gholipour, "Learning to estimate the fiber orientation distribution function from diffusion-weighted MRI," *NeuroImage*, vol. 239, Oct. 2021.

[15] Thomas Schultz, "Learning a Reliable Estimate of the Number of Fiber Directions in Diffusion MRI," in *MICCAI 2012*, 2012.

[16] Emmanuelle Renauld, Antoine Théberge, Laurent Petit, Jean-Christophe Houde, and Maxime Descoteaux, "Validate your white matter tractography algorithms with a reappraised ISMRM 2015 Tractography Challenge scoring system," *Sci Rep*, vol. 13, no. 1, pp. 1–11, Feb. 2023.

[17] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Medical Image Analysis*, vol. 12, no. 1, pp. 26–41, Feb. 2008.

[18] Peter F. Neher, Frederik B. Laun, Bram Stieltjes, and Klaus H. Maier-Hein, "Fiberfox: Facilitating the creation of realistic white matter software phantoms," *Magn. Reson. Med.*, vol. 72, no. 5, pp. 1460–1470, 2014.